

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ



**ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ
ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ**

**ΜΕΤΑ-ΑΝΑΛΥΣΗ:
ΣΤΑΤΙΣΤΙΚΗ ΜΕΘΟΔΟΛΟΓΙΑ
ΚΑΙ ΘΕΜΕΛΙΩΣΗ**

Αναστασία Γ. Ελευθεράκη

Διπλωματική Εργασία

που υποβλήθηκε στο Τμήμα Στατιστικής και Ασφαλιστικής
Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των
απαιτήσεων για την απόκτηση του Μεταπτυχιακού
Διπλώματος Ειδίκευσης στην Εφαρμοσμένη Στατιστική

Πειραιάς
Ιούνιος 2004

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ



**ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ
ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ**

**ΜΕΤΑ-ΑΝΑΛΥΣΗ:
ΣΤΑΤΙΣΤΙΚΗ ΜΕΘΟΔΟΛΟΓΙΑ
ΚΑΙ ΘΕΜΕΛΙΩΣΗ**

Αναστασία Γ. Ελευθεράκη

Διπλωματική Εργασία

που υποβλήθηκε στο Τμήμα Στατιστικής και Ασφαλιστικής
Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των
απαιτήσεων για την απόκτηση του Μεταπτυχιακού
Διπλώματος Ειδίκευσης στην Εφαρμοσμένη Στατιστική

Πειραιάς
Ιούνιος 2004

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίστηκε από τη ΓΣΕΣ του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ' αριθμ. συνεδρίασή του σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών στην Εφαρμοσμένη Στατιστική

Τα μέλη της Επιτροπής ήταν:

- Τ. Παπαϊωάννου (Επιβλέπων)
- Μ. Κούτρας
- Μ. Κατέρη

Η έγκριση της Διπλωματικής Εργασίας από το Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς δεν υποδηλώνει αποδοχή των γνώμων του συγγραφέα.

UNIVERSITY OF PIRAEUS



**DEPARTMENT OF STATISTICS
AND INSURANCE SCIENCE**

**POSTGRADUATE PROGRAM IN
APPLIED STATISTICS**

**META-ANALYSIS:
STATISTICAL METHODOLOGY AND
FOUNDATION**

By

Anastasia G. Eleftheraki

MSc Dissertation

submitted to the Department of Statistics and Insurance
Science of the University of Piraeus in partial fulfillment of
the requirements for the degree of Master of Science in
Applied Statistics

Piraeus, Greece
June 2004

*Στους γονείς μου
& στην αδερφή μου*

Ευχαριστίες

Αρκετοί ήταν αυτοί που συνέβαλλαν με τον ένα ή με τον άλλο τρόπο στην πραγματοποίηση της παρούσας διπλωματικής εργασίας και νιώθω την ανάγκη να τους ευχαριστήσω.

Καταρχήν ευχαριστώ θερμά τον κ. Τάκη Παπαϊωάννου, Καθηγητή του τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιά για την συνεχή επίβλεψη και καθοδήγηση στη διαμόρφωση της διπλωματικής εργασίας. Επίσης ευχαριστώ τους δύο καθηγητές που συμπληρώνουν τη τριμελή επιτροπή, την κα. Μαρία Κατέρη, Επίκουρος Καθηγήτρια του τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιά και τον κ. Μάρκο Κούτρα, Καθηγητή του τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιά, γιατί με τις παρατηρήσεις τους βοήθησαν στην ολοκλήρωση της παρούσας εργασίας. Επιπλέον θα ήθελα να ευχαριστήσω τον κ. Marlos Viana, Καθηγητή του Πανεπιστημίου του Ιλινόις για τις σκέψεις και τους προβληματισμούς του σχετικά με την εφαρμογή της μετανάλυσης.

Τέλος ευχαριστώ τους φίλους και συμφοιτητές μου Σταυρούλα Πουλοπούλου και Γιώργο Μακρυγιώτη για τις πολύτιμες συμβουλές τους και για την ηθική υποστήριξη σε όλη την διάρκεια της διπλωματικής εργασίας.

Αναστασία Ελευθεράκη

Ιούνιος 2004

Περίληψη

Από τις αρχές του 20^{ου} αιώνα διαπιστώθηκε η ανάγκη για την σύνθεση των αποτελεσμάτων ανεξάρτητων αλλά παρόμοιων πειραμάτων. Η μετανάλυση είναι απόρροια αυτής της ανάγκης και παρόλο που αναπτύχθηκε στον χώρο της κοινωνικής επιστήμης, καθιερώθηκε επίσημα από τη χρήση της στην ιατρική έρευνα. Με τον όρο μετανάλυση ονομάζουμε το σύνολο των στατιστικών τεχνικών με τις οποίες εξετάζονται τα αποτελέσματα ανεξάρτητων αλλά σχετικών κλινικών μελετών, με σκοπό την εξαγωγή ενός γενικού συμπεράσματος.

Σε αυτή τη διπλωματική εργασία παρουσιάσαμε το μεθοδολογικό πλαίσιο της μετανάλυσης. Ξεκινώντας περιγράφουμε τα βασικά συστατικά της, όπως είναι η κατασκευή του πρωτοκόλλου, η αναζήτηση και η επιλογή των σχετικών μελετών. Στη συνέχεια παρουσιάζονται οι στατιστικές τεχνικές που χρησιμοποιούνται στην μετανάλυση, δίνοντας έμφαση στις συνέπειες της επιλογής μιας ιδιαίτερης προσέγγισης και της ερμηνείας των αποτελεσμάτων. Εξετάζουμε τα μοντέλα σταθερών και τυχαίων επιδράσεων, τόσο την περίπτωση που είναι γνωστά τα συνοπτικά στατιστικά (*summary statistics*) όσο και την περίπτωση που γνωρίζουμε τα ατομικά δεδομένα των ασθενών. Επίσης παρουσιάζονται και απλές μεθοδολογίες για τον συνδυασμό δεδομένων, όπως είναι για παράδειγμα ο συνδυασμός των *p-values*. Περιγράφεται το πρόβλημα της ετερογένειας και της μεροληψίας, ενώ δίνονται και οι αντίστοιχες στατιστικές τεχνικές. Αν και αναφερόμαστε μόνο στη περίπτωση που τα δεδομένα είναι συνεχή ή δίτιμα, οι μεθοδολογίες αυτές μπορούν να εφαρμοστούν και στην περίπτωση άλλου τύπου δεδομένων. Στη συνέχεια γίνεται αναφορά στο λογισμικό που είναι διαθέσιμο για την πραγματοποίηση μιας μετανάλυσης, ενώ μέσα από ένα παράδειγμα διαπιστώνονται οι ελλείψεις αυτών των προγραμμάτων για τη σωστή εφαρμογή της μετανάλυσης. Στο τελευταίο κεφάλαιο διατυπώνουμε τα συμπεράσματα και τους προβληματισμούς μας σε σχέση με την εφαρμογή των μεθόδων της μετανάλυσης.

Abstract

The need for the composition of results of independent but similar experiments was realised by the beginning of 20th century. The meta-analysis is a consequence of this need and even if it was developed in the field of social sciences, it was officially established by its use in the medical research. By the term meta-analysis we name the collection of statistical techniques with which the results of independent but relative clinical trials are examined, aiming to the statement of a general conclusion.

In this project we reviewed the methodological frame of meta-analysis. In the beginning we describe the basic components of meta-analysis, such as the design of the protocol, the search and the selection of relative studies. Then we present the statistical techniques that are used in the meta-analysis, emphasizing the consequences of choosing a particular approach and the interpretation of the results. We examine the models of fixed and random effects in the case of summary statistics and when the individual patient data are available. Furthermore we present simple techniques for combining data, as for example the combination of p-values. We describe the problem of heterogeneity and bias, while the corresponding statistical techniques are given. Even if we are only referred in the case where the data are continuous or binary, these methodologies can be also applied in the case of other data types. We continue with the software that is available for the realization of meta-analysis, while through an example we detect the lack of these programs for the right application of meta-analysis. At the end we state our conclusions concerning the application of the meta-analytic methods.

Περιεχόμενα

| | |
|--|--------------|
| Ευχαριστίες | ix |
| Περίληψη | xi |
| Abstract | xiii |
| | |
| 1. Εισαγωγή | 1-6 |
| 1.1 Τι είναι η μετανάλυση; | 1 |
| 1.2 Αναδρομή στην ιστορία της μετανάλυσης | 2 |
| 1.3 Ο ρόλος της μετανάλυσης στην ιατρική επιστήμη | 3 |
| 1.4 Το περιεχόμενο της διπλωματικής | 4 |
| | |
| 2. Ξεκινώντας μια μετανάλυση | 7-11 |
| 2.1 Ανάπτυξη πρωτοκόλλου | 7 |
| 2.2 Αναζήτηση της σχετικής βιβλιογραφίας | 8 |
| 2.3 Αξιολόγηση της ανακτημένης βιβλιογραφίας | 9 |
| | |
| 3. Επίδραση της θεραπείας (<i>treatment effect</i>) | 13-26 |
| 3.1 Εισαγωγή | 13 |
| 3.2 Συνεχή μεταβλητή απόκρισης | 14 |
| 3.2.1 Διαφορά μέσων τιμών | 16 |
| 3.2.2 Τυποποιημένη διαφορά μέσων τιμών | 17 |
| 3.2.3 Έλεγχος της κοινής διακύμανσης μεταξύ των μελετών | 18 |
| 3.2.4 Συγκρίσεις κατά ζεύγη | 19 |
| 3.3 Δίτιμη μεταβλητή απόκρισης | 20 |
| 3.3.1 Διαφορά ποσοστών | 21 |
| 3.3.2 Λόγος ποσοστών ή σχετικός κίνδυνος | 22 |
| 3.3.3 Λόγος σχετικών πιθανοτήτων | 23 |

| | | |
|-----------|---|--------------|
| 3.3.4 | Σύγκριση των μέτρων για δίτιμα δεδομένα | 24 |
| 3.3.5 | Σύγκριση ποσοστών κατά ζεύγη | 25 |
| 4. | Συνδυασμός των εκτιμητών | 27-49 |
| 4.1 | Εισαγωγή | 27 |
| 4.2 | Μοντέλα σταθερών επιδράσεων | 28 |
| 4.2.1 | Μέθοδος Mantel-Haenszel | 28 |
| 4.2.1.1 | Διαφορά ποσοστών | 30 |
| 4.2.1.2 | Λόγος ποσοστών | 31 |
| 4.2.1.3 | Λόγος σχετικών πιθανοτήτων | 32 |
| 4.2.2 | Peto's Assumption Free Method | 33 |
| 4.2.3 | Inverse Variance-weighted method (W) | 34 |
| 4.2.4 | Σύγκριση των μεθόδων σταθερής επίδρασης | 38 |
| 4.3 | Μοντέλο Τυχαίων επιδράσεων | 39 |
| 4.3.1 | Εκτιμητής του τ^2 με την μέθοδο των ροπών | 42 |
| 4.3.2 | Εκτιμητής Μέγιστης Πιθανοφάνειας του τ^2 | 44 |
| 4.3.3 | Εκτιμητής Περιορισμένης Μέγιστης Πιθανοφάνειας του τ^2 | 45 |
| 4.4 | Σύγκριση των μοντέλων μετανάλυσης | 47 |
| 5. | Ετερογένεια | 51-65 |
| 5.1 | Τι είναι η ετερογένεια; | 51 |
| 5.2 | Γραφική διερεύνηση της ετερογένειας | 52 |
| 5.2.1 | Γράφημα των z-scores | 52 |
| 5.2.2 | Forest Plot | 53 |
| 5.2.3 | Radial Plot | 54 |
| 5.2.4 | L' Abbe Plot | 55 |
| 5.3 | Στατιστικός έλεγχος ετερογένειας | 56 |
| 5.4 | Πώς θα χειριστούμε την ετερογένεια; | 57 |
| 5.4.1 | Μετα-παλινδρόμηση | 59 |
| 5.4.1.1 | Μέθοδος σταθερών επιδράσεων | 60 |
| 5.4.1.2 | Μέθοδος Τυχαίων Επιδράσεων | 61 |

| | | |
|-----------|--|---------------|
| 5.4.1.2.1 | Εκτιμητής του τ^2 με τη μέθοδο των ροπών | 62 |
| 5.4.1.2.2 | Εκτιμητής ML του τ^2 | 63 |
| 5.4.1.2.3 | Εκτιμητής REML του τ^2 | 63 |
| 5.4.2 | Ανάλυση υποομάδων (<i>subgroup analysis</i>) | 64 |
| 6. | Απλές μέθοδοι για τον συνδυασμό μελετών | 67-72 |
| 6.1 | Μέθοδοι “ <i>Vote-counting</i> ” | 67 |
| 6.1.1 | Conventional “ <i>vote-counting</i> ” | 67 |
| 6.1.2 | Προσημικό τεστ | 68 |
| 6.2 | Συνδυασμός των <i>p-values</i> | 68 |
| 6.2.1 | Μέθοδος του <i>Tippet</i> | 70 |
| 6.2.2 | Μέθοδος του <i>Fisher</i> | 70 |
| 6.2.3 | Μέθοδος του <i>Stouffer</i> | 71 |
| 6.2.4 | Μέθοδος των <i>Mosteller</i> και <i>Bush</i> | 71 |
| 7. | Μετανάλυση με τα ατομικά δεδομένα των ασθενών | 73-81 |
| 7.1 | Εισαγωγή | 73 |
| 7.2 | Μοντέλο σταθερών επιδράσεων | 75 |
| 7.2.1 | Συνεχής μεταβλητή απόκρισης | 75 |
| 7.2.2 | Δίτιμη μεταβλητή απόκρισης | 77 |
| 7.3 | Μοντέλο τυχαίων επιδράσεων | 79 |
| 7.2.1 | Συνεχής μεταβλητή απόκρισης | 79 |
| 7.2.2 | Δίτιμη μεταβλητή απόκρισης | 80 |
| 8. | Αξιολόγηση του μοντέλου | 83-100 |
| 8.1 | Εισαγωγή | 83 |
| 8.2 | Ανάλυση Ευαισθησίας | 83 |
| 8.3 | Μεροληψία | 84 |
| 8.3.1 | Γράφημα “χοάνη” (<i>Funnel plot</i>) | 85 |

| | | |
|------------|--|----------------|
| 8.3.2 | Μη παραμετρική μέθοδος των Begg και Mazumdar | 88 |
| 8.3.3 | Egger's regression method | 89 |
| 8.3.4 | Εναλλακτική μέθοδος παλινδρόμησης | 91 |
| 8.3.5 | Μέθοδος " <i>Trim and Fill</i> " | 91 |
| 8.3.5.1 | Επαναληπτική μέθοδος " <i>Trim and Fill</i> " | 92 |
| 8.3.5.2 | Εφαρμογή της μεθόδου " <i>Trim and Fill</i> " | 94 |
| 8.3.6 | Rosenthal's file-drawer method | 97 |
| 8.3.7 | Weighted distributions approach | 99 |
| 9. | Λογισμικό Μετανάλυσης (<i>Software</i>) | 101-107 |
| 9.1 | Εισαγωγή | 101 |
| 9.2 | Εμπορικά προγράμματα (<i>Commercial software</i>) | 101 |
| 9.3 | Δωρεάν προγράμματα (<i>Free-ware software</i>) | 103 |
| 9.4 | Ρουτίνες μετανάλυσης διαθέσιμες σε γενικά στατιστικά πακέτα. | 106 |
| 9.5 | Σχετική Βιβλιογραφία | 106 |
| 10. | Εφαρμογή | 109-115 |
| 10.1 | Εισαγωγή | 109 |
| 10.2 | Ανάλυση | 109 |
| 11. | Συμπεράσματα | 117-120 |
| | Βιβλιογραφία | 121-124 |

ΚΕΦΑΛΑΙΟ 1

Εισαγωγή

1.1 Τι είναι η Μετανάλυση;

Αρχικά θα προσπαθήσουμε να αποδώσουμε την εννοιολογική σημασία του όρου “μετανάλυση”. Η λέξη στα Αγγλικά είναι σύνθετη και αποτελείται από δυο ελληνικές λέξεις την πρόθεση “μετα-”, η οποία δηλώνει χρονική ακολουθία, και το ουσιαστικό “ανάλυση” το οποίο δηλώνει την λεπτομερειακή και σε βάθος εξέταση. Επομένως ο όρος μετανάλυση περιγράφει μια διαδικασία μεταγενέστερης ανάλυσης. Στην συνέχεια θα αναφέρουμε και κάποιους ενδεικτικούς ορισμούς, οι οποίοι έχουν προταθεί στη διεθνή βιβλιογραφία, με σκοπό να αποκτήσουμε μια πιο ολοκληρωμένη ιδέα για την έννοια και τη χρήση της μετανάλυσης:

1. Μετανάλυση είναι η στατιστική ανάλυση μιας μεγάλης συλλογής αποτελεσμάτων, τα οποία προέρχονται από μεμονωμένες μελέτες, με σκοπό την ενσωμάτωση των συμπερασμάτων [Glass, 1978].
2. Μετανάλυση είναι η διαδικασία με την οποία οι στατιστικές μέθοδοι χρησιμοποιούνται για να συνδυαστούν τα αποτελέσματα διαφορετικών μελετών [Last, 1995].
3. Μετανάλυση είναι η μαθηματική σύνθεση των αποτελεσμάτων δύο ή περισσότερων αρχικών μελετών που εξετάζουν την ίδια υπόθεση με τον ίδιο τρόπο [Greenhaigh, 1997].
4. Μετανάλυση είναι η ποσοτική ανασκόπηση με σκοπό τη σύνθεση των αποτελεσμάτων, τα οποία προέρχονται από παρόμοιες αλλά ανεξάρτητες μελέτες [Normand, 1999].
5. Μετανάλυση είναι η ποσοτική στατιστική ανάλυση η οποία εφαρμόζεται σε παρόμοια πειράματα, ανεξάρτητων συνήθως ερευνητών με σκοπό να συγκεντρωθούν τα δεδομένα και με βάση αυτά να ελεγχθεί η αποτελεσματικότητα των συμπερασμάτων. [U.S National Library of Medicine]

1.2 Αναδρομή στην ιστορία της μετανάλυσης

Η στατιστική βάση της μετανάλυσης φθάνει πίσω στο 17^ο αιώνα όταν στην αστρονομία διαπίστωσαν, με οδηγό τη διαίσθηση και την εμπειρία, ότι ο συνδυασμός των στοιχείων είναι καλύτερος από την προσπάθεια επιλογής ενός στοιχείου μεταξύ όλων. Το 1904 ο διακεκριμένος στατιστικός *Karl Pearson* ήταν πιθανότατα ο πρώτος στην ιατρική έρευνα ο οποίος χρησιμοποίησε επίσημες στατιστικές τεχνικές για να συνδυάσει συντελεστές συσχέτισης από μελέτες που αφορούσαν τον εμβολιασμό κατά του τύφου. Ο *Tippet* το 1931 και ο *Fisher* το 1932 παρουσίασαν μεθόδους για τον συνδυασμό *p-values*, ενώ ο *Yates* και ο *Cochran* το 1938 μελέτησαν τον συνδυασμό εκτιμητών από διαφορετικά γεωργικά πειράματα. Εντούτοις, τέτοιες τεχνικές δεν χρησιμοποιήθηκαν ευρέως στην ιατρική για πολλά από τα επόμενα έτη. Σε αντίθεση με την ιατρική, οι κοινωνικές επιστήμες και ειδικότερα η ψυχολογία και η εκπαιδευτική έρευνα, έδειξαν πρόωρο ενδιαφέρον για τη σύνθεση των ερευνητικών συμπερασμάτων. Το 1976 ο ψυχολόγος *Gene Glass* έπλασε τον όρο "Μετανάλυση" στο άρθρο με τίτλο "*Primary, Secondary and Meta-Analysis of Research*". Αν και ο *Glass* αναφερόταν σε έρευνες της κοινωνικής επιστήμης, ο όρος μετανάλυση υιοθετήθηκε και έγινε ευρύτατα δημοφιλής στις κλινικές μελέτες, με αποτέλεσμα τη δεκαετία του '80 οι μέθοδοι της μετανάλυσης να ανθίζουν στον ιατρικό χώρο. Σημαντικά ιατρικά προβλήματα που αφορούσαν τις καρδιακές παθήσεις και τον καρκίνο λύθηκαν με χρήση των μεθόδων της μετανάλυσης. Το 1989 η "μετανάλυση" συμπεριλήφθηκε ως ιατρικός όρος στην βάση δεδομένων MEDLINE. Την δεκαετία του '90 δημοσιευμένες μεταanalύσεις έκαναν την εμφάνισή τους παντού. Αρκεί να αναφέρουμε ότι ο αριθμός των δημοσιευμένων μεταanalύσεων στις ιατρικές μελέτες αυξήθηκε από τις 18 την δεκαετία του '70, στις 406 την δεκαετία του '80, ενώ ο *Altman* [2000] σημείωσε ότι μόνο το 1997 η Medline είχε 589 δημοσιευμένες μεταanalύσεις ενώ το 1999 οι μεταanalύσεις έφτασαν τις 713.

Η γρήγορη αύξηση του αριθμού των μεταanalύσεων κατά την διάρκεια της τελευταίας δεκαετίας οφείλεται κυρίως στο ότι τα τελευταία χρόνια δίνεται ιδιαίτερη έμφαση στην "βασισμένη σε ενδείξεις ιατρική" (*evidence-based medicine*). Η συστηματική ανασκόπηση των ερευνών που αφορούν το ίδιο ιατρικό ζήτημα είναι αναγκαία για τη σφαιρική θεώρηση ενός προβλήματος, αφού ο αριθμός των δημοσιευμένων ιατρικών μελετών ολοένα και αυξάνει. Σε αυτό άλλωστε το σημείο έρχεται να βοηθήσει και η μετανάλυση, αφού όχι μόνο προϋποθέτει την συστηματική ανασκόπηση των σχετικών μελετών αλλά καταλήγει και στην

ποσοτική συγκέντρωση των αποτελεσμάτων τα οποία προέρχονται από τις μελέτες αυτές. Η διεθνής συνεργασία “*Cochrane Collaboration*”, η οποία αναπτύχθηκε το 1993, δίνει σημαντική ώθηση στην *evidence-based* ιατρική, αφού δεσμεύεται για την προετοιμασία, τη διατήρηση και τη διάδοση της συστηματικής ανασκόπησης της ιατρικής έρευνας.

1.3 Ο ρόλος της μετανάλυσης στην ιατρική επιστήμη

Τα επιτεύγματα της μετανάλυσης στη σφαίρα της κλινικής έρευνας είναι εντυπωσιακά. Κατ' αρχάς, η μετανάλυση βοήθησε να ξεπεραστεί το πρόβλημα που πρώτα προσδιορίστηκε από τον *Pearson*: "οποιοσδήποτε από τις ομάδες ... είναι πάρα πολύ μικρές για να επιτρέψουν την διατύπωση οποιασδήποτε καθορισμένης άποψης, λαμβάνοντας υπόψη το μέγεθος του πιθανού σφάλματος που εμπεριέχεται". Αν και το πλήθος των κλινικών μελετών που δημοσιεύονται στα ιατρικά περιοδικά έχει αυξηθεί, πολλές δοκιμές αποτυγχάνουν να καταλήξουν σε στατιστικά σημαντικά και γενικεύσιμα συμπεράσματα για την επίδραση μιας θεραπείας. Αυτό οφείλεται κυρίως στο ότι οι μελέτες αυτές είναι είτε πολύ μικρές είτε πολύ περιορισμένες, λόγω των κριτηρίων που ορίζονται στο πρωτόκολλο. Με τον συνδυασμό πληροφοριών από διαφορετικές μελέτες, η μετανάλυση αποκτά μεγαλύτερη στατιστική ισχύ για να ανιχνεύσει την επίδραση της θεραπείας, σε σχέση με μια ανάλυση που βασίζεται μόνο σε μια κλινική μελέτη.

Συχνά μελέτες που αντιμετωπίζουν το ίδιο ιατρικό πρόβλημα καταλήγουν σε διαφορετικά συμπεράσματα, τα οποία μπορούν να προκαλέσουν σύγχυση σε εκείνους, που με βάση τη βιβλιογραφία, επιδιώκουν την καθοδήγηση στο ιατρικό ζήτημα. Η μετανάλυση αποτελεί μια ελκυστική λύση αφού μπορεί να χρησιμοποιηθεί για να υπολογίσει την μέση επίδραση της θεραπείας στο σύνολο των κλινικών μελετών ή για να προσδιορίσει ένα υποσύνολο μελετών οι οποίες συνδέονται με μια ευεργετική-θετική επίδραση της θεραπείας.

Η έκρηξη των ερευνητικών στοιχείων τις τελευταίες δεκαετίες, υπό μορφή δημοσιευμένων κλινικών δοκιμών, δεν επιτρέπει την αφομοίωσή τους χωρίς επίσημη ανασκόπηση. Για να αξιολογηθούν τα οφέλη μιας θεραπευτικής αγωγής, θα πρέπει να βασιστούμε στο σύνολο των αξιόπιστων στοιχείων, όπως είναι για παράδειγμα το σύνολο των τυχαιοποιημένων κλινικών δοκιμών. Επιπλέον είναι ανάγκη να εντοπιστούν εκείνες οι ιατρικές περιοχές στις οποίες υπάρχει έλλειψη επαρκών στοιχείων προκειμένου διεξαχθούν επιπρόσθετες κλινικές

μελέτες. Ο *Chalmers* [2001] επισήμανε ότι οι συστηματικές ανασκοπήσεις των ήδη πραγματοποιημένων αλλά και των υπό εξέλιξη κλινικών δοκιμών πρέπει να είναι προϋποθέσεις για τον επιστημονικό και τον ηθικό σχεδιασμό νέων μελετών. Οι τεχνικές επομένως της μετανάλυσης αποτελούν το εργαλείο με το οποίο η ιατρική επιστήμη μπορεί να επιλύσει αυτά τα προβλήματα.

Στις βιοστατιστικές οδηγίες (*guidelines*) η μετανάλυση θεωρείται ως η επίσημη αξιολόγηση των ποσοτικών στοιχείων τα οποία προέρχονται από δυο ή περισσότερες κλινικές δοκιμές και αφορούν το ίδιο ζητούμενο. Οι τεχνικές της μετανάλυσης αναγνωρίζονται και εφαρμόζονται για τη διατύπωση ενός γενικού συμπεράσματος σε σχέση με την αποτελεσματικότητα ενός φαρμάκου, αλλά και για την ανάλυση των λιγότερο συχνών εκβάσεων στην αξιολόγηση της ασφάλειας. Έτσι, σύμφωνα με τις κατευθυντήριες οδηγίες του Ευρωπαϊκού Οργανισμού Αξιολόγησης των Φαρμακευτικών Προϊόντων (CPMP), η μετανάλυση μπορεί να εφαρμοστεί επιτυχώς:

1. Για να παρέχει μια ακριβέστερη εκτίμηση των γενικών αποτελεσμάτων.
2. Για να αξιολογήσει εάν τα θετικά αποτελέσματα μιας θεραπείας φαίνονται και σε προ-διευκρινισμένες υποομάδες ασθενών.
3. Για να αξιολογήσει μια πρόσθετη αποτελεσματικότητα της θεραπείας, η οποία απαιτεί μεγαλύτερη ισχύ απ' ότι έχουν οι μεμονωμένες μελέτες.
4. Για να αξιολογήσει την ασφάλεια σε μια υποομάδα ασθενών, ή ένα σπάνιο δυσμενές γεγονός σε όλους τους ασθενείς.
5. Για να βελτιώσει την σχέση δόσης –αντίδρασης του φαρμάκου.
6. Για να αξιολογήσει και να διερευνήσει τα συγκρουόμενα αποτελέσματα μελετών.

1.4 Το περιεχόμενο της διπλωματικής

Στην παρούσα διπλωματική παρουσιάζουμε το μεθοδολογικό πλαίσιο της μετανάλυσης, δίνοντας όμως μεγαλύτερη έμφαση στις στατιστικές μεθόδους που έχουν αναπτυχθεί. Επειδή η μετανάλυση είναι μια δομημένη διαδικασία θεωρούμε σκόπιμο να αναφέρουμε συνοπτικά τα βασικά βήματα για την διεξαγωγή της, προκειμένου να έχουμε μια πληρέστερη εικόνα:

1. Διατύπωση της ερευνητικής υπόθεσης
2. Ανάπτυξη του πρωτοκόλλου

3. Προσδιορισμός της πηγής των στοιχείων, ανάπτυξη της στρατηγικής αναζήτησης των σχετικών μελετών
 - Αναζήτηση δημοσιευμένης και αδημοσίευτης έρευνας
4. Καθιέρωση των κριτηρίων επιλογής ή αποκλεισμού των μελετών
 - Αξιολόγηση της ποιότητας των μελετών
5. Εξαγωγή και καταγραφή των στοιχείων που προκύπτουν από τις μελέτες
6. Στατιστική ανάλυση
 - Εκτίμηση της επίδρασης της θεραπείας στις επιμέρους μελέτες
 - Έλεγχος και αναζήτηση των πηγών της ετερογένειας
 - Επιλογή μοντέλου
 - Εκτίμηση της ολικής επίδρασης της θεραπείας
 - Έλεγχος και διόρθωση της μεροληψίας
 - Ανάλυση ευαισθησίας
 - Δευτερεύουσες αναλύσεις, *subgroup analyses*
7. Ερμηνεία των αποτελεσμάτων και συμπεράσματα

Στο επόμενο κεφάλαιο περιγράφεται συνοπτικά το πρώτο και βασικό σημαντικό κομμάτι της μετανάλυσης, η συστηματική ανασκόπηση (*systematic review*), στην οποία περιλαμβάνεται ο σχεδιασμός του πρωτοκόλλου, η αναζήτηση των μελετών και η αξιολόγησή τους. Στην συνέχεια παρουσιάζονται οι στατιστικές προσεγγίσεις για την ανάλυση των δεδομένων. Στο 3^ο κεφάλαιο περιγράφονται οι μέθοδοι για την εκτίμηση της επίδρασης της θεραπείας σε κάθε μελέτη, ενώ στο 4^ο κεφάλαιο παρουσιάζονται και συγκρίνονται τα δυο στατιστικά μοντέλα μετανάλυσης (μοντέλο σταθερών και μοντέλο τυχαίων επιδράσεων), τα οποία έχουν προταθεί για τον υπολογισμό της ολικής επίδρασης της θεραπείας. Οι μεθοδολογίες αναφέρονται στην περίπτωση που η κάθε μελέτη περιλαμβάνει δυο ομάδες θεραπείας (*parallel design*) με σκοπό την σύγκριση της νέας θεραπευτικής μεθόδου με το αντίστοιχο *control*, ενώ θα ασχοληθούμε μόνο με την περίπτωση που η μεταβλητή απόκρισης είναι συνεχής ή δίτιμη. Στο 5^ο κεφάλαιο εξετάζουμε το θέμα της ετερογένειας των μελετών και πώς μπορεί αυτή να ερμηνευτεί με χρήση συμμεταβλητών (*covariates*). Στα κεφάλαια 6 και 7 περιγράφονται μέθοδοι οι οποίες μπορούν να χρησιμοποιηθούν όταν οι μελέτες δίνουν ελάχιστη πληροφορία για την επίδραση της θεραπείας, ή αντίστοιχα, αν υπάρχουν διαθέσιμα τα ατομικά δεδομένα των ασθενών. Το 8^ο

κεφάλαιο πραγματεύεται το θέμα της μεροληψίας που υπεισέρχεται στην ανάλυση των δεδομένων λόγω της επιλογής των μελετών. Στο κεφάλαιο προτείνονται μοντέλα για την διόρθωση της μεροληψίας, ενώ γίνεται αναφορά στην ανάλυση ευαισθησίας που πρέπει να πραγματοποιηθεί για τον έλεγχο της ανθεκτικότητας των συμπερασμάτων. Στο κεφάλαιο 9 δίνονται πληροφορίες σχετικά με το λογισμικό που είναι διαθέσιμο για την διεξαγωγή μιας μετανάλυσης, ενώ στο κεφάλαιο 10, μέσα από ένα παράδειγμα μετανάλυσης, διαπιστώνονται πρακτικά προβλήματα που προκύπτουν χρησιμοποιώντας ένα εξειδικευμένο πρόγραμμα μετανάλυσης. Τέλος στο 11^ο κεφάλαιο καταλήγουμε σε συμπεράσματα σχετικά με τη χρήση των μεθόδων της μετανάλυσης, προβάλλοντας τα θετικά σημεία της ενώ παράλληλα διατυπώνουμε και κάποιους προβληματισμούς.

ΚΕΦΑΛΑΙΟ 2

Ξεκινώντας μια μετανάλυση

2.1 Ανάπτυξη πρωτοκόλλου

Πριν αρχίσει μια οποιαδήποτε κλινική μελέτη, είναι ανάγκη να κατασκευαστεί ένα πρωτόκολλο στο οποίο θα διευκρινίζονται λεπτομερώς όλες οι διαδικασίες που θα ακολουθηθούν από την έναρξη μέχρι την λήξη της μελέτης. Το ίδιο αναγκαίο είναι να κατασκευαστεί ένα πρωτόκολλο και για την εκτέλεση μιας μετανάλυσης, αφού και αυτή είναι μια σύνθετη διαδικασία. Αν χρειαστεί να γίνουν αλλαγές στο πρωτόκολλο της μετανάλυσης λόγω απρόβλεπτων περιστάσεων, θα πρέπει να έχει προηγηθεί σαφής τεκμηρίωση και στη συνέχεια να εξεταστεί το πώς αυτές οι αλλαγές επηρεάζουν τα αποτελέσματα της ανάλυσης. Στο πρωτόκολλο της μετανάλυσης, όπως και σε κάθε μελέτη, είναι απαραίτητο να δηλωθεί η βασική υπόθεση που μας ενδιαφέρει. Επιπλέον όταν γίνει η καταγραφή των συμπερασμάτων, είναι σημαντικό να υπάρχει διάκριση μεταξύ των προσχεδιασμένων αναλύσεων και των αναλύσεων που έγιναν για διερευνητικούς σκοπούς.

Η ιδανική κατάσταση είναι η δημιουργία ενός πρωτοκόλλου μετανάλυσης, πριν από την έναρξη των κλινικών δοκιμών οι οποίες θα πραγματοποιηθούν στο μέλλον για να εξετάσουν την ίδια θεραπευτική μέθοδο. Σε μια τέτοια περίπτωση τα πρωτόκολλα των μεμονωμένων μελετών σχεδιάζονται με βάση το πρωτόκολλο της μετανάλυσης, έτσι ώστε η καταγραφή και η συλλογή των στοιχείων να γίνει με τον ίδιο τρόπο, ενώ οι οποιοσδήποτε αλλαγές εξετάζονται σε παγκόσμια κλίμακα. Η μετανάλυση αυτής της μορφής έχει κοινά χαρακτηριστικά με μια πολυκεντρική μελέτη (*multicenter trial*) και ονομάζεται *προοπτική*.

Όμως οι μετανalύσεις συχνά γίνονται *αναδρομικά*, βασίζονται δηλαδή σε κλινικές μελέτες οι οποίες έχουν ήδη διεξαχθεί. Σε αυτή την περίπτωση η προετοιμασία του πρωτοκόλλου είναι ακόμα πιο κρίσιμη, αφού ο αριθμός των προβλημάτων που μπορεί να προκύψουν είναι μεγάλος. Επειδή η επιλογή των μελετών που θα συμπεριληφθούν στην μετανάλυση επηρεάζει την ορθότητα των αποτελεσμάτων της, στο πρωτόκολλο θα πρέπει να αναφέρονται τα σαφή

κριτήρια με βάση τα οποία θα γίνει αυτή η επιλογή έτσι ώστε να αποφευχθούν πιθανές μεροληψίες σχετικά με την επιλογή των μελετών (*selection bias*). Αν συμπεριληφθούν μόνο οι δημοσιευμένες μελέτες μπορεί να οδηγηθούμε σε μεροληψία, η οποία είναι γνωστή ως σφάλμα δημοσίευσης (*publication bias*), αφού η πιθανότητα έκδοσης μιας μελέτης εξαρτάται από την στατιστική σημασία του αποτελέσματος. Επειδή οι διαθέσιμες μελέτες μπορεί να διαφέρουν στο σχεδιασμό, στα χαρακτηριστικά των ασθενών, στην θεραπευτική αγωγή που εφαρμόζεται, στις κύριες υποθέσεις, στα στοιχεία που καταγράφονται αλλά και στην ποιότητα, στο πρωτόκολλο θα πρέπει να καθορίζονται τα κριτήρια για τον αποκλεισμό κάποιων μελετών από την μετανάλυση καθώς και ο τρόπος χειρισμού των δεδομένων. Πιο συγκεκριμένα στο πρωτόκολλο θα πρέπει να διευκρινίζονται οι εξής παράμετροι:

- Ο κύριος στόχος ή οι στόχοι της ανάλυσης.
- Κριτήρια για τον συνυπολογισμό και τον αποκλεισμό των μελετών π.χ πληθυσμός της μελέτης, σχέδιο μελέτης (τυχαιοποιημένες, *blinded*), είδος θεραπείας κλπ.
- Χειρισμός μεροληψίας λόγω της επιλογής των μελετών (*selection bias*).
- Διευκρινίσεις για την στρατηγική προσδιορισμού των μελετών. Τρόποι αντιμετώπισης της μεροληψίας λόγω δημοσίευσης (*publication bias*).
- Υποθέσεις και εξαρτημένες μεταβλητές (*endpoints*).
- Στατιστική μεθοδολογία (συμπεριλαμβανομένων και των μεθόδων για τον έλεγχο της ετερογένειας). Αν οι μελέτες περιλαμβάνουν διαφορετικά σχέδια ανάλυσης θα πρέπει να διευκρινιστεί η μεθοδολογία που θα ακολουθηθεί.
- Προσεγγίσεις με στόχο την μεγιστοποίηση της ποιότητας των δεδομένων όλων των συμπεριλαμβανόμενων μελετών. Έλεγχος για πιθανές επιρροές στην μεταβλητότητα της ποιότητας των δεδομένων και μέθοδοι εξέτασης των ελλειπουσών παρατηρήσεων (*missing values*).
- Σχέδιο για την αξιολόγηση της συνέπειας και της ανθεκτικότητας των αποτελεσμάτων (*sensitivity analysis*). Αυτό το σχέδιο θα πρέπει να στοχεύει να καταδείξει παρόμοια αποτελέσματα σε εναλλακτικές αναλύσεις διαφορετικών *endpoints*, διαφορετικών υποσυνόλων πληθυσμού και διαφορετικών υποσυνόλων μελετών.

2.2 Αναζήτηση της σχετικής βιβλιογραφίας

Αφού καθοριστούν οι στόχοι της μετανάλυσης και κατασκευαστεί το πρωτόκολλο, θα πρέπει να ξεκινήσει μια βιβλιογραφική αναζήτηση προκειμένου να ανακτηθούν οι παρόμοιες μελέτες. Για να ελαχιστοποιηθούν τα προβλήματα τα οποία σχετίζονται με το σφάλμα δημοσίευσης (*publication bias*), είναι αναγκαίο να καταστρώσουμε κάποια στρατηγική “αναζήτησης”. Η στρατηγική αυτή θα πρέπει να προσδιορίζεται λεπτομερώς στο πρωτόκολλο της μετανάλυσης. Η αναζήτηση θα είναι ανεπαρκής αν οι σχετικές μελέτες προσδιοριστούν μόνο μέσω αυτοματοποιημένων βιβλιογραφικών αναζητήσεων, όπως είναι για παράδειγμα το δίκτυο MEDLINE ή το EMBASE. Είναι γεγονός ότι αν στην μετανάλυση βασιστούμε μόνο σε δημοσιευμένες μελέτες θα οδηγηθούμε σε υπερεκτίμηση της επίδρασης της θεραπείας, αφού υπάρχει η τάση να δημοσιεύονται αποτελέσματα τα οποία είναι στατιστικά σημαντικά. Για την εξάλειψη αυτής της μεροληψίας ο αναλυτής θα πρέπει λάβει υπόψη του και την αδημοσίευτη έρευνα. Οι κατάλογοι κλινικής έρευνας και κλινικών δοκιμών είναι άλλη μια σημαντική πηγή πληροφοριών, αφού παρέχουν πληροφορίες για όλες τις μελέτες που έχουν ολοκληρωθεί ή είναι σε εξέλιξη. Σημαντική επίσης πηγή πληροφορίας είναι και τα αρχεία των φαρμακευτικών εταιριών, αν και σε αυτή την περίπτωση η πρόσβαση είναι αρκετά δύσκολη. Τέλος οι αδημοσίευτες διατριβές και οι διπλωματικές μεταπτυχιακών φοιτητών μπορεί να παρέχουν εξίσου σημαντικό όγκο πληροφορίας, επομένως η αναζήτηση θα πρέπει να περιλαμβάνει και τις βάσεις δεδομένων των πανεπιστημίων.

Συμπερασματικά, η αναζήτηση των σχετικών μελετών θα πρέπει να είναι απόρροια μιας καλά οργανωμένης προσπάθειας. Η ομάδα που θα ασχοληθεί με αυτό το βασικό κομμάτι της μετανάλυσης οφείλει να είναι σωστά ενημερωμένη και συντονισμένη. Η επιτυχία της αναζήτησης των μελετών είναι αποτέλεσμα της συνεργασίας βιβλιοθηκάρων και ιατρών οι οποίοι μπορούν, ο καθένας με βάση την επιστημονική του κατάρτιση, να βοηθήσουν ανάλογα.

2.3 Αξιολόγηση της βιβλιογραφίας

Αφού ανακτηθεί η σχετική βιβλιογραφία, στην συνέχεια θα πρέπει να αποφασίσουμε ποιες μελέτες θα συμπεριληφθούν στη μετανάλυση και ποιες θα απορριφθούν. Από κάθε μελέτη θα

πρέπει να είναι δυνατή η συλλογή κάποιων βασικών στοιχείων, όπως: στοιχεία που αφορούν το τυπικό μέρος της έκθεσης (συντάκτης, έτος, πηγή δημοσίευσης), στοιχεία της μελέτης (γενικά χαρακτηριστικά γνωρίσματα του πληθυσμού, όπως για παράδειγμα το κοινωνικό και δημογραφικό προφίλ, μέγεθος δείγματος), στοιχεία που αφορούν τους ασθενείς (δημογραφικά και κλινικά γνωρίσματα), στοιχεία για το ερευνητικό σχέδιο (παρατηρητικός ή τυχαίος μηχανισμός δειγματοληψίας, τυχαιοποίηση στην θεραπεία, ποσοστά συμμόρφωσης με το πρωτόκολλο κλπ), στοιχεία για την θεραπευτική αγωγή (διάρκεια, δόση, συγχρονισμός δόσεων), στοιχεία για το μέγεθος της επίδρασης της θεραπείας (εκτίμηση, τυπικό σφάλμα). Η ποσότητα και η ποιότητα των πληροφοριών τα οποία παρέχονται σε οποιαδήποτε μελέτη εξαρτάται από διάφορους παράγοντες οι οποίοι σχετίζονται κυρίως με την κατάρτιση και την εμπειρία του συντάκτη της μελέτης καθώς και με το περιοδικό στο οποίο γίνεται η δημοσίευση.

Οι κανόνες για τον έλεγχο της ποιότητας των μελετών δεν είναι καθορισμένοι με σαφήνεια. Ο *Chalmer* [1981] περιγράφει ένα σύνολο κριτηρίων για την λήψη αποφάσεων σχετικά με τον συνυπολογισμό ή τον αποκλεισμό μιας τυχαιοποιημένης μελέτης από την μετανάλυση. Οι τρεις περιοχές που προτείνει ότι πρέπει να εξεταστούν είναι: ο σχεδιασμός της μελέτης, η εφαρμογή του πρωτοκόλλου και η στατιστική ανάλυση. Επομένως η αξιολόγηση μιας μελέτης βασίζεται κυρίως σε ζητήματα σχεδιασμού, όπως είναι η διαδικασία τυχαιοποίησης των ασθενών στις ομάδες, το αν η μελέτη είναι τυφλή (*blinding*), πώς έγινε η καταγραφή της απόκρισης του κάθε ασθενή, πώς έγινε η παρακολούθηση των ασθενών, αν υπήρχαν αποκλίσεις από το πρωτόκολλο και πώς έγινε ο χειρισμός των ασθενών που αποχώρησαν από την μελέτη. Αν και αυτές οι πτυχές μιας ανάλυσης δεν διατυπώνονται με σαφήνεια στις δημοσιεύσεις, η αξιολόγηση γίνεται συνήθως από μια επιτροπή εμπειρογνομόνων, σύμφωνα με την εξής διαδικασία: δυο ειδικοί οι οποίοι δεν γνωρίζουν τους συντάκτες, την πηγή και τα αποτελέσματα μιας μελέτης κάνουν, ανεξάρτητα ο ένας από τον άλλο, τις εκτιμήσεις τους σχετικά με την ποιότητα της μελέτης. Το ποσοστό το οποίο προκύπτει καθορίζει τον αποκλεισμό ή τον συνυπολογισμό της μελέτης. Στην συνέχεια τα ποσοστά αυτά μπορούν να ενσωματωθούν στην ανάλυση ευαισθησίας της μετανάλυσης, προκειμένου να εξεταστεί αν για παράδειγμα υπάρχει αλλαγή στο τελικό συμπέρασμα όταν η ανάλυση περιοριστεί μόνο σε υψηλής ποιότητας μελέτες.

Συνήθως για κάθε υπόθεση που εξετάζεται, κατασκευάζονται δυο ομάδες μελετών. Η πρώτη ομάδα περιλαμβάνει όλες τις κύριες μελέτες πάνω στις οποίες θα βασιστεί η

μετανάλυση, ενώ η δεύτερη ομάδα περιλαμβάνει μελέτες οι οποίες αναφέρονται σε διαφορετικούς πληθυσμούς ασθενών ή συγκρίνουν διαφορετικές θεραπείες σε σχέση με τις κύριες μελέτες. Οι μελέτες αυτές μπορούν επίσης να χρησιμοποιηθούν στην ανάλυση ευαισθησίας.

Αν τα κριτήρια επιλογής των μελετών είναι αυστηρά, τότε τα συμπεράσματα της μετανάλυσης θα είναι εφαρμόσιμα μόνο σε ένα μικρό σύνολο ασθενών ή θα αναφέρονται σε μια πολύ συγκεκριμένη θεραπευτική αγωγή. Αφετέρου, αν τα κριτήρια επιλογής των μελετών είναι πολύ ελαστικά, δεν θα μπορούν να συνδυαστούν τα αποτελέσματα των μεμονωμένων μελετών αφού οι μελέτες θα διαφέρουν αρκετά στο σχεδιασμό.

Αυτό που θα πρέπει να γνωρίζουμε είναι ότι η ποιότητα μιας μετανάλυσης είναι υψηλή μόνο εάν η ποιότητα των συμπεριλαμβανόμενων μελετών είναι υψηλή. Εάν οι κλινικές μελέτες είναι μεροληπτικές, το ίδιο μεροληπτική θα είναι και η μετανάλυση. Ο σχεδιασμός του πρωτοκόλλου και οι μέθοδοι ποιοτικής αξιολόγησης των μελετών αποτελούν μια διαδικασία η οποία αναφέρεται ως συστηματική ανασκόπηση (*systematic review*). Η διαδικασία αυτή αποτελεί τη βάση πάνω στην οποία θα στηριχτεί η μετανάλυση, προκειμένου να συνδυαστούν κατάλληλα τα αποτελέσματα των μεμονωμένων μελετών με απώτερο σκοπό τη διατύπωση ενός γενικότερου συμπεράσματος για την επίδραση της θεραπείας.

ΚΕΦΑΛΑΙΟ 3

Επίδραση της θεραπείας *(Treatment effect)*

3.1 Εισαγωγή

Μετά την συγκέντρωση των μελετών, το επόμενο βήμα είναι να εκτιμήσουμε σε κάθε μελέτη την επίδραση της θεραπείας σε σχέση με το *control*, βασιζόμενοι σε μια ή περισσότερες μεταβλητές απόκρισης. Επειδή η επίδραση της θεραπείας (*treatment effect*) πρέπει να εκτιμηθεί με τον ίδιο τρόπο σε όλες τις μελέτες, τα διαθέσιμα στοιχεία των μεμονωμένων κλινικών μελετών θα πρέπει να είναι επαρκή γι' αυτόν τον υπολογισμό. Έτσι, ανάλογα με το ποια είναι η μορφή των διαθέσιμων στοιχείων, έχουν αναπτυχθεί διαφορετικές μεθοδολογικές προσεγγίσεις για τη διεξαγωγή της μετανάλυσης.

Η πιο περιορισμένη ανάλυση, σχετικά με το ποσό της πληροφορίας που είναι διαθέσιμο από την κάθε μελέτη, είναι αυτή στην οποία οι μεμονωμένες μελέτες αναφέρουν μόνο αν η επίδραση της θεραπείας είναι θετική ή αρνητική, ή αναφέρουν μόνο το *p-value* του αντίστοιχου ελέγχου. Στη περίπτωση αυτή έχουν αναπτυχθεί απλές μεθοδολογίες για να ελέγξουν την ολική επίδραση της θεραπείας, οι οποίες θα εξεταστούν στο 6^ο κεφάλαιο. Διαφορετικά, οι μελέτες μπορεί να παρέχουν τον εκτιμητή της επίδρασης της θεραπείας με το αντίστοιχο τυπικό σφάλμα. Αν και για να διεξαχθεί η μετανάλυση η γνώση αυτής της πληροφορίας είναι αρκετή, το πρόβλημα που μπορεί να προκύψει είναι ότι θα πρέπει όλες οι μελέτες να έχουν υπολογίσει τον ίδιο εκτιμητή της επίδρασης της θεραπείας, ο οποίος ταυτίζεται με τον εκτιμητή βάση του οποίου έχει αποφασιστεί να διεξαχθεί η μετανάλυση. Για παράδειγμα, αν η μεταβλητή απόκρισης είναι δίτιμη και θέλουμε να εκτιμήσουμε τον λογάριθμο του λόγου των σχετικών πιθανοτήτων (*log-odds ratio*), τότε μια μελέτη στην οποία αναφέρεται ο εκτιμητής της διαφοράς των ποσοστών των δύο ομάδων δεν μπορεί να συνεισφέρει στην μετανάλυση.

Όταν στις μεμονωμένες μελέτες αναφέρονται πίνακες με συνοπτικά στατιστικά στοιχεία (*summary statistics*) για τις δυο ομάδες θεραπείας, τότε μπορούμε να εκτιμήσουμε την επίδραση της θεραπείας με όποιο μέτρο θέλουμε. Για παράδειγμα, αν η μεταβλητή απόκρισης είναι συνεχής τότε η γνώση του πλήθους των ασθενών, της μέσης τιμής και της τυπικής απόκλισης των δύο ομάδων επιτρέπουν την σύγκριση των θεραπειών χρησιμοποιώντας την απόλυτη ή την τυποποιημένη διαφορά των μέσων τιμών των δύο ομάδων. Στη συνέχεια οι εκτιμητές αυτοί συνδυάζονται κατάλληλα για να εκτιμηθεί η ολική επίδραση τη θεραπείας. Η μετανάλυση με χρήση συνοπτικών στατιστικών θα παρουσιαστεί στο 4^ο κεφάλαιο.

Τέλος αν στις μεμονωμένες μελέτες αναφέρονται τα αρχεία με τα ατομικά στοιχεία των ασθενών, τότε μπορούμε και πάλι να επιλέξουμε οποιοδήποτε μέτρο σύγκρισης των δυο θεραπειών. Στην περίπτωση όμως αυτή, η γνώση των αρχικών δεδομένων επιτρέπει μια πιο λεπτομερή μέθοδο ανάλυσης την οποία θα αναπτύξουμε στο κεφάλαιο 7.

Στο παρόν κεφάλαιο θα εξετάσουμε την περίπτωση που η μεταβλητή απόκρισης είναι συνεχής ή δίτιμη και για κάθε περίπτωση θα παρουσιάσουμε τον τρόπο με τον οποίο μπορούμε να εκτιμήσουμε την επίδραση της θεραπείας, για να εφαρμόσουμε τις τεχνικές της μετανάλυσης του 4^{ου} κεφαλαίου. Χρησιμοποιώντας αντίστοιχες μεθοδολογίες, η μετανάλυση μπορεί να εφαρμοστεί όταν οι αποκριτικές μεταβλητές είναι κατηγορικές ή όταν τα δεδομένα προέρχονται από ανάλυση επιβίωσης.

3.2 Συνεχής μεταβλητή απόκρισης (*Continuous data*)

Σε κάθε μια από τις k μελέτες υποθέτουμε ότι υπάρχουν δυο ανεξάρτητες ομάδες ασθενών (*parallel design*), η ομάδα στην οποία εφαρμόζεται η θεραπευτική αγωγή που μας ενδιαφέρει (*treatment*) και η ομάδα στην οποία εφαρμόζεται η θεραπεία σύγκρισης (*control*). Αν η μεταβλητή απόκρισης είναι συνεχής, τότε μπορούμε υποθέσουμε ότι ακολουθεί κανονική κατανομή. Ακόμη κι αν αυτό δεν ισχύει, μπορούμε να το επιτύχουμε με κατάλληλο μετασχηματισμό. Θεωρούμε επομένως ότι τα δεδομένα των ασθενών οι οποίοι ανήκουν στην ομάδα θεραπείας (*treatment*) προέρχονται από κανονική κατανομή με μέση τιμή μ_T και τυπική απόκλιση σ_T , ενώ τα δεδομένα των ασθενών της ομάδας *control* προέρχονται από κανονική κατανομή με μέση τιμή μ_C και τυπική απόκλιση σ_C .

Έστω n_T το πλήθος των ασθενών της ομάδας *treatment* και n_C το πλήθος των ασθενών της ομάδας *control*. Αν y_{jT} , $j=1,2,\dots,n_T$ είναι οι αποκρίσεις των ασθενών της ομάδας *treatment* και y_{jC} , $j=1,2,\dots,n_C$ είναι οι αποκρίσεις των ασθενών της ομάδας *control*, τότε η δειγματική μέση τιμή και η δειγματική τυπική απόκλιση για τις δύο ομάδες, δίνονται από τις σχέσεις:

$$\text{Ομάδα "treatment": } \bar{y}_T = \frac{1}{n_T} \sum_{j=1}^{n_T} y_{jT} \quad \text{με} \quad s_T^2 = \frac{\left(\sum_{j=1}^{n_T} y_{jT}^2\right) - n_T \bar{y}_T^2}{n_T - 1}$$

$$\text{Ομάδα "control": } \bar{y}_C = \frac{1}{n_C} \sum_{j=1}^{n_C} y_{jC} \quad \text{με} \quad s_C^2 = \frac{\left(\sum_{j=1}^{n_C} y_{jC}^2\right) - n_C \bar{y}_C^2}{n_C - 1}$$

Έστω ότι τα δεδομένα της i μελέτης συνοψίζονται σε πίνακα της μορφής:

| | Ομάδα "treatment" | Ομάδα "control" | Σύνολο |
|-----------------|-------------------|-----------------|--------|
| Αριθμός ασθενών | n_T | n_C | n |
| Μέση τιμή | \bar{y}_T | \bar{y}_C | |
| Τυπική απόκλιση | s_T | s_C | |

Πίνακας 3.1 Συνοπτικά στατιστικά για συνεχείς μεταβλητές

Όταν η αποκριτική μεταβλητή είναι συνεχής και τα δεδομένα είναι κανονικά κατανομημένα, χρησιμοποιούμε συνήθως δυο μέτρα για να εκφράσουμε την διαφορά των δυο θεραπειών. Το ένα είναι η διαφορά των μέσων τιμών $\theta = \mu_T - \mu_C$ και το δεύτερο είναι η τυποποιημένη διαφορά των μέσων τιμών $\theta = \frac{\mu_T - \mu_C}{\sigma}$. Η διαφορά των μέσων τιμών έχει το πλεονέκτημα ότι μπορεί να ερμηνευτεί εύκολα ενώ είναι κατάλληλη όταν σε όλες τις μελέτες έχει χρησιμοποιηθεί η ίδια κλίμακα μέτρησης στην καταγραφή της απόκρισης. Η τυποποιημένη διαφορά μπορεί να χρησιμοποιηθεί όταν πρέπει να συνδυαστούν μετρήσεις προέρχονται από ίδιες μεταβλητές αλλά που έχουν χρησιμοποιηθεί διαφορετικές κλίμακες μέτρησης.

3.2.1 Διαφορά μέσων τιμών (*Difference between means*)

Ας υποθέσουμε ότι θέλουμε να εκτιμήσουμε την διαφορά των μέσων τιμών των δυο ομάδων, δηλαδή θέλουμε να εκτιμήσουμε την παράμετρο $\theta = \mu_T - \mu_C$. Ο εκτιμητής μέγιστης πιθανοφάνειας (MLE) της παραμέτρου θ είναι ο $\hat{\theta} = \bar{y}_T - \bar{y}_C$.

Για να μπορέσουμε να υπολογίσουμε τη διακύμανση του εκτιμητή $\hat{\theta}$, θα διακρίνουμε τις εξής περιπτώσεις:

A) Αν οι διακυμάνσεις σ_T^2 και σ_C^2 είναι γνωστές τότε η διακύμανση του εκτιμητή $\hat{\theta}$ είναι ίση με:

$$Var(\hat{\theta}) = \frac{\sigma_T^2}{n_T} + \frac{\sigma_C^2}{n_C}$$

B) Αν οι διακυμάνσεις σ_T^2 και σ_C^2 είναι άγνωστες αλλά ισχύει η υπόθεση $\sigma_T^2 = \sigma_C^2 = \sigma^2$, τότε η διακύμανση του εκτιμητή $\hat{\theta}$ δίνεται από την σχέση:

$$Var(\hat{\theta}) = \sigma^2 \left(\frac{1}{n_T} + \frac{1}{n_C} \right)$$

και ο εκτιμητής μέγιστης πιθανοφάνειας της κοινής διακύμανσης σ^2 είναι ο:

$$\hat{\sigma}^2 = \frac{(n_T - 1)s_T^2 + (n_C - 1)s_C^2}{n_T + n_C}.$$

Επειδή ο εκτιμητής αυτός δεν είναι αμερόληπτος, συνήθως χρησιμοποιούμε τον λεγόμενο “αθροιστικό εκτιμητή” της διακύμανσης σ^2 , ο οποίος είναι αμερόληπτος εκτιμητής και δίνεται από τη σχέση:

$$s^2 = \frac{(n_T - 1)s_T^2 + (n_C - 1)s_C^2}{n_T + n_C - 2}$$

Ο τελευταίος εκτιμητής συχνά στην βιβλιογραφία ονομάζεται και “εκτιμητής περιορισμένης μέγιστης πιθανοφάνειας”.

Επομένως η εκτιμημένη διακύμανση του εκτιμητή $\hat{\theta}$ είναι ίση με:

$$Var(\hat{\theta}) = s^2 \left(\frac{1}{n_T} + \frac{1}{n_C} \right)$$

Στο σημείο αυτό θα πρέπει να σημειώσουμε ότι στις κλινικές μελέτες συνήθως γίνεται η υπόθεση ότι οι διακυμάνσεις των δυο ομάδων θεραπείας είναι άγνωστες άλλα ίσες μεταξύ τους. Γι’ αυτό τον λόγο, στις επόμενες παραγράφους θα θεωρήσουμε ότι η υπόθεση των ίσων διακυμάνσεων είναι αληθής, χωρίς βέβαια αυτό να μας περιορίζει στο να εφαρμόσουμε την αντίστοιχη μεθοδολογία για την περίπτωση που οι διακυμάνσεις είναι άνισες.

3.2.2 Τυποποιημένη διαφορά μέσων τιμών (*Standardized difference between means*)

Ας υποθέσουμε ότι θέλουμε να εκτιμήσουμε την διαφορά των μέσων τιμών των δυο ομάδων, δηλαδή θέλουμε να εκτιμήσουμε την παράμετρο $\theta = \frac{\mu_T - \mu_C}{\sigma}$, όπου σ^2 είναι η κοινή διακύμανση των δυο ομάδων θεραπείας.

Ο εκτιμητής περιορισμένης μέγιστης πιθανοφάνειας (REML) της παραμέτρου θ είναι ο $\hat{\theta} = \frac{\bar{y}_T - \bar{y}_C}{s}$, όπου $s^2 = \frac{(n_T - 1)s_T^2 + (n_C - 1)s_C^2}{n_T + n_C - 2}$. Ο εκτιμητής $\hat{\theta}$ έχει επικρατήσει στην βιβλιογραφία με το όνομα *Cohen’s d*.

Η ασυμπτωτική διακύμανση του $\hat{\theta}$ δίνεται από τη σχέση: $Var(\hat{\theta}) = \frac{n_T + n_C}{n_T \cdot n_C}$.

Επειδή για μικρό μέγεθος δείγματος ο εκτιμητής $\hat{\theta}$ είναι μεροληπτικός, έχει προταθεί ένας σταθμισμένος αμερόληπτος εκτιμητής ο οποίος ονομάζεται *Hedges' g* [Deeks, 1999] και δίνεται από την σχέση:

$$\hat{\theta} = \frac{\bar{y}_T - \bar{y}_C}{s} \left(1 - \frac{3}{4(n_T + n_C) - 9} \right)$$

Ο εκτιμητής της διακύμανσης του $\hat{\theta}$ είναι ίσος με:

$$Var(\hat{\theta}) = \frac{n_T + n_C}{n_T \cdot n_C} + \frac{\hat{\theta}^2}{2(n_T + n_C)}$$

3.2.3 Έλεγχος κοινής διακύμανσης μεταξύ των μελετών

Κατά την εκτέλεση της μετανάλυσης είναι πολλές φορές χρήσιμο να υποθέσουμε ότι σε όλες τις μελέτες ισχύει ότι $\sigma_i^2 = \sigma^2$, $\forall i = 1, \dots, k$, όπου σ_i^2 είναι η κοινή διακύμανση των δυο ομάδων θεραπείας στην i μελέτη. Ο εκτιμητής της κοινής διακύμανσης των k ανεξάρτητων μελετών δίνεται από την σχέση:

$$s_p^2 = \frac{\sum_{i=1}^k (n_i - 1) s_i^2}{\sum_{i=1}^k (n_i - 1)},$$

όπου n_i είναι το συνολικό πλήθος των ασθενών της i μελέτης και s_i^2 είναι ο αθροιστικός εκτιμητής της διακύμανσης σ_i^2 ο οποίος βασίζεται στην i μελέτη, δηλαδή

$$s_i^2 = \frac{(n_{iT} - 1) s_{iT}^2 + (n_{iC} - 1) s_{iC}^2}{n_{iT} + n_{iC} - 2}, \quad i = 1, 2, \dots, k.$$

Η υπόθεση της κοινής διακύμανσης μεταξύ όλων των μελετών, μπορεί να εξεταστεί χρησιμοποιώντας το τεστ του *Bartlett* [Bartlett, 1937]. Το στατιστικό του ελέγχου δίνεται από την σχέση:

$$T = \frac{1}{c} \left\{ (n - 2\kappa) \log s_p^2 - \sum_{i=1}^{\kappa} (n_i - 2) \log s_i^2 \right\}$$

όπου $c = 1 + \frac{1}{3(\kappa - 1)} \left\{ \left(\sum_{i=1}^{\kappa} \frac{1}{n_i - 2} \right) - \frac{1}{n - 2\kappa} \right\}$ και n είναι το σύνολο των ασθενών σε όλες τις μελέτες. Κάτω από την μηδενική υπόθεση της κοινής διακύμανσης το στατιστικό ακολουθεί ασυμπτωτικά $X_{\kappa-1}^2$.

3.2.4 Συγκρίσεις κατά ζεύγη (*Paired samples*)

Αρκετά συχνά στις κλινικές μελέτες προκύπτουν συγκρίσεις κατά ζεύγη. Αυτό μπορεί να συμβεί είτε εφαρμόζοντας δυο θεραπείες στο ίδιο άτομο (ή σε διαφορετικά άτομα παρόμοιων χαρακτηριστικών) είτε εφαρμόζοντας μια θεραπεία και λαμβάνοντας από κάθε άτομο δυο μετρήσεις (“πριν-μετά”). Σε κάθε περίπτωση θα έχουμε n ζεύγη ατόμων.

Η παράμετρος που εκφράζει την διαφορά των θεραπειών (ή την επίδραση της θεραπείας όταν οι μετρήσεις είναι “πριν” και “μετά”) είναι η $\theta = \mu_T - \mu_C$ όπου μ_T είναι η μέση τιμή της θεραπείας που εξετάζουμε και μ_C είναι η μέση τιμή της θεραπείας *control* ή μ_T είναι η μέση τιμή πριν τη θεραπεία και μ_C είναι η μέση τιμή μετά τη θεραπεία. Η παράμετρος θ

εκτιμάται αμερόληπτα από την $\hat{\theta} = \bar{y}_T - \bar{y}_C$, όπου $\bar{y}_T = \frac{1}{n} \sum_{j=1}^n y_{jT}$ είναι η δειγματική μέση

τιμή των μετρήσεων πριν τη θεραπεία και $\bar{y}_C = \frac{1}{n} \sum_{j=1}^n y_{jC}$ είναι η δειγματική μέση τιμή των

μετρήσεων μετά τη θεραπεία.

Η εκτιμώμενη διασπορά του εκτιμητή $\hat{\theta}$ είναι ίση με $Var(\hat{\theta}) = \frac{s_D^2}{n}$ όπου

$$s_D^2 = \frac{1}{n-1} \sum_{j=1}^n (D_j - \hat{\theta})^2 \quad \text{και} \quad D_j = y_{jT} - y_{jC}, \quad j=1, 2, \dots, n.$$

3.3 Δίτιμη μεταβλητή απόκρισης (*Binary data*)

Έστω δυο ανεξάρτητες ομάδες ασθενών, η ομάδα στην οποία εφαρμόζεται η θεραπευτική αγωγή που μας ενδιαφέρει (*treatment*) και η ομάδα στην οποία εφαρμόζεται η θεραπεία σύγκρισης (*control*). Μια δίτιμη μεταβλητή παίρνει δυο τιμές που συχνά αναφέρονται ως “επιτυχία” και “αποτυχία”. Έστω p_{iT} η πιθανότητα επιτυχίας στην θεραπευτική αγωγή ενός ατόμου το οποίο ανήκει στην ομάδα θεραπείας (*treatment*) και p_{iC} η αντίστοιχη πιθανότητα ενός ατόμου το οποίο ανήκει στην ομάδα σύγκρισης (*control*). Αν γνωρίζουμε το σύνολο των ασθενών των δυο ομάδων, n_{iT} και n_{iC} , και τον αριθμό των ασθενών στις δυο ομάδες με θετικό αποτέλεσμα στην θεραπεία, τότε έχουμε επαρκή πληροφορία για να εκτιμήσουμε την επίδραση της θεραπείας.

Έστω ότι τα δεδομένα της i μελέτης συνοψίζονται σε πίνακα της μορφής:

| | Επιτυχία | Αποτυχία | Σύνολο |
|-------------------------------|----------|----------|-------------------------|
| Treated group | a_i | b_i | n_{iT} |
| Control group | c_i | d_i | n_{iC} |
| Σύνολο ασθενών i μελέτης | | | $N_i = n_{iT} + n_{iC}$ |

Πίνακας 3.2 Συνοπτικά στατιστικά για δίτιμες μεταβλητές

Στις κλινικές μελέτες, όταν η αποκριτική μεταβλητή είναι δίτιμη, χρησιμοποιούνται κυρίως τρία μέτρα για την σύγκριση των θεραπειών: η διαφορά των ποσοστών επιτυχίας των δυο ομάδων $p_{iT} - p_{iC}$, ο λόγος των ποσοστών επιτυχίας $\frac{p_{iT}}{p_{iC}}$ (*relative risk*) και ο λόγος των σχετικών πιθανοτήτων $\frac{p_{iT}/1-p_{iT}}{p_{iC}/1-p_{iC}}$ (*odds ratio*).

3.3.1 Διαφορά ποσοστών (*Difference of proportions*)

Ας υποθέσουμε ότι στη μελέτη i θέλουμε να εκτιμήσουμε την διαφορά των ποσοστών επιτυχίας των δυο ομάδων, δηλαδή θέλουμε να εκτιμήσουμε το $\theta_i = p_{iT} - p_{iC}$. Ο εκτιμητής μέγιστης πιθανοφάνειας (MLE) του θ_i είναι ο $\hat{\theta}_i = \frac{a_i}{n_{iT}} - \frac{c_i}{n_{iC}}$.

Η ασυμπτωτική διακύμανση του εκτιμητή $\hat{\theta}_i$ είναι ίση με:

$$Var(\hat{\theta}) = \frac{p_{iT} \cdot (1 - p_{iT})}{n_{iT}} + \frac{p_{iC} \cdot (1 - p_{iC})}{n_{iC}}$$

Επειδή $\hat{p}_{iT} = \frac{a_i}{n_{iT}}$ και $\hat{p}_{iC} = \frac{c_i}{n_{iC}}$, ο εκτιμητής της ασυμπτωτικής διακύμανσης είναι:

$$Var(\hat{\theta}_i) = \frac{a_i \cdot b_i}{n_{iT}^3} + \frac{c_i \cdot d_i}{n_{iC}^3}$$

όπου $b_i = n_{iT} - a_i$ και $d_i = n_{iC} - c_i$ είναι το πλήθος των αποτυχιών στις αντίστοιχες ομάδες *treatment* και *control*.

Αν έχουμε ορίσει ως επιτυχία ένα ανεπιθύμητο συμβάν όπως για παράδειγμα την εμφάνιση καρκίνου, η διαφορά των ποσοστών αναφέρεται στην βιβλιογραφία και ως “*risk difference*”.

Η επίδραση θ_i της θεραπείας σε σχέση με το *control*, παίρνει τιμές στο συμμετρικό διάστημα $[-1,1]$, ενώ αν ισχύει η υπόθεση ότι $\theta_i = 0$, τότε στην i μελέτη οι δυο θεραπείες είναι ισοδύναμες.

3.3.2 Λόγος ποσοστών (*Relative risk*)

Ας υποθέσουμε ότι στην μελέτη i θέλουμε να εκτιμήσουμε τον σχετικό κίνδυνο των δυο ομάδων, δηλαδή θέλουμε να εκτιμήσουμε το $\theta_i = \frac{p_{iT}}{p_{iC}}$. Ο εκτιμητής μέγιστης πιθανοφάνειας

(MLE) του σχετικού κινδύνου είναι ο $\hat{\theta}_i = \frac{a_i/n_{iT}}{c_i/n_{iC}}$ ή $\hat{\theta}_i = \frac{a_i \cdot n_{iC}}{c_i \cdot n_{iT}}$.

Επειδή ο λογάριθμος του σχετικού κινδύνου, συγκλίνει πιο γρήγορα στην κανονική κατανομή, συνήθως το μέτρο σύγκρισης που χρησιμοποιούμε είναι το $\theta'_i = \log\left(\frac{p_{iT}}{p_{iC}}\right)$ με αντίστοιχο εκτιμητή τον $\hat{\theta}'_i = \log\left(\frac{a_i \cdot n_{iC}}{c_i \cdot n_{iT}}\right)$.

Η ασυμπτωτική διακύμανση του εκτιμητή $\hat{\theta}'_i$ με χρήση της μεθόδου Δέλτα [βλ. *Agresti*, 2002, σελ. 73] είναι ίση με:

$$Var(\hat{\theta}'_i) = \frac{(1-p_{iT})}{p_{iT} \cdot n_{iT}} + \frac{(1-p_{iC})}{p_{iC} \cdot n_{iC}}$$

Επειδή $\hat{p}_{iT} = \frac{a_i}{n_{iT}}$ και $\hat{p}_{iC} = \frac{c_i}{n_{iC}}$, ο εκτιμητής της ασυμπτωτικής διακύμανσης είναι ίσος με:

$$\widehat{Var}(\hat{\theta}'_i) = \frac{b_i}{a_i \cdot n_{iT}} + \frac{d_i}{c_i \cdot n_{iC}} = \frac{1}{a_i} - \frac{1}{n_{iT}} + \frac{1}{c_i} - \frac{1}{n_{iC}}$$

όπου $b_i = n_{iT} - a_i$ και $d_i = n_{iC} - c_i$ είναι το πλήθος των αποτυχιών στις αντίστοιχες ομάδες *treatment* και *control*.

Αν ισχύει η υπόθεση $\theta'_i = 0$ (ή αντίστοιχα $\theta_i = 1$) τότε οι δυο θεραπείες είναι ισοδύναμες, ενώ αν $\theta' > 1$ (ή αντίστοιχα $\theta > 1$) η θεραπεία που εξετάζουμε στη μελέτη i έχει μεγαλύτερο ποσοστό επιτυχίας από το *control*.

3.3.3 Λόγος σχετικών πιθανοτήτων (*Odds ratio*)

Ας υποθέσουμε ότι θέλουμε να εκτιμήσουμε τον λόγο των σχετικών πιθανοτήτων των δυο ομάδων, δηλαδή θέλουμε να εκτιμήσουμε το $\theta_i = \frac{p_{iT}/1-p_{iT}}{p_{iC}/1-p_{iC}}$. Ο εκτιμητής μέγιστης πιθανοφάνειας (MLE) του λόγου των σχετικών πιθανοτήτων είναι ο $\hat{\theta}_i = \frac{a_i \cdot d_i}{c_i \cdot b_i}$, όπου $b_i = n_{iT} - a_i$ και $d_i = n_{iC} - c_i$ είναι το πλήθος των αποτυχιών στις αντίστοιχες ομάδες *treatment* και *control* αντίστοιχα.

Όπως συμβαίνει και με τον σχετικό κίνδυνο, έτσι και ο λογάριθμος του λόγου των σχετικών πιθανοτήτων συγκλίνει πιο γρήγορα στην κανονική κατανομή. Επομένως το μέτρο σύγκρισης των θεραπειών που χρησιμοποιείται συνήθως είναι το $\theta'_i = \log\left(\frac{p_{iT}/1-p_{iT}}{p_{iC}/1-p_{iC}}\right)$ με αντίστοιχο εκτιμητή τον $\hat{\theta}'_i = \log\left(\frac{a_i \cdot d_i}{c_i \cdot b_i}\right)$.

Η ασυμπτωτική διακύμανση του εκτιμητή $\hat{\theta}'_i$ με χρήση της μεθόδου Δέλτα [βλ. *Agresti*, 2002, σελ. 76] είναι ίση με:

$$Var(\hat{\theta}'_i) = \frac{1}{n_{iT} \cdot p_{iT}} + \frac{1}{n_{iC} \cdot p_{iC}} + \frac{1}{n_{iT} \cdot (1-p_{iT})} + \frac{1}{n_{iC} \cdot (1-p_{iC})}$$

Επειδή $\hat{p}_{iT} = \frac{a_i}{n_{iT}}$ και $\hat{p}_{iC} = \frac{c_i}{n_{iC}}$, ο εκτιμητής της ασυμπτωτικής διακύμανσης του εκτιμητή

$\hat{\theta}_i'$ είναι ίσος με:

$$\widehat{Var}(\hat{\theta}_i') = \frac{1}{a_i} + \frac{1}{c_i} + \frac{1}{b_i} + \frac{1}{d_i}$$

Αν ισχύει η υπόθεση $\theta_i' = 0$ (ή αντίστοιχα $\theta_i = 1$) τότε στην μελέτη i οι δυο θεραπείες είναι ισοδύναμες. Εδώ θα πρέπει να προσθέσουμε ότι αν $b_i = 0$ ή $c_i = 0$ τότε δεν μπορεί να υπολογιστεί το *odds ratio*, ενώ για να υπολογίσουμε την αντίστοιχη διακύμανση θα πρέπει να είναι $a_i, b_i, c_i, d_i \neq 0$. Μια καλή πρακτική είναι να προσθέσουμε το 0.5 σε κάθε κελί πριν προχωρήσουμε στην ανάλυση.

3.3.4 Σύγκριση των μέτρων για δίτιμα δεδομένα

Η χρήση της διαφοράς των ποσοστών για την σύγκριση των δυο θεραπειών έχει το πλεονέκτημα ότι μπορεί να ερμηνευτεί εύκολα. Συγκλίνει στην κανονική κατανομή για μέτρια μεγέθη δειγμάτων [Normand, 1999] και παίρνει τιμές στο συμμετρικό διάστημα $[-1, 1]$.

Ο λόγος των σχετικών πιθανοτήτων (*odds ratio*) και ο σχετικός κίνδυνος (*relative risk*) χρησιμοποιούνται περισσότερο στην λογαριθμική κλίμακα, λόγω του ότι οι δειγματικές τους κατανομές συγκλίνουν πιο γρήγορα στην κανονική κατανομή και επιπλέον η κλίμακα των τιμών τους γίνεται συμμετρική. Το *odds ratio* όμως υπερέχει σε σχέση με το *relative risk* γιατί αν αντί της πιθανότητας επιτυχίας των ατόμων στην θεραπεία θεωρήσουμε την πιθανότητα αποτυχίας, το *odds ratio* δίνει συνάφεια ίδιας έντασης με αντίθετο πρόσημο, ενώ το *relative risk* δίνει συνάφεια διαφορετικής έντασης και αντίθετου πρόσημου. Όταν όμως η εμφάνιση του γεγονότος που εξετάζουμε είναι σπάνια, τα δυο μέτρα ταυτίζονται. Η ερμηνεία του *odds ratio* είναι όμως πιο δύσκολη από την ερμηνεία του *relative risk*. Σε αρκετές κλινικές μελέτες συχνά χρησιμοποιείται ο όρος *relative risk* ενώ γίνεται χρήση του *odds ratio*, γι' αυτό και θα πρέπει να είμαστε ιδιαίτερα προσεκτικοί τόσο στην ανάγνωση των κλινικών μελετών όσο και στην ερμηνεία των αποτελεσμάτων.

Άλλος ένας λόγος που χρησιμοποιούνται περισσότερο το *relative risk* και το *odds ratio* για να περιγράψουν την επίδραση της θεραπείας σε μια μετανάλυση είναι γιατί μετα από εμπειρικές παρατηρήσεις διαπιστώθηκε ότι με τα μέτρα αυτά είναι λιγότερο πιθανό να εντοπιστεί ετερογένεια μεταξύ των μελετών [Deeks, 1997]. Για το θέμα της ετερογένειας και την επίδραση που έχει σε αυτή η επιλογή του μέτρου επίδρασης της θεραπείας θα μιλήσουμε με περισσότερες λεπτομέρειες στο 5^ο κεφάλαιο.

3.3.5 Σύγκριση ποσοστών κατά ζεύγη (*Paired Proportions*)

Αρκετά συχνά στις κλινικές μελέτες έχουμε ένα πληθυσμό ασθενών τα μέλη του οποίου υποβάλλονται και στις δυο θεραπείες με σκοπό τη σύγκριση της αποτελεσματικότητας των δυο θεραπειών. Συνήθως τα δεδομένα συνοψίζονται σε πίνακα της μορφής:

| 1 ^η Θεραπεία | 2 ^η Θεραπεία | | Σύνολο |
|-------------------------|-------------------------|----------|----------|
| | Επιτυχία | Αποτυχία | |
| Επιτυχία | n_{SS} | n_{SF} | n_{S+} |
| Αποτυχία | n_{FS} | n_{FF} | n_{F+} |
| Σύνολο | n_{+S} | n_{+F} | n |

Πίνακας 3.3 Συνοπτικά στατιστικά για ποσοστά κατά ζεύγη

Έστω p_{S+} η πιθανότητα επιτυχίας της πρώτης θεραπείας και p_{+S} η πιθανότητα επιτυχίας της δεύτερης θεραπείας. Παρατηρούμε ότι οι δυο πιθανότητες είναι συσχετισμένες. Η παράμετρος $\theta = p_{S+} - p_{+S}$ εκφράζει την διαφορά των δυο θεραπειών και εκτιμάται αμερόληπτα από $\hat{\theta} = \frac{n_{S+}}{n} - \frac{n_{+S}}{n}$.

Η διακύμανση της παραμέτρου θ δίνεται από την σχέση [βλ. Agresti, 2002, σελ. 411]:

$$Var(\hat{\theta}) = \frac{1}{n} \{ p_{S+}(1-p_{S+}) + p_{+S}(1-p_{+S}) - 2(p_{SS}p_{FF} - p_{SF}p_{FS}) \}$$

Ο εκτιμητής της διακύμανσης $Var(\hat{\theta})$ είναι ίσος με:

$$\hat{Var}(\hat{\theta}) = \frac{1}{n} \left\{ \left(\hat{p}_{SF} + \hat{p}_{FS} \right) - \left(\hat{p}_{SF} - \hat{p}_{FS} \right)^2 \right\} \text{ όπου } \hat{p}_{SF} = \frac{n_{SF}}{n} \text{ και } \hat{p}_{FS} = \frac{n_{FS}}{n}.$$

Αν η υπόθεση $p_{S+} = p_{+S}$ ή $\theta = 0$ είναι αληθής, τότε οι δυο θεραπείες είναι ισοδύναμες. Για $I \times I$ η περίπτωση αυτή είναι γνωστή ως *περιθώρια ομοιογένεια* (*marginal homogeneity*), ενώ για $I = 2$ είναι το γνωστό τεστ *Mc Nemar*.

ΚΕΦΑΛΑΙΟ 4

Συνδυασμός των εκτιμητών

4.1 Εισαγωγή

Έστω ότι σε κάθε μια από τις k μελέτες έχουμε υπολογίσει τους εκτιμητές των παραμέτρων θ_i , $i = 1, 2, \dots, k$, οι οποίοι εκφράζουν την επίδραση της θεραπείας σε σχέση με το *control* (3^ο κεφάλαιο). Στην συνέχεια θα πρέπει να συνδυάσουμε όλη αυτή την πληροφορία προκειμένου να υπολογίσουμε την ολική επίδραση θ της εξεταζόμενης θεραπείας. Ο εκτιμητής της ολικής επίδρασης της θεραπείας προκύπτει με βάση τον σταθμισμένο μέσο όρο των επιδράσεων της θεραπείας στις k μελέτες.

Στο παρόν κεφάλαιο θα παρουσιάσουμε τα δυο βασικά μοντέλα μετανάλυσης με τα οποία μπορούμε να συνδυάσουμε τις αρχικές μελέτες: το μοντέλο σταθερών επιδράσεων και το μοντέλο τυχαίων επιδράσεων. Στο μοντέλο σταθερών επιδράσεων υποθέτουμε ότι η επίδραση της θεραπείας είναι η ίδια σε όλες τις μελέτες, δηλαδή ισχύει $\theta_i = \theta$ για $i = 1, 2, \dots, k$, ενώ στο μοντέλο τυχαίων επιδράσεων η επίδραση της θεραπείας στις k μελέτες διαφέρει. Ο έλεγχος της υπόθεσης $H_0 : \theta_i = \theta, \forall i$ ονομάζεται *έλεγχος ετερογένειας* και θα παρουσιαστεί αναλυτικότερα στο 5^ο κεφάλαιο.

Στην περίπτωση του μοντέλου σταθερών επιδράσεων θα παρουσιαστούν τρεις μεθοδολογίες: τη μέθοδο των *Mantel-Haenszel*, τη μέθοδο του *Peto* και τη μέθοδο της αντίστροφης διακύμανσης (*Inverse variance method*). Στην περίπτωση του μοντέλου τυχαίων επιδράσεων θα παρουσιαστεί η μέθοδος των *DerSimonian* και *Laird*.

Σε κάθε μελέτη υποθέτουμε ότι έχουμε παράλληλο σχεδιασμό (*parallel design*) με δυο ομάδες ασθενών. Στην πρώτη ομάδα εφαρμόζεται η υπό εξέταση θεραπεία (ομάδα “*treatment*”) ενώ στη δεύτερη εφαρμόζεται η καθιερωμένη θεραπεία (ομάδα “*control*”). Όπως αναφέρθηκε και στο προηγούμενο κεφάλαιο, η επίδραση της θεραπείας θ_i θα πρέπει να εκφράζεται με το ίδιο μέτρο σε όλες τις μελέτες.

4.2 Μοντέλα σταθερών επιδράσεων - *Fixed Effects Models*

Χρησιμοποιώντας το μοντέλο σταθερών επιδράσεων για την εκτίμηση της ολικής επίδρασης της θεραπείας, υποθέτουμε ότι οι μελέτες είναι ομοιογενείς, δηλαδή $\theta_i = \theta$ για $i = 1, 2, \dots, k$. Αυτό σημαίνει ότι όλες οι μελέτες εκτιμάνε την ίδια επίδραση της θεραπείας. Επομένως η παρατηρούμενη διακύμανση μεταξύ των k εκτιμητών της θεραπείας οφείλεται αποκλειστικά στο σφάλμα δειγματοληψίας. Το μοντέλο επομένως δίνει απάντηση στην ερώτηση: “*Ποια είναι η επίδραση της θεραπείας στο σύνολο των μελετών που εξετάζουμε?*”

4.2.1 Μέθοδος Mantel-Haenszel (MH)

Η μέθοδος MH [Mantel, Haenszel, 1959] είναι ευρέως γνωστή για τον υπολογισμό ενός ολικού εκτιμητή όταν υπάρχουν k στρώματα (*strata*). Επειδή οι μελέτες που συμπεριλαμβάνονται στην μετανάλυση μπορούν να θεωρηθούν ως ανεξάρτητα στρώματα, η μέθοδος MH είναι κατάλληλη για την ανάλυση αυτών των δεδομένων, με την προϋπόθεση ότι η μεταβλητή απόκρισης που εξετάζουμε είναι δίτιμη. Θεωρούμε επομένως ότι τα δεδομένα συνοψίζονται στη μορφή του πίνακα 3.2 (σελ. 20).

Στην περίπτωση αυτή, η επίδραση θ_i της θεραπείας στην i μελέτη μπορεί να εκφραστεί με βάση την διαφορά των ποσοστών των δυο ομάδων, τον λόγο των ποσοστών επιτυχίας ή αποτυχίας (*relative risk*) ή τον λόγο των σχετικών πιθανοτήτων επιτυχίας (*odds ratio*). Η προσέγγιση MH μπορεί να εφαρμοστεί και με τα τρία προηγούμενα μέτρα.

Στον επόμενο πίνακα γίνεται μια περιγραφή των μέτρων αυτών, ενώ για περισσότερες λεπτομέρειες ο αναγνώστης παραπέμπεται στις αντίστοιχες παραγράφους του προηγούμενου κεφαλαίου.

| | Διαφορά ποσοστών | Λόγος ποσοστών | Odds ratio |
|---------------------|---|--|--|
| Παράμετρος | $\theta_i = p_{iT} - p_{iC}$ | $\theta_i = p_{iT} / p_{iC}$ | $\theta_i = \frac{p_{iT} / (1 - p_{iT})}{p_{iC} / (1 - p_{iC})}$ |
| Εκτιμητής | $\hat{\theta}_i = \frac{a_i}{n_{iT}} - \frac{c_i}{n_{iC}}$ | $\hat{\theta}_i = \frac{a_i \cdot n_{iC}}{c_i \cdot n_{iT}}$ | $\hat{\theta}_i = \frac{a_i \cdot d_i}{c_i \cdot b_i}$ |
| Δειγματική διασπορά | $\hat{Var}(\hat{\theta}_i) = \frac{a_i \cdot b_i}{n_{iT}^3} + \frac{c_i \cdot d_i}{n_{iC}^3}$ | $\hat{Var}(\log \hat{\theta}_i) = \frac{1}{a_i} - \frac{1}{n_{iT}} + \frac{1}{c_i} - \frac{1}{n_{iC}}$ | $\hat{Var}(\log \hat{\theta}_i) = \frac{1}{a_i} + \frac{1}{c_i} + \frac{1}{b_i} + \frac{1}{d_i}$ |

Πίνακας 4.1 Μέτρα επίδρασης της θεραπείας για δίτιμα δεδομένα

Ο ΜΗ εκτιμητής της ολικής επίδρασης της θεραπείας σε σχέση με το *control* δίνεται γενικά από την σχέση:

$$\hat{\theta}_{MH} = \frac{\sum_{i=1}^{\kappa} \hat{\theta}_i \cdot w_i}{\sum_{i=1}^{\kappa} w_i}$$

όπου $\hat{\theta}_i$ είναι ο εκτιμητής της παραμέτρου θ_i που προκύπτει από την i μελέτη και w_i είναι το αντίστοιχο βάρος της μελέτης.

Στην συνέχεια για κάθε μέτρο θα παρουσιάσουμε το αντίστοιχο βάρος w_i που δίνεται σε κάθε μελέτη. Επιπλέον μαζί με τον εκτιμητή της ολικής επίδρασης της θεραπείας, θα δοθεί η διακύμανσή του, ένα ασυμπτωτικό τεστ για τον έλεγχο της υπόθεσης $H_0 : \theta = 0$ καθώς και το αντίστοιχο διάστημα εμπιστοσύνης.

4.2.1.1 Διαφορά ποσοστών (*risk difference*)

Αν εκφράσουμε την επίδραση της θεραπείας σε σχέση με το *control* χρησιμοποιώντας την διαφορά των αντίστοιχων ποσοστών, $\theta = p_T - p_C$, το βάρος της i μελέτης δίνεται από την σχέση $w_i = \frac{n_{iT} \cdot n_{iC}}{N_i}$, όπου n_{iT} είναι το πλήθος των ασθενών της ομάδας θεραπείας, n_{iC} είναι το πλήθος των ασθενών της ομάδας *control* και $N_i = n_{iT} + n_{iC}$.

Με αντικατάσταση στον τύπο $\hat{\theta}_{MH}^{RD} = \frac{\sum_{i=1}^{\kappa} \hat{\theta}_i \cdot w_i}{\sum_{i=1}^{\kappa} w_i}$, βρίσκουμε τον ζητούμενο εκτιμητή της ολικής επίδρασης της θεραπείας σε σχέση με το *control*.

Η διακύμανση του εκτιμητή $\hat{\theta}_{MH}^{RD}$ δίνεται από την σχέση [Alderson, 2004]:

$$Var\left(\hat{\theta}_{MH}^{RD}\right) = \frac{\sum_{i=1}^{\kappa} \left(\alpha_i b_i n_{iC}^3 + c_i d_i n_{iT}^3\right) / \left(n_{iC} n_{iT} N_i^2\right)}{\left(\sum_{i=1}^{\kappa} \left(n_{iC} n_{iT} / N_i\right)\right)^2}.$$

Για τον έλεγχο της υπόθεσης $H_0 : \theta = 0$ μπορούμε να χρησιμοποιήσουμε το προσεγγιστικό στατιστικό:

$$Z = \frac{\hat{\theta}_{MH}^{RD}}{\sqrt{Var\left(\hat{\theta}_{MH}^{RD}\right)}} \stackrel{H_0}{\sim} N(0,1)$$

το οποίο κάτω από την μηδενική υπόθεση ακολουθεί τυπική κανονική κατανομή.

Το 95% ασυμπτωτικό διάστημα εμπιστοσύνης για την παράμετρο $\theta = p_T - p_C$ είναι το εξής:

$$\left(\hat{\theta}_{MH}^{RD} \pm z_{\alpha/2} \cdot \sqrt{Var\left(\hat{\theta}_{MH}^{RD}\right)} \right).$$

4.2.1.2 Λόγος ποσοστών (*Relative risk*)

Αν χρησιμοποιήσουμε ως μέτρο επίδρασης της θεραπείας τον λόγο των ποσοστών,

$$\theta = \frac{p_T}{p_C}, \text{ τότε σε κάθε μελέτη δίνεται βάρος ίσο με: } w_i = \frac{c_i(a_i + b_i)}{N_i}.$$

Με αντικατάσταση στον τύπο $\hat{\theta}_{MH}^{RR} = \frac{\sum_{i=1}^K \hat{\theta}_i \cdot w_i}{\sum_{i=1}^K w_i}$, βρίσκουμε τον ζητούμενο ολικό εκτιμητή

του λόγου των ποσοστών επιτυχίας μεταξύ των ομάδων “*treatment*” και “*control*”.

Η διακύμανση του λογαρίθμου του εκτιμητή $\hat{\theta}_{MH}^{RR}$ δίνεται από την σχέση:

$$Var\left(\log \hat{\theta}_{MH}^{RR}\right) = \frac{\sum_{i=1}^K (n_{iC}n_{iT}(a_i + c_i) - a_i c_i N_i) / N_i^2}{\left(\sum_{i=1}^K a_i n_{iC} / N_i\right) \cdot \left(\sum_{i=1}^K c_i n_{iT} / N_i\right)} \quad [\text{βλ. Deeks, 1999}]$$

Για τον έλεγχο της υπόθεσης $H_0: \log \theta = 0$ μπορούμε να χρησιμοποιήσουμε το προσεγγιστικό στατιστικό:

$$Z = \frac{\log \hat{\theta}_{MH}^{RR}}{\sqrt{Var\left(\log \hat{\theta}_{MH}^{RR}\right)}} \stackrel{H_0}{\sim} N(0,1)$$

το οποίο κάτω από την μηδενική υπόθεση ακολουθεί τυπική κανονική κατανομή.

Το 95% ασυμπτωτικό διάστημα εμπιστοσύνης για την παράμετρο $\theta = \frac{p_T}{p_C}$ είναι το εξής:

$$\exp\left\{\log \hat{\theta}_{MH}^{RR} \pm z_{\alpha/2} \cdot \sqrt{Var\left(\log \hat{\theta}_{MH}^{RR}\right)}\right\}.$$

4.2.1.3 Λόγος των σχετικών πιθανοτήτων (*Odds ratio*)

Αν χρησιμοποιήσουμε ως μέτρο επίδρασης της θεραπείας τον λόγο των σχετικών ποσοστών, $\theta = \frac{p_T / 1 - p_T}{p_C / 1 - p_C}$, τότε το βάρος της i μελέτης τίθεται ίσο με: $w_i = \frac{b_i c_i}{N_i}$.

Με αντικατάσταση στον τύπο $\hat{\theta}_{MH}^{OR} = \frac{\sum_{i=1}^K \hat{\theta}_i \cdot w_i}{\sum_{i=1}^K w_i}$, βρίσκουμε τον ζητούμενο ολικό εκτιμητή του λόγου των σχετικών ποσοστών επιτυχίας μεταξύ των ομάδων “*treatment*” και “*control*”.

Η διακύμανση του λογαρίθμου εκτιμητή $\hat{\theta}_{MH}^{OR}$ προτάθηκε από τους *Robins*, *Breslow* και *Greenland* [1986] και δίνεται από την σχέση:

$$Var\left(\log \hat{\theta}_{MH}^{OR}\right) = \frac{1}{2} \times \left\{ \frac{\sum_{i=1}^K (a_i + d_i) a_i d_i / N_i^2}{\left(\sum_{i=1}^K a_i d_i / N_i\right)^2} + \frac{\sum_{i=1}^K (b_i + c_i) b_i c_i / N_i^2}{\left(\sum_{i=1}^K b_i c_i / N_i\right)^2} + \frac{\sum_{i=1}^K (a_i + d_i) b_i c_i / N_i^2 + \sum_{i=1}^K (b_i + c_i) a_i d_i / N_i^2}{\left(\sum_{i=1}^K a_i d_i / N_i\right) \cdot \left(\sum_{i=1}^K b_i c_i / N_i\right)} \right\}$$

Για τον έλεγχο της υπόθεσης $H_0: \log \theta = 0$ μπορούμε να χρησιμοποιήσουμε το προσεγγιστικό στατιστικό:

$$Z = \frac{\log \hat{\theta}_{MH}^{OR}}{\sqrt{Var\left(\log \hat{\theta}_{MH}^{OR}\right)}} \stackrel{H_0}{\sim} N(0,1)$$

το οποίο κάτω από την μηδενική υπόθεση ακολουθεί τυπική κανονική κατανομή.

Το 95% ασυμπτωτικό διάστημα εμπιστοσύνης για την παράμετρο $\theta = \frac{p_T / 1 - p_T}{p_C / 1 - p_C}$ είναι το

εξής:

$$\exp \left\{ \log \hat{\theta}_{MH}^{OR} \pm z_{\alpha/2} \cdot \sqrt{Var \left(\log \hat{\theta}_{MH}^{OR} \right)} \right\}$$

4.2.2 Peto's Assumption Free Method

Η μέθοδος που θα παρουσιάσουμε σε αυτή τη παράγραφο, περιγράφηκε αρχικά από τους *Peto et al.* [1977] και αργότερα από τους *Yusuf et al.* [1985]. Η μέθοδος αυτή μπορεί να θεωρηθεί ως μια παραλλαγή της μεθόδου MH, αλλά εφαρμόζεται μόνο όταν χρησιμοποιούμε το *odds ratio* ως μέτρο επίδρασης της θεραπείας.

Έστω λοιπόν ότι η μετανάλυση συμπεριλαμβάνει k ανεξάρτητες μελέτες και ότι τα δεδομένα συνοψίζονται στη μορφή του πίνακα 3.2 (σελ. 20).

Η μέθοδος του *Peto* εκτιμά ασυμπτωτικά το *odds ratio* με την ποσότητα

$\hat{\theta}_i = \exp \left\{ (a_i - E[a_i]) / v_i \right\}$. Ο όρος $E[a_i] = \frac{n_{iT}(a_i + c_i)}{N_i}$ εκφράζει την αναμενόμενη τιμή της

παραμέτρου a_i (πλήθος επιτυχιών της ομάδας θεραπείας) και $v_i = \frac{n_{iT}n_{iC}(a_i + c_i)(b_i + d_i)}{N_i^2(N_i - 1)}$

είναι η αντίστοιχη υπεργεωμετρική διακύμανση.

Η μέθοδος του *Peto* προκύπτει ως εξής: αν στα δεδομένα του πίνακα 3.2 δεσμεύσουμε ως προς τον συνολικό αριθμό των επιτυχιών της ομάδας θεραπείας, τότε κάτω από την μηδενική υπόθεση ότι η επίδραση της θεραπείας είναι μηδενική ($\log \hat{\theta}_i = 0$), ο αριθμός των επιτυχιών της ομάδας θεραπείας a_i θα ακολουθεί υπεργεωμετρική κατανομή (βλ. *Whitehead* [2002], σελ. 29-30). Η προσέγγιση του *odds ratio* είναι ικανοποιητική όταν η επίδραση της θεραπείας είναι μικρή (κοντά στη μονάδα) και όταν στις δυο ομάδες ο αριθμός των γεγονότων είναι ίδιου μεγέθους [*Fleiss*, 1994].

Η ασυμπτωτική διακύμανση του λογαρίθμου του εκτιμητή $\hat{\theta}_i$ είναι ίση με:

$$Var\left(\log \hat{\theta}_i\right) = 1/v_i.$$

Ο ολικός εκτιμητής του λόγου των σχετικών πιθανοτήτων των δυο ομάδων και η αντίστοιχη διακύμανση του λογαρίθμου του εκτιμητή δίνονται από τις σχέσεις:

$$\hat{\theta}_p = \frac{\sum_{i=1}^{\kappa} \hat{\theta}_i \cdot w_i}{\sum_{i=1}^{\kappa} w_i} \quad \text{και} \quad Var\left(\log \hat{\theta}_p\right) = \frac{1}{\sum_{i=1}^{\kappa} v_i}.$$

Το βάρος w_i της κάθε μελέτης είναι ίσο με την υπεργεωμετρική διακύμανση v_i της παραμέτρου a_i , δηλαδή $w_i = v_i$.

Για τον έλεγχο της υπόθεσης $H_0: \log \theta = 0$ μπορούμε να χρησιμοποιήσουμε το

προσεγγιστικό στατιστικό:
$$Z = \frac{\log \hat{\theta}_p}{\sqrt{Var\left(\log \hat{\theta}_p\right)}} \stackrel{H_0}{\sim} N(0,1)$$

το οποίο κάτω από την μηδενική υπόθεση ακολουθεί τυπική κανονική κατανομή.

Το 95% ασυμπτωτικό διάστημα εμπιστοσύνης για την παράμετρο $\theta = \frac{p_T/1-p_T}{p_C/1-p_C}$ είναι το εξής:

$$\exp\left\{\log \hat{\theta}_p \pm z_{\alpha/2} \cdot \sqrt{Var\left(\log \hat{\theta}_p\right)}\right\}.$$

4.2.3 Inverse Variance-weighted method (W)

Στη συνέχεια θα παρουσιάσουμε την μέθοδο των αντίστροφων διακυμάνσεων η οποία προτάθηκε από τον *Cochran* τη δεκαετία του '30. Η μέθοδος αυτή είναι αρκετά απλή και μπορεί να εφαρμοστεί για όλα τα είδη των μεταβλητών απόκρισης (δίτιμες, συνεχείς, διατακτικές κλπ). Το όνομά της προκύπτει από το γεγονός ότι το βάρος που δίνεται σε κάθε

μελέτη είναι ίσο με τον αντίστροφο της διακύμανσης της επίδρασης της θεραπείας. Όσο πιο μεγάλη σε μέγεθος είναι μια μελέτη τόσο μεγαλύτερο θα είναι το βάρος που της δίνεται, αφού το τυπικό σφάλμα του εκτιμητή θα είναι μικρό. Τα μόνα στοιχεία επομένως που χρειάζονται για την εφαρμογή της μεθόδου είναι ο εκτιμητής της επίδρασης της θεραπείας και η αντίστοιχη διακύμανσή του.

Έστω λοιπόν ότι έχουμε k ανεξάρτητες μελέτες. Σε κάθε μια από τις μελέτες συγκρίνονται δυο ομάδες: η ομάδα στην οποία εφαρμόζεται η θεραπευτική αγωγή η οποία μας ενδιαφέρει (T) και η ομάδα στην οποία εφαρμόζεται η θεραπεία σύγκρισης-*control* (C). Υποθέτουμε ότι η μεταβλητή απόκρισης που εξετάζουμε είναι η ίδια σε όλες τις μελέτες.

Θεωρούμε ότι η παράμετρος θ_i εκφράζει την επίδραση της θεραπείας σε σχέση με το *control* στην i μελέτη. Για δίτιμα δεδομένα το θ_i , $i = 1, 2, \dots, k$ μπορεί να είναι η διαφορά δυο ποσοστών, ο σχετικός κίνδυνος, το *odds ratio* ενώ για συνεχή δεδομένα θα είναι αντίστοιχα η διαφορά μέσων τιμών ή η τυποποιημένη διαφορά μέσων τιμών. Βασική προϋπόθεση είναι ότι αυτή η παράμετρος θα πρέπει να ακολουθεί, έστω και ασυμπτωτικά, κανονική κατανομή.

Έστω $\hat{\theta}_i$ ο εκτιμητής του θ_i που προκύπτει από την i μελέτη και w_i ο αντίστροφος της ασυμπτωτικής διακύμανσης του $\hat{\theta}_i$. Οι τιμές των w_i θεωρούνται γνωστές για όλες τις μελέτες.

Επειδή στο μοντέλο σταθερών επιδράσεων έχουμε υποθέσει ότι $\theta_i = \theta$ για κάθε $i = 1, 2, \dots, k$, μπορούμε να το γράψουμε στην αλγεβρική μορφή:

$$\hat{\theta}_i = \theta + e_i, \text{ για } i = 1, 2, \dots, k$$

Τα σφάλματα e_i ακολουθούν κανονική κατανομή με μέση τιμή μηδέν και διακύμανση σ_i^2 , δηλαδή ισχύει $e_i \sim N(0, \sigma_i^2)$, όπου σ_i^2 είναι η κοινή διακύμανση των ομάδων θεραπείας εντός της i μελέτης.

Επομένως ασυμπτωτικά για κάθε μελέτη θα ισχύει η σχέση:

$$\hat{\theta}_i \sim N(\theta, \sigma_i^2), i = 1, 2, \dots, k \quad (4.1)$$

Στην συνέχεια υποθέτουμε ότι η εκτιμημένη διασπορά των $\hat{\theta}_i$, $Var(\hat{\theta}_i)$, είναι ίση με την πραγματική διασπορά σ_i^2 των σφαλμάτων του μοντέλου.

Αν $w_i = \frac{1}{Var(\hat{\theta}_i)}$, η σχέση (4.1) παίρνει τη μορφή:

$$\hat{\theta}_i \sim N\left(\theta_i, \frac{1}{w_i}\right) \Rightarrow \hat{\theta}_i \cdot w_i \sim N(\theta_i \cdot w_i, w_i), \text{ για } i=1,2,\dots,k \quad (4.2)$$

Η προηγούμενη υπόθεση είναι αληθής όταν το μέγεθος του δείγματος της κάθε μελέτης είναι μεγάλο. Σε διαφορετική περίπτωση η εκτίμηση της διακύμανσης εμπεριέχει σφάλμα.

Η σχέση (4.2) κάτω από την μηδενική υπόθεση ότι στην i μελέτη η επίδραση της θεραπείας είναι μηδέν, $H_{0i} : \theta_i = 0$ δίνει:

$$\hat{\theta}_i \cdot w_i \sim N(0, w_i), \quad i=1,2,\dots,k.$$

Κάτω από την συνδυασμένη μηδενική υπόθεση ότι σε όλες τις μελέτες η επίδραση της θεραπείας είναι μηδέν, $H_0 : \theta_1 = \theta_2 = \dots = \theta_k = 0$, και με βάση το γεγονός ότι οι εκτιμητές $\hat{\theta}_i$, $i=1,2,\dots,k$ είναι ανεξάρτητοι προκύπτει ότι :

$$\sum_{i=1}^k \hat{\theta}_i \cdot w_i \sim N\left(0, \sum_{i=1}^k w_i\right), \quad i=1,2,\dots,k \quad (4.3)$$

Από την σχέση (4.3) προκύπτει το στατιστικό U , το οποίο κάτω από την μηδενική υπόθεση ότι η επίδραση της θεραπείας είναι μηδέν σε όλες τις μελέτες, ακολουθεί χ^2 κατανομή με 1 βαθμό ελευθερίας.

$$U = \frac{\left(\sum_{i=1}^k \hat{\theta}_i \cdot w_i\right)^2}{\sum_{i=1}^k w_i} \underset{H_0}{\sim} \chi_1^2$$

Υποθέτοντας ότι ισχύει η μηδενική υπόθεση $H_0 : \theta_1 = \theta_2 = \dots = \theta_k = \theta$, η σχέση (4.2) μετατρέπεται:

$$\sum_{i=1}^k \hat{\theta}_i \cdot w_i \sim N \left(\theta \sum_{i=1}^k w_i, \sum_{i=1}^k w_i \right), \quad i = 1, 2, \dots, k$$

και επομένως η ολική επίδραση θ της θεραπείας μπορεί να εκτιμηθεί με βάση τον τύπο

$$\hat{\theta} = \frac{\sum_{i=1}^k \hat{\theta}_i \cdot w_i}{\sum_{i=1}^k w_i}$$

Η διακύμανση του εκτιμητή $\hat{\theta}$ είναι ίση με: $Var(\hat{\theta}) = \frac{1}{\sum_{i=1}^k w_i}$

Οι Li *et al* [1994] έδειξαν ότι η διακύμανση του ολικού εκτιμητή $\hat{\theta}$ είναι μερικές φορές μεροληπτική και ευαίσθητη στην μικρότερη τιμή των διακυμάνσεων των εκτιμητών $\hat{\theta}_i$. Αν επομένως η μικρότερη διακύμανση έχει τόσο μικρή τιμή από λάθος, τότε η επίδρασή της στην διακύμανση του ολικού εκτιμητή είναι μεγάλη και οδηγεί στην υποεκτίμηση της πραγματικής τιμής.

Εδώ θα πρέπει να σημειώσουμε ότι αν τα βάρη w_i είναι οι αντίστροφοι των πραγματικών διασπορών των $\hat{\theta}_i$ κι όχι οι εκτιμητές τους, τότε ο εκτιμητής $\hat{\theta}$ είναι ίσος με τον ΕΜΠ του θ [βλ. Whitehead, 2002].

Το 95% ασυμπτωτικό διάστημα εμπιστοσύνης για την παράμετρο θ είναι:

$$\left(\hat{\theta} \pm z_{\alpha/2} \cdot \sqrt{\frac{1}{\sum_{i=1}^k w_i}} \right).$$

4.2.4 Σύγκριση των μεθόδων σταθερής επίδρασης

Για να μπορέσουμε να συγκρίνουμε τις τρεις μεθοδολογίες σταθερών επιδράσεων θεωρούμε ότι η μεταβλητή απόκρισης είναι δίτιμη και το μέτρο που εκφράζει την επίδραση της θεραπείας σε σχέση με το *control* είναι ο λόγος των σχετικών πιθανοτήτων (*odds ratio*). Γενικά αν τα μεγέθη των δειγμάτων σε κάθε μελέτη είναι μεγάλα, για παράδειγμα αν σε κάθε κελί του 2x2 πίνακα συνάφειας υπάρχουν περισσότερες από πέντε παρατηρήσεις, τότε και οι τρεις μεθοδολογίες δίνουν παρόμοια αποτελέσματα.

Στη συνέχεια θα αναφέρουμε κάποια θετικά και αρνητικά σημεία των τριών μεθοδολογιών τα οποία θα πρέπει να προσέξουμε για να αποφασίσουμε ποια μέθοδο θα πρέπει να χρησιμοποιήσουμε.

- **Mantel-Haenszel μέθοδος:**

- i. Είναι υπολογιστικά εύκολη μέθοδος.
- ii. Δίνει ακριβή αποτελέσματα αν ο αριθμός των μελετών που πρόκειται να συνδυαστούν είναι μεγάλος αλλά το μέγεθος δείγματος σε κάθε μελέτη είναι μικρό [Fleiss, 1981].
- iii. Δεν μπορεί να εφαρμοστεί όταν υπάρχουν κενά κελιά στον 2x2 πίνακα εκτός και αν χρησιμοποιηθεί διόρθωση συνέχειας.

- **Μέθοδος του Peto's :**

- i. Είναι πιο βελτιωμένη από την μέθοδο MH, γιατί μπορεί να εφαρμοστεί ακόμα και όταν υπάρχουν κενά κελιά στους πίνακες.
- ii. Δίνει ικανοποιητικά αποτελέσματα όταν το γεγονός είναι σπάνιο.
- iii. Όταν τα *odds ratio* είναι κοντά στη μονάδα ο ολικός εκτιμητής υποεκτιμά την πραγματική τιμή [Fleiss, 1995].
- iv. Οι εκτιμητές των *odds ratios* δεν είναι αμερόληπτοι όταν ο αριθμός των ασθενών στις δυο ομάδες δεν είναι ισορροπημένος (*balanced*) [Greenland, Salvan 1990].

- **Μέθοδος της αντίστροφης διακύμανσης :**
 - i. Είναι πολύ εύκολη στον υπολογισμό.
 - ii. Δίνει ακριβή αποτελέσματα αν το πλήθος των μελετών που πρόκειται να συνδυαστούν είναι μικρό, αλλά το μέγεθος δείγματος σε κάθε μελέτη είναι μεγάλο [Fleiss, 1981].

4.3 Μοντέλο Τυχαίων επιδράσεων – *Random Effects Model*

Το μοντέλο τυχαίων επιδράσεων παρουσιάστηκε από τους *DerSimonian* και *Laird* το 1986. Η βασική υπόθεση του μοντέλου είναι ότι οι επιδράσεις της θεραπείας στις k ανεξάρτητες μελέτες, θ_i , $i = 1, 2, \dots, k$, είναι ένα τυχαίο δείγμα ανεξάρτητων παρατηρήσεων οι οποίες προέρχονται από την κανονική κατανομή με μέση τιμή θ και διακύμανση τ^2 .

Στο μοντέλο τυχαίων επιδράσεων έχουμε δυο στάδια δειγματοληψίας: Αρχικά παίρνουμε ένα δείγμα από μελέτες από τον πληθυσμό όλων των πιθανών μελετών (αυτών που έχουν ήδη πραγματοποιηθεί αλλά και αυτών που θα πραγματοποιηθούν στο μέλλον) στον οποίο η μέση επίδραση της θεραπείας είναι θ . Στην συνέχεια παίρνουμε δείγμα παρατηρήσεων από την i μελέτη στην οποία η επίδραση της θεραπείας είναι θ_i . Στο μοντέλο αυτό λοιπόν θεωρούμε ότι οι μελέτες που εξετάζουμε προέρχονται από έναν υπερπληθυσμό μελετών. Έτσι μπορούμε να πούμε ότι δίνει απάντηση στο ερώτημα: “*Ποιος είναι ο καλύτερος εκτιμητής της πραγματικής επίδρασης της θεραπείας?*”

Αν $\hat{\theta}_i$ είναι ο εκτιμητής της παραμέτρου θ_i για την i μελέτη, τότε το μοντέλο τυχαίων επιδράσεων δίνεται μαθηματικά από τη σχέση: $\hat{\theta}_i = \theta + \nu_i + e_i$, $i = 1, 2, \dots, k$ όπου $\nu_i \sim N(0, \tau^2)$ και $e_i \sim N(0, \sigma_i^2)$. Οι τυχαίοι παράγοντες ν_i και e_i είναι μεταξύ τους ανεξάρτητοι. Επομένως η παράμετρος τ^2 εκφράζει την μεταβλητότητα μεταξύ των μελετών, ενώ το σ_i^2 είναι η μεταβλητότητα εντός της i μελέτης.

Με βάση τις προηγούμενες σχέσεις θα ισχύει $\hat{\theta}_i \sim N(\theta, \sigma_i^2 + \tau^2)$, για $i = 1, 2, \dots, k$.

Οι διασπορές σ_i^2 θεωρούνται ίσες με τις δειγματικές διακυμάνσεις $\hat{Var}(\hat{\theta}_i)$ και οι οποίες είναι γνωστές για κάθε μελέτη. Όπως αναφέρθηκε και στο μοντέλο σταθερών επιδράσεων, η υπόθεση αυτή είναι αληθής όταν το μέγεθος του δείγματος της κάθε μελέτης είναι μεγάλο. Σε διαφορετική περίπτωση η εκτίμηση της διακύμανσης εμπεριέχει σφάλμα.

Αν θεωρήσουμε ότι τα βάρη είναι ίσα με $w_i = \frac{1}{\hat{Var}(\hat{\theta}_i)}$ τότε η προηγούμενη σχέση

μετατρέπεται στην :

$$\hat{\theta}_i \sim N\left(\theta, \frac{1}{w_i} + \tau^2\right), \quad i = 1, 2, \dots, k.$$

Ο εκτιμητής “σταθερών επιδράσεων” $\hat{\theta} = \frac{\sum_{i=1}^k \hat{\theta}_i \cdot w_i}{\sum_{i=1}^k w_i}$ εξακολουθεί να είναι αμερόληπτος

εκτιμητής του θ , αλλά η διακύμανσή του είναι ίση με :

$$Var(\hat{\theta}) = \frac{\sum_i w_i^2 \cdot Var(\hat{\theta}_i)}{\left(\sum_i w_i\right)^2} = \frac{\sum_i w_i^2 \cdot \left(\frac{1}{w_i} + \tau^2\right)}{\left(\sum_i w_i\right)^2} \Rightarrow Var(\hat{\theta}) = \frac{1}{\sum_i w_i} + \frac{\tau^2 \sum_i w_i^2}{\left(\sum_i w_i\right)^2}.$$

Η διακύμανση τ^2 του πληθυσμού των μελετών, είναι άγνωστη παράμετρος και θα πρέπει να εκτιμηθεί από τα δεδομένα.

Αν θεωρήσουμε ότι το $\hat{\tau}^2$ είναι ο ζητούμενος εκτιμητής τότε μπορούμε να χρησιμοποιήσουμε το ασυμπτωτικό αποτέλεσμα:

$$\hat{\theta}_i \sim N\left(\theta, \frac{1}{w_i} + \hat{\tau}^2\right), \quad i = 1, 2, \dots, k$$

$$\text{ή} \quad \hat{\theta}_i \sim N\left(\theta, \frac{1}{w_i^*}\right), \quad i = 1, 2, \dots, k \quad \text{όπου} \quad w_i^* = \frac{1}{w_i^{-1} + \hat{\tau}^2}.$$

Επειδή η προηγούμενη σχέση δίνει $\hat{\theta}_i \cdot w_i^* \sim N(\theta \cdot w_i^*, w_i^*)$, κάτω από την μηδενική υπόθεση ότι η ολική επίδραση της θεραπείας είναι μηδέν, $H_0 : \theta = 0$, θα είναι

$$\hat{\theta}_i \cdot w_i^* \sim N(0, w_i^*).$$

Οι εκτιμητές $\hat{\theta}_i, i = 1, 2, \dots, k$ είναι ανεξάρτητοι επομένως προκύπτει ότι :

$$\sum_{i=1}^k \hat{\theta}_i \cdot w_i^* \sim N\left(0, \sum_{i=1}^k w_i^*\right), i = 1, 2, \dots, k.$$

Από την τελευταία σχέση προκύπτει το στατιστικό U^* το οποίο ακολουθεί χ^2 κατανομή με 1 βαθμό ελευθερίας.

$$U^* = \frac{\left(\sum_{i=1}^k \hat{\theta}_i \cdot w_i^*\right)^2}{\sum_{i=1}^k w_i^*} \underset{H_0}{\sim} \chi_1^2.$$

Ένας νέος αμερόληπτος εκτιμητής του θ , ο οποίος έχει μικρότερη διακύμανση από τον αντίστοιχο εκτιμητή σταθερών επιδράσεων, είναι ο εξής:

$$\hat{\theta}^* = \frac{\sum_{i=1}^k \hat{\theta}_i \cdot w_i^*}{\sum_{i=1}^k w_i^*}, \text{ όπου } w_i^* = \frac{1}{w_i^{-1} + \hat{\tau}^2} \text{ με } \text{Var}(\hat{\theta}^*) = \frac{1}{\sum_{i=1}^k w_i^*}.$$

Αν τα βάρη $(w_i^*)^{-1}$ είναι οι πραγματικές διακυμάνσεις των $\hat{\theta}_i$, τότε ο ολικός εκτιμητής $\hat{\theta}^*$ είναι ΕΜΠ του θ [Whitehead, 2002].

Το 95% ασυμπτωτικό διάστημα εμπιστοσύνης για την ολική επίδραση της θεραπείας θ είναι:

$$\left(\hat{\theta}^* \pm z_{\alpha/2} \cdot \sqrt{\frac{1}{\sum_{i=1}^k w_i^*}} \right).$$

Επειδή στο μοντέλο τυχαίων επιδράσεων το βάρος της κάθε μελέτης είναι μικρότερο σε σχέση με το βάρος που δίνει το μοντέλο σταθερών επιδράσεων, η διακύμανση του εκτιμητή της ολικής επίδρασης της θεραπείας θα είναι μεγαλύτερη. Επομένως το διάστημα εμπιστοσύνης του μοντέλου τυχαίων επιδράσεων θα είναι πλατύτερο σε σχέση με το δ.ε του μοντέλου σταθερών επιδράσεων.

4.3.1 Εκτιμητής του τ^2 με την μέθοδο των ροπών

Αυτός ο εκτιμητής της πληθυσμιακής διασποράς τ^2 προκύπτει από το τεστ ετερογένειας, εξισώνοντας το στατιστικό Q με την αντίστοιχη αναμενόμενη τιμή του. Ο εκτιμητής του τ^2 με τη μέθοδο των ροπών είναι μη επαναληπτικός και προτάθηκε από τους *DerSimonian* and *Laird* το 1986.

Το στατιστικό του τεστ ετερογένειας (βλ. 5^ο κεφάλαιο, σελ. 56) δίνεται από την σχέση:

$$Q = \sum_i w_i \left(\hat{\theta}_i - \hat{\theta} \right)^2 = \sum_i w_i \left(\hat{\theta}_i - \theta \right)^2 - \left(\sum_i w_i \right) \cdot \left(\hat{\theta} - \theta \right)^2$$

όπου $\hat{\theta}_i$ είναι ο εκτιμητής της επίδρασης της θεραπείας στην i μελέτη, $\hat{\theta}$ είναι ο εκτιμητής σταθερών επιδράσεων του θ και w_i είναι το βάρος της i μελέτης.

Η αναμενόμενη τιμή του στατιστικού είναι:

$$\begin{aligned}
 E(Q) &= \sum_i w_i \left(\hat{\theta}_i - \theta \right)^2 - \left(\sum_i w_i \right) \cdot E \left(\hat{\theta} - \theta \right)^2 \\
 &= \sum_i w_i \text{Var} \left(\hat{\theta}_i \right) - \sum_i w_i \left\{ \frac{1}{\sum_i w_i} + \frac{\tau^2 \sum_i w_i^2}{\left(\sum_i w_i \right)^2} \right\} = \sum_i w_i \left(w_i^{-1} + \tau^2 \right) - \left\{ 1 + \frac{\tau^2 \sum_i w_i^2}{\sum_i w_i} \right\} \\
 &= \tau^2 \sum_i w_i + \kappa - 1 - \frac{\tau^2 \sum_i w_i^2}{\sum_i w_i} = (\kappa - 1) + \tau^2 \left(\sum_i w_i - \frac{\sum_i w_i^2}{\sum_i w_i} \right)
 \end{aligned}$$

Σύμφωνα με την μέθοδο των ροπών, ο εκτιμητής του τ^2 είναι ο:

$$\hat{\tau}^2 = \frac{Q - (\kappa - 1)}{\sum_i w_i - \frac{\sum_i w_i^2}{\sum_i w_i}}$$

Επειδή με αυτή τη μέθοδο μπορεί να καταλήξουμε σε αρνητικό εκτιμητή του τ^2 , στην πράξη χρησιμοποιούμε ως εκτιμητή τον: $\hat{\tau}^2 = \max \{0, \hat{\tau}^2\}$. Λόγω όμως του μετασχηματισμού που θεωρήσαμε ο εκτιμητής $\hat{\tau}^2$ που προκύπτει δεν είναι αμερόληπτος εκτιμητής του τ^2 .

Αν τιμή του $\hat{\tau}^2$ είναι μηδέν, τότε τα βάρη w_i^* θα είναι ίσα με τα αρχικά βάρη w_i . Σε αυτή την περίπτωση το τυπικό σφάλμα του εκτιμητή $\hat{\theta}$ και το αντίστοιχο διάστημα εμπιστοσύνης θα είναι τα ίδια με αυτά του μοντέλου σταθερών επιδράσεων. Επιπλέον η τιμή των εκτιμητών $\hat{\theta}$ και $\hat{\theta}^*$ που προκύπτουν από τα δυο μοντέλα θα είναι η ίδια. Αν η τιμή του $\hat{\tau}^2$ είναι μεγάλη, τότε το τυπικό σφάλμα και το διάστημα εμπιστοσύνης του μοντέλου τυχαίων επιδράσεων είναι πολύ μεγαλύτερα σε σχέση με το μοντέλο σταθερών επιδράσεων.

4.3.2 Εκτιμητής Μέγιστης Πιθανοφάνειας του τ^2 (ML Estimator)

Στο μοντέλο τυχαίων επιδράσεων έχουμε υποθέσει ότι: $\hat{\theta}_i \sim N(\theta, \sigma_i^2 + \tau^2)$. Όπως έχουμε αναφέρει, η διακύμανση σ_i^2 των εκτιμητών θ_i αντικαθίσταται από την εκτιμημένη διακύμανση $\hat{Var}(\hat{\theta}_i)$. Τα βάρη w_i είναι ίσα με $w_i = \hat{Var}(\hat{\theta}_i)^{-1}$ και θεωρούνται γνωστά για κάθε μελέτη.

Επειδή $\hat{\theta}_i \sim N(\theta, w_i^{-1} + \tau^2)$, η πιθανοφάνεια της i μελέτης είναι ίση με [Whitehead,2002]:

$$L(\theta, \tau^2; \hat{\theta}_i) = \frac{1}{\sqrt{2\pi(w_i^{-1} + \tau^2)}} \exp \left\{ -\frac{(\hat{\theta}_i - \theta)^2}{2(w_i^{-1} + \tau^2)} \right\}$$

Για k ανεξάρτητες μελέτες, ο λογάριθμος της σύνθετης πιθανοφάνειας δίνεται από την σχέση:

$$l(\theta, \tau^2; \hat{\theta}_i \quad i=1,2,\dots,k) = \mathbf{constant} - \frac{1}{2} \sum_i \log(w_i^{-1} + \tau^2) - \frac{1}{2} \sum_i \frac{(\hat{\theta}_i - \theta)^2}{(w_i^{-1} + \tau^2)}$$

Οι εκτιμητές μέγιστης πιθανοφάνειας των θ και τ^2 , βρίσκονται χρησιμοποιώντας μια επαναληπτική διαδικασία η οποία σε κάθε βήμα περιλαμβάνει δυο στάδια:

Αρχικά θεωρούμε ότι η παράμετρος τ^2 είναι σταθερά και υπολογίζουμε την τιμή της παραμέτρου θ η οποία μεγιστοποιεί τον λογάριθμο της πιθανοφάνειας. Στην συνέχεια υπολογίζεται η τιμή του θ και στη συνέχεια θεωρείται ως σταθερά για να υπολογιστεί η νέα τιμή του τ^2 που μεγιστοποιεί τον λογάριθμο της πιθανοφάνειας.

Ο εκτιμητής του θ στην $(t+1)$ επανάληψη του αλγορίθμου δίνεται από τη σχέση:

$$\hat{\theta}_{t+1}^* = \frac{\sum_i w_{it}^* \cdot \hat{\theta}_i}{\sum_i w_{it}^*}, \quad \text{για } t=0,1,\dots \text{ όπου } w_{it}^* = \left(w_i^{-1} + \tau_{M,t}^2 \right)^{-1} \quad \text{και } \tau_{M,t}^2 \text{ είναι ο εκτιμητής}$$

μέγιστης πιθανοφάνειας του τ^2 στην t επανάληψη του αλγορίθμου.

Ο εκτιμητής του τ^2 στην (t+1) επανάληψη του αλγορίθμου μπορεί να βρεθεί χρησιμοποιώντας την επαναληπτική διαδικασία των *Newton – Raphson*.

Ο προσεγγιστικός εκτιμητής του τ^2 είναι:

$$\hat{\tau}_{M,t+1}^2 = \frac{\sum_i (w_{it}^*)^2 \left\{ \left(\hat{\theta}_i - \theta_{t+1}^* \right)^2 - w_i^{-1} \right\}}{\sum_i (w_{it}^*)^2}.$$

Για να αρχίσει η επαναληπτική διαδικασία μπορούμε να χρησιμοποιήσουμε ως αρχική τιμή του τ^2 , την τιμή του εκτιμητή που βρήκαμε με την μέθοδο των ροπών.

Οι εκτιμητές μέγιστης πιθανοφάνειας έχουν το πλεονέκτημα ότι ασυμπτωτικά ακολουθούν κανονική κατανομή. Είναι γνωστό ότι ο εκτιμητής μέγιστης πιθανοφάνειας ενός δείγματος k ανεξάρτητων και ισόνομα κατανεμημένων τυχαίων μεταβλητών ακολουθεί κανονική κατανομή όταν το $k \rightarrow \infty$. Εδώ θα πρέπει να σημειώσουμε ότι στην μετανάλυση οι k τυχαίες μεταβλητές $\hat{\theta}_i$ αν και είναι ανεξάρτητες, δεν είναι ισόνομα κατανεμημένες αφού $Var(\hat{\theta}_i) = \hat{\sigma}_i^2 + \tau^2$. Παρόλα αυτά για μεγάλο πλήθος μελετών η ασυμπτωτικές υποθέσεις θα ισχύουν [Brockwell, 2001].

Ο εκτιμητής μέγιστης πιθανοφάνειας ML του τ^2 συνήθως υποεκτιμά την πραγματική τιμή, γιατί δεν λαμβάνει υπόψη την πληροφορία η οποία χρησιμοποιήθηκε στην εκτίμηση του θ .

4.3.3 Εκτιμητής Περιορισμένης Μέγιστης Πιθανοφάνειας του τ^2 (*REML Estimator*)

Οι *Patterson* και *Thompson* [1971] παρουσίασαν τον εκτιμητή περιορισμένης πιθανοφάνειας ως μέθοδο εκτίμησης των παραμέτρων μεταβλητότητας σε μη ισορροπημένους και μη πλήρεις σχεδιασμούς (*unbalanced and incomplete block designs*). Ο εκτιμητής περιορισμένης μέγιστης πιθανοφάνειας, είναι προτιμότερος από τον εκτιμητή μέγιστης πιθανοφάνειας γιατί λαμβάνει υπόψη την απώλεια βαθμών ελευθερίας επειδή από

το ίδιο σύνολο δεδομένων εκτιμώνται και το τ^2 και το θ . Επιπλέον παράγει αμερόληπτους εκτιμητές των παραμέτρων μεταβλητότητας.

Η συνάρτηση του λογαρίθμου της σύνθετης πιθανοφάνειας βασίζεται στα υπόλοιπα

$\left(\hat{\theta}_i - \hat{\theta}_{t+1}^*\right)$ και δίνεται από την σχέση, [Whitehead, 2002]:

$$lR\left(\tau^2; \left(\hat{\theta}_i - \hat{\theta}_{t+1}^*\right) i = 1, 2, \dots, k\right) = \text{constant} - \frac{1}{2} \sum_i \log(w_i^{-1} + \tau^2) - \frac{1}{2} \sum_i \frac{\left(\hat{\theta}_i - \hat{\theta}_{t+1}^*\right)^2}{(w_i^{-1} + \tau^2)} - \frac{1}{2} \log\left\{\sum_i \frac{1}{(w_i^{-1} + \tau^2)}\right\}$$

Οι REML εκτιμητές βρίσκονται με ανάλογη μεθοδολογία όπως οι ML εκτιμητές, όπου όμως :

$$w_{it}^* = \left(w_i^{-1} + \tau_{R,t}^2\right)^{-1}.$$

Στην (t+1) επανάληψη του αλγορίθμου, ο εκτιμητής του θ είναι:

$$\hat{\theta}_{t+1}^* = \frac{\sum_i w_{it}^* \cdot \hat{\theta}_i}{\sum_i w_{it}^*}, \text{ για } t = 0, 1, \dots$$

Ο REML εκτιμητής του τ^2 στην (t+1) επανάληψη του αλγορίθμου, μπορεί να υπολογιστεί με την επαναληπτική διαδικασία των *Newton – Raphson*.

Ο προσεγγιστικός εκτιμητής του τ^2 δίνεται από τη σχέση:

$$\hat{\tau}_{R,t+1}^2 = \frac{\sum_i (w_{it}^*)^2 \left\{ r \cdot \left(\hat{\theta}_i - \hat{\theta}_{t+1}^*\right)^2 / (r-1) - w_i^{-1} \right\}}{\sum_i (w_{it}^*)^2}.$$

Για περισσότερες πληροφορίες σχετικά με την χρήση των REML εκτιμητών στις κλινικές μελέτες, ο αναγνώστης παραπέμπεται στο άρθρο των *Brown και Kempton* [1994].

4.4 Σύγκριση των μοντέλων μετανάλυσης

Ένα σημαντικό κομμάτι της μετανάλυσης αποτελεί η επιλογή του μοντέλου μεταξύ σταθερών και τυχαίων επιδράσεων. Στο μοντέλο σταθερών επιδράσεων η επίδραση της θεραπείας θ_i θεωρείται ότι είναι η ίδια για όλες τις μελέτες και η μόνη πηγή αβεβαιότητας είναι αυτή που οφείλεται στην δειγματοληψία των ασθενών εντός της κάθε μελέτης (*within study variation*). Τα συμπεράσματα επομένως είναι ανάλογα με εκείνα που προκύπτουν με την εκτέλεση ανάλυσης διασποράς (ANOVA) στην περίπτωση που δεν υπάρχει *interstudy* μεταβλητότητα. Το μοντέλο τυχαίων επιδράσεων στηρίζεται στο γεγονός ότι επειδή υπάρχουν διαφορετικές προσεγγίσεις για να πραγματοποιηθεί μια μελέτη, μπορούν να προκύψουν διαφορετικά αποτελέσματα ως προς την επίδραση της θεραπείας. Επομένως στο μοντέλο τυχαίων επιδράσεων, η επίδραση της θεραπείας θ_i θεωρείται ότι είναι μια τυχαία μεταβλητή κανονικά κατανοημένη. Αυτή η κατάσταση αντιστοιχεί σε ένα μοντέλο ανάλυσης διακύμανσης (ANOVA) στο οποίο υπάρχει *inter-study* μεταβλητότητα.

Ο πληθυσμός στον οποίο γίνονται οι γενικεύσεις του μοντέλου σταθερών επιδράσεων περιορίζεται στο σύνολο των μελετών που έχουν τα ίδια χαρακτηριστικά με τις εξεταζόμενες μελέτες. Αντίθετα, τα αποτελέσματα του μοντέλου τυχαίων επιδράσεων γενικεύονται σε έναν υπερπληθυσμό παρόμοιων μελετών οι οποίες έχουν πραγματοποιηθεί ή πρόκειται να πραγματοποιηθούν. Συμπεραίνουμε λοιπόν ότι ενώ στο μοντέλο σταθερών επιδράσεων ο πληθυσμός στον οποίο γενικεύονται τα αποτελέσματα μπορεί να είναι πολύ μικρός για να είναι οποιασδήποτε πρακτικής σημασίας, στο μοντέλο τυχαίων επιδράσεων ο πληθυσμός αυτός μπορεί να είναι πολύ μεγάλος για να μπορέσουμε να τον προσδιορίσουμε.

Η κύρια διαφωνία μεταξύ των ερευνητών είναι αν θα πρέπει να ενσωματώσουμε την μεταβλητότητα που υπάρχει μεταξύ των μελετών για να εκτιμήσουμε την ολική επίδραση της θεραπείας. Αν υπάρχουν στοιχεία που μας αποδεικνύουν ότι η μεταβλητότητα αυτή είναι πολύ μικρή τότε η επιλογή μεταξύ των δυο μοντέλων δεν έχει διαφορά. Αν όμως υπάρχει σημαντική μεταβλητότητα (*ετερογένεια*), τότε μια ανάλυση που την αγνοεί, δηλαδή το μοντέλο σταθερών επιδράσεων, θα δίνει πιο μικρά διαστήματα εμπιστοσύνης σε σχέση με την προσέγγιση των τυχαίων επιδράσεων. Περισσότερες λεπτομέρειες για τον έλεγχο της ετερογένειας αναφέρονται στο 5^ο κεφάλαιο.

Αν και το μοντέλο τυχαίων επιδράσεων παράγει σαφώς πιο γενικεύσιμα συμπεράσματα, έχουν αναπτυχθεί κάποιοι προβληματισμοί σχετικά με την χρήση του μοντέλου. Πρώτον, το

μοντέλο αυτό υποθέτει ότι οι μελέτες που περιλαμβάνονται στην μετανάλυση είναι αντιπροσωπευτικές ενός υποθετικού πληθυσμού μελετών. Η υπόθεση όμως αυτή πόσο ρεαλιστική μπορεί να είναι; Επιπλέον παρατηρούμε ότι στο μοντέλο τυχαίων επιδράσεων η ετερογένεια των μελετών εκφράζεται μόνο με τον τυχαίο παράγοντα τ^2 .

Ο *Greenland* [1994] επισήμανε ότι το μοντέλο τυχαίων επιδράσεων είναι πιο ευαίσθητο στο σφάλμα δημοσίευσης (*publication bias*), αφού δίνει μεγαλύτερο βάρος στις μελέτες με μικρό μέγεθος δείγματος. Η μεροληψία αυτή προκύπτει από το γεγονός ότι αυτές οι μελέτες δεν δημοσιεύονται γιατί καταλήγουν συνήθως σε στατιστικά μη σημαντικά συμπεράσματα. Περισσότερα για το συστηματικό σφάλμα δημοσίευσης και την διόρθωσή του αναφέρονται στο 8^ο κεφάλαιο.

Άλλο ένα μειονέκτημα του μοντέλου τυχαίων επιδράσεων είναι το εξής: αν στην μετανάλυση συμπεριλαμβάνονται μόνο λίγες μελέτες δεν μπορούμε να πούμε με σιγουριά ότι ο εκτιμητής της μεταβλητότητας τ^2 μεταξύ των μελετών (*inter-study variability*) είναι αξιόπιστος. Στη περίπτωση αυτή το μοντέλο σταθερών επιδράσεων είναι πιο κατάλληλη επιλογή.

Αν οι μελέτες έχουν τα ίδια χαρακτηριστικά (*ομοιογενείς μελέτες*) τότε τα δυο μοντέλα καταλήγουν στον ίδιο ολικό εκτιμητή του θ αφού είναι $\hat{\tau}^2 = 0$. Επιπλέον τα τυπικά σφάλματα των ολικών εκτιμητών των δυο μοντέλων θα είναι τα ίδια. Όσο πιο ετερογενείς είναι οι μελέτες που περιλαμβάνονται στην μετανάλυση, τόσο το τυπικό σφάλμα του ολικού εκτιμητή που δίνει το μοντέλο τυχαίων επιδράσεων αυξάνει σε σχέση με το αντίστοιχο σφάλμα του μοντέλου σταθερών επιδράσεων. Αξίζει να παρατηρήσουμε ότι ακόμα και όταν οι μελέτες είναι ετερογενείς ο εκτιμητής της ολικής επίδρασης της θεραπείας που δίνει το μοντέλο σταθερών επιδράσεων αποτελεί και πάλι τον μέσο όρο των επιδράσεων της θεραπείας σε κάθε μελέτη.

Στο άρθρο των *Brockwell* και *Gordon* [2001], έγινε σύγκριση των δύο μοντέλων με χρήση της πιθανότητας επικάλυψης (*coverage probability*) και μεθόδων προσομοίωσης. Αν για παράδειγμα το (A,B) είναι ένα τυχαίο διάστημα εμπιστοσύνης της μέσης τιμής μ τότε ως πιθανότητα επικάλυψης ορίζεται η $\Pr(\mu \in (A,B))$, η οποία για $\alpha = 5\%$ θα πρέπει να είναι ίση με 0.95. Κατέληξαν στο συμπέρασμα ότι το μοντέλο σταθερών επιδράσεων συμπεριφέρεται καλά μόνο όταν υπάρχει πολύ μικρή μεταβλητότητα μεταξύ των μελετών. Επιπλέον το τεστ για τον έλεγχο της ετερογένειας, για το οποίο θα μιλήσουμε πιο

συγκεκριμένα στο επόμενο κεφάλαιο, καταλήγει κατά μέσο όρο σε διαστήματα εμπιστοσύνης αρκετά στενά, γι' αυτό προτείνεται η χρήση του μοντέλου τυχαίων επιδράσεων ανεξάρτητα από την τιμή του στατιστικού τεστ. Αν ο εκτιμητής της μεταβλητότητας $\hat{\tau}^2$ είναι μηδέν τότε το μοντέλο τυχαίων επιδράσεων θα ταυτιστεί με το μοντέλο σταθερών επιδράσεων, ενώ αν το πλήθος των μελετών είναι σχετικά μικρό, δηλαδή $k < 20$ τότε το μοντέλο τυχαίων επιδράσεων δεν δίνει ακριβή αποτελέσματα.

Μερικές φορές είναι χρήσιμο να εξεταστούν τα αποτελέσματα και από τα δυο μοντέλα. Εάν πραγματοποιηθεί όμως μια τέτοια ανάλυση θα πρέπει να έχει αποφασιστεί εκ των προτέρων πως θα χειριστούμε τα αποτελέσματά της, προκειμένου να αποφύγουμε πιθανή μεροληψία η οποία σχετίζεται με την επιλογή του μοντέλου. Γενικά το μοντέλο τυχαίων επιδράσεων δίνει πιο συντηρητικούς εκτιμητές με πιο πλατιά διαστήματα εμπιστοσύνης. Εάν τα μοντέλα καταλήγουν σε σημαντικές διαφορές τότε θα πρέπει να είμαστε προσεκτικοί ως προς την ποιότητα των μελετών που περιλαμβάνονται, τις διαφορές στα πρωτόκολλα ή τις διαφορές που πιθανόν υπάρχουν στους υπό εξέταση πληθυσμούς.

ΚΕΦΑΛΑΙΟ 5

Ετερογένεια

5.1 Τι είναι η ετερογένεια;

Τις περισσότερες φορές οι μεταanalύσεις βασίζονται σε ένα σύνολο μελετών, οι οποίες έχουν πραγματοποιηθεί στο παρελθόν και ακολουθούν διαφορετικό πρωτόκολλο η κάθε μια. Πριν συνδυάσουμε επομένως τα δεδομένα θα πρέπει να αναρωτηθούμε: τα αποτελέσματα των μελετών που εξετάζουμε είναι παρόμοια; Οποιοδήποτε είδος μεταβλητότητας μεταξύ των μελετών μπορεί να θεωρηθεί ότι αποτελεί ετερογένεια. Επομένως είναι αναγκαίο να κάνουμε ένα διαχωρισμό μεταξύ των διαφορετικών τύπων ετερογένειας. Η μεταβλητότητα που οφείλεται στα διαφορετικά χαρακτηριστικά των ασθενών, στο είδος της θεραπευτικής μεθόδου και στην αποκριτική μεταβλητή που μετρήθηκε σε κάθε μελέτη ονομάζεται κλινική ποικιλομορφία ή κλινική ετερογένεια (*clinical diversity* ή *clinical heterogeneity*). Η μεταβλητότητα όμως ως προς τον σχεδιασμό και την ποιότητα της μελέτης περιγράφεται ως μεθοδολογική ποικιλομορφία ή μεθοδολογική ετερογένεια (*methodological diversity* ή *methodological heterogeneity*). Η μεταβλητότητα ως προς την επίδραση της θεραπείας μεταξύ των μελετών που εξετάζουμε είναι γνωστή ως στατιστική ετερογένεια (*statistical heterogeneity*) και είναι αποτέλεσμα και της κλινικής και της μεθοδολογικής ποικιλομορφίας των μελετών [Thompson, 1994]. Ο Sutton *et al* [1998] όρισε την ετερογένεια ως “την μεταβλητότητα στο μέγεθος των εκτιμητών της επίδρασης της θεραπείας, η οποία δεν οφείλεται στο σφάλμα δειγματοληψίας”. Στη συνέχεια, κατά σύμβαση, θα αναφερόμαστε στην στατιστική ετερογένεια χρησιμοποιώντας μόνο τον όρο ετερογένεια.

Στο σημείο αυτό θα πρέπει να αναφερθούμε στο διαχωρισμό της ετερογένειας σε ποιοτική ετερογένεια (*qualitative interaction*) και σε ποσοτική ετερογένεια (*quantitative interaction*). Η ποιοτική ετερογένεια υπάρχει όταν η κατεύθυνση της επίδρασης της θεραπείας αντιστρέφεται, δηλαδή η θεραπεία σε κάποιες μελέτες είναι ευεργετική-θετική και σε κάποιες άλλες είναι βλαβερή-αρνητική. Η ποσοτική ετερογένεια εμφανίζεται όταν ποικίλλει το μέγεθος της επίδρασης της θεραπείας αλλά όχι η κατεύθυνσή της, δηλαδή η θεραπεία είναι ευεργετική σε όλες τις μελέτες αλλά σε διαφορετικό βαθμό.

Η ύπαρξη ετερογένειας είναι ένα από τα σημαντικότερα θέματα το οποίο επηρεάζει την ανάλυση των μελετών και επιπλέον καθορίζει σε μεγάλο βαθμό την μεθοδολογία που θα πρέπει να ακολουθήσουμε για την επεξεργασία των δεδομένων. Το πρώτο βήμα είναι να ανιχνεύσουμε είτε γραφικά είτε με στατιστικό τεστ αν η ετερογένεια είναι σημαντική και στη συνέχεια να αποφασίσουμε με ποιο τρόπο θα την χειριστούμε ή ακόμα και να ελέγξουμε αν μπορούμε να εξηγήσουμε ποιοι είναι οι παράγοντες που την προκαλούν.

Στην παράγραφο 5.2 παρουσιάζονται οι γραφικές μέθοδοι ανίχνευσης της ετερογένειας ενώ στην παράγραφο 5.3 δίνεται ο αντίστοιχος στατιστικός έλεγχος. Στη συνέχεια στη παράγραφο 5.4 προτείνονται τρόποι για το πώς μπορούμε να χειριστούμε την ετερογένεια σε σχέση με το μοντέλο ανάλυσης των δεδομένων, ενώ στις παραγράφους 5.4.1 και 5.4.2 παρουσιάζονται αντίστοιχα η μέθοδος της μετα-παλινδρόμησης (*meta-regression*) και η μέθοδος ανάλυσης υποομάδων (*subgroup analysis*), με τις οποίες μπορούμε να ανακαλύψουμε πιθανούς λόγους που προκαλούν την ετερογένεια.

5.2 Γραφική διερεύνηση της ετερογένειας

5.2.1 Γράφημα των z-scores

Για την κατασκευή αυτού του γραφήματος υπολογίζουμε για κάθε μια μελέτη το αντίστοιχο *z-score* σύμφωνα με τη σχέση:

$$z_i = \frac{\hat{\theta}_i - \hat{\theta}}{\sqrt{\text{Var}(\hat{\theta}_i)}}, \quad \text{με} \quad \hat{\theta} = \frac{\sum_{i=1}^k w_i \hat{\theta}_i}{\sum_{i=1}^k w_i} \quad \text{και} \quad w_i = \frac{1}{\text{Var}(\hat{\theta}_i)}, \quad i = 1, 2, \dots, k$$

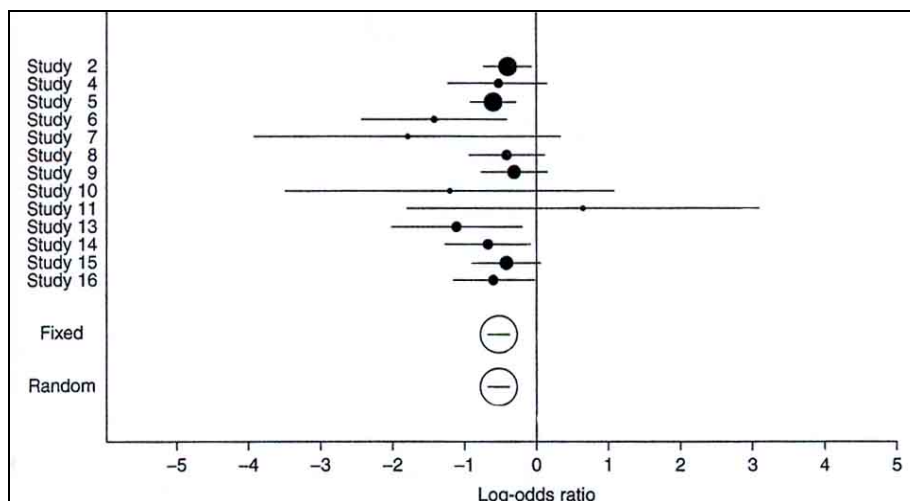
$\hat{\theta}_i$ είναι η επίδραση της θεραπείας στην i μελέτη, $\hat{\theta}$ είναι ο ολικός εκτιμητής της επίδρασης της θεραπείας (με βάση το μοντέλο σταθερών επιδράσεων) και w_i είναι το βάρος της i μελέτης. Στην συνέχεια κατασκευάζουμε ένα ιστόγραμμα με τα *z-scores* και το συγκρίνουμε με τη τυπική κανονική κατανομή, αφού θα πρέπει να ισχύει $\hat{\theta}_i \sim N(\hat{\theta}, w_i^{-1})$ δηλαδή

$\frac{\hat{\theta}_i - \hat{\theta}}{\sqrt{1/w_i}} \sim N(0,1)$ (βλέπε παρ. 4.2.3, σελ. 34). Αν το ιστόγραμμα που κατασκευάσαμε δεν

ταυτίζεται με την τυπική κανονική κατανομή τότε έχουμε μεγαλύτερη μεταβλητότητα απ' ό τι θα περιμέναμε και πιθανόν αυτή η μεταβλητότητα να οφείλεται στην ετερογένεια των μελετών. Η γραφική αυτή μέθοδος μπορεί να χρησιμοποιηθεί μόνο για μεγάλο πλήθος μελετών.

5.2.2 Forest Plot

Στο γράφημα *forest plot* απεικονίζονται τα διαστήματα εμπιστοσύνης της επίδρασης της θεραπείας σε κάθε μελέτη με μια οριζόντια γραμμή. Για παράδειγμα αν το επίπεδο σημαντικότητας είναι $\alpha=5\%$ στο γράφημα απεικονίζονται τα διαστήματα $\hat{\theta}_i \pm 1.96\sqrt{Var(\hat{\theta}_i)}$ για $i = 1, 2, \dots, k$. Σε κάθε διάστημα εμπιστοσύνης, το σημείο στο οποίο αντιστοιχεί η τιμή $\hat{\theta}_i$ συμβολίζεται με ένα κύκλο, η διάμετρος του οποίου είναι ανάλογη με την ακρίβεια του εκτιμητή δηλαδή είναι ανάλογη με το βάρος $w_i = 1/Var(\hat{\theta}_i)$ (Σχήμα 5.1). Όσο πιο ακριβής είναι ένας εκτιμητής τόσο πιο μεγάλος θα είναι ο κύκλος στο κέντρο του διαστήματος της αντίστοιχης μελέτης. Επομένως αν οι κύκλοι στα αντίστοιχα διαστήματα είναι μεγάλοι και επιπλέον υπάρχει μεταβλητότητα στην επικάλυψη των διαστημάτων εμπιστοσύνης, τότε προφανώς οι εκτιμητές διαφέρουν μεταξύ τους και επομένως υπάρχει ετερογένεια μεταξύ των μελετών.



Σχήμα 5.1: Forest plot of the log-odds ratio of a stroke on antihypertensive treatment relative to control

(Πηγή: A. Whitehead [2002])

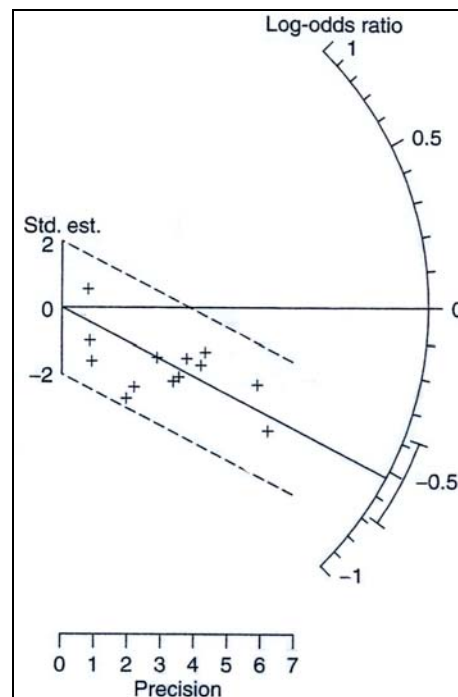
5.2.3 Radial Plot

Το *radial plot* περιγράφηκε από τον *Galbraith* [1988]. Είναι ένα διμεταβλητό διάγραμμα διασποράς (x, y) στο οποίο απεικονίζεται ο τυποποιημένος εκτιμητής της διαφοράς των θεραπειών σε σχέση με την ακρίβεια του εκτιμητή για κάθε μια μελέτη.

Ο τυποποιημένος εκτιμητής για την μελέτη i δίνεται από τον τύπο $\hat{\theta}_i / s.e.(\hat{\theta}_i)$ ενώ η αντίστοιχη ακρίβεια σύμφωνα με τον *Galbraith* είναι ίση με $\sqrt{w_i} = 1 / s.e.(\hat{\theta}_i)$, δηλαδή ισούται με τη ρίζα της συνηθισμένης έκφρασης της ακρίβειας. Στην συνέχεια στο διάγραμμα

κατασκευάζεται η ευθεία $y = \hat{\theta} x$, όπου $\hat{\theta} = \frac{\sum_{i=1}^k w_i \cdot \hat{\theta}_i}{\sum_{i=1}^k w_i}$ καθώς και το διάστημα $y = \hat{\theta} x \pm 2$

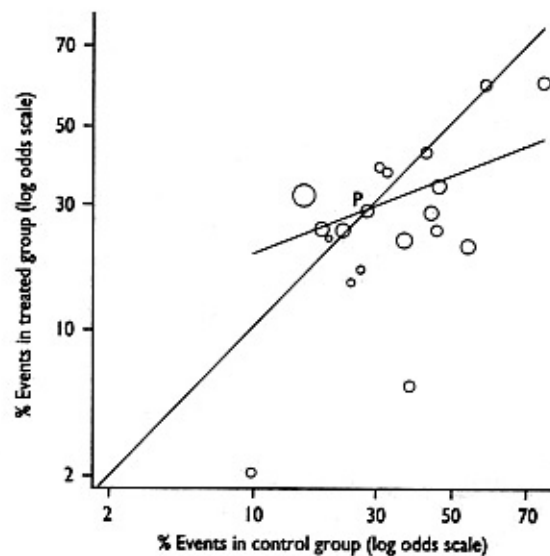
(Σχήμα 5.2). Αν υπάρχουν μελέτες οι οποίες βρίσκονται έξω από το προηγούμενο διάστημα, τότε μπορούμε να αποφανθούμε ότι οι μελέτες που εξετάζουμε είναι ετερογενείς.



Σχήμα 5.2: *Radial plot of the study” standardized estimates” of the log-odds ratio of a stroke on antihypertensive treatment relative to control, against “precision” (Πηγή: A. Whitehead [2002])*

5.2.4 L' Abbe Plot

Ο *L'Abbe* [1987] πρότεινε το ομώνυμο διάγραμμα προκειμένου να εντοπιστεί η μεταβλητότητα μεταξύ των παρατηρούμενων εκτιμητών της διαφοράς των θεραπειών. Το διάγραμμα *L'Abbe* είναι χρήσιμο για δίτιμα κυρίως δεδομένα. Είναι ένα είδος διαγράμματος διασποράς με σημεία τα (x, y) όπου x είναι το ποσοστό των ασθενών που εμφάνισαν το γεγονός στην ομάδα θεραπείας της i μελέτης και y είναι το αντίστοιχο ποσοστό των ασθενών για την ομάδα *control*. Αν οι μελέτες είναι ομοιογενείς τότε τα σημεία θα πρέπει να βρίσκονται πάνω σε μια νοητή ευθεία γραμμή, με κλίση ίση με τον ολικό εκτιμητή της θεραπείας $\hat{\theta}$, παράλληλη της ευθείας $y = x$. Σε διαφορετική περίπτωση υπάρχει ένδειξη για ύπαρξη ετερογένειας. Τα σημεία του διαγράμματος είναι κύκλοι των οποίων η διάμετρος είναι ανάλογη με την ακρίβεια των αντίστοιχων εκτιμητών.



Σχήμα 5.3: *L' Abbe plot for sclerotherapy trials. (Πηγή: Sharp et al. [1996])*

5.3 Στατιστικός έλεγχος ετερογένειας

Έστω $\theta_1, \theta_2, \dots, \theta_k$ η επίδραση της θεραπείας σε σχέση με το *control* στις k μελέτες αντίστοιχα. Το στατιστικό τεστ ετερογένειας ελέγχει την υπόθεση ότι η επίδραση της θεραπείας είναι η ίδια σε όλες τις μελέτες, δηλαδή ελέγχει την μηδενική υπόθεση $H_0: \theta_1 = \theta_2 = \dots = \theta_k$ έναντι της εναλλακτικής υπόθεσης H_a : τουλάχιστον ένα θ_i είναι διαφορετικό. Ουσιαστικά, αυτός ο έλεγχος εξετάζει εάν όλες οι μελέτες υπολογίζουν μια ενιαία επίδραση της θεραπείας, έστω θ , και εάν η διαφορά στις εκτιμήσεις της θεραπείας από την κάθε μελέτη είναι τυχαία ή συστηματική.

Το τεστ ετερογένειας Q βασίζεται στο στατιστικό του *Cochran* [Cochran, 1954 και Fleiss, 1981] και δίνεται από την σχέση:

$$Q = \sum_{i=1}^k w_i \left(\hat{\theta}_i - \hat{\theta} \right)^2 = \sum_{i=1}^k w_i \hat{\theta}_i^2 - \frac{\left(\sum_{i=1}^k w_i \hat{\theta}_i \right)^2}{\sum_{i=1}^k w_i},$$

όπου $\hat{\theta}_i$ είναι ο εκτιμητής της επίδρασης θ_i της εξεταζόμενης θεραπείας στην i μελέτη και w_i είναι ο αντίστροφος της διακύμανσης του εκτιμητή $\hat{\theta}_i$, δηλαδή $w_i = 1/Var(\hat{\theta}_i)$. Μόνο στην περίπτωση που εφαρμόζεται η μέθοδος του *Peto* τα βάρη είναι ίσα με $w_i = v_i$.

Κάτω από την μηδενική υπόθεση, το στατιστικό Q ακολουθεί ασυμπτωτικά χ^2 κατανομή με $k-1$ βαθμούς ελευθερίας.

Η μονομερής χρήση του στατιστικού τεστ για τον έλεγχο της ετερογένειας των μελετών δεν συνίσταται. Οι *Hardy* και *Thompson* [1998] εξέτασαν την ισχύ του τεστ ετερογένειας κάτω από διαφορετικά σενάρια. Κατέληξαν στο συμπέρασμα ότι η ισχύς του είναι χαμηλή, ειδικά στην περίπτωση που τα δεδομένα είναι ελλιπή (*sparse data*) ή όταν κάποια μελέτη έχει πολύ μεγαλύτερο βάρος από τις υπόλοιπες. Το ίδιο ισχύει αν συνδυάζονται λίγες μελέτες ή μελέτες με μικρό μέγεθος δείγματος η κάθε μια. Ενώ ένα στατιστικά σημαντικό αποτέλεσμα μας προειδοποιεί για την ύπαρξη ετερογένειας, ένα μη στατιστικά σημαντικό αποτέλεσμα δεν θα πρέπει να θεωρηθεί ως απόδειξη για την μη ύπαρξη ετερογένειας. Άλλο ένα πρόβλημα του ελέγχου ετερογένειας προκύπτει όταν συνδυάζονται πολλές μελέτες. Στην περίπτωση αυτή το

τεστ έχει μεγάλη ισχύ και επομένως μπορεί να εντοπίσει ένα μικρό ποσό ετερογένειας ακόμα κι αν αυτό είναι κλινικά μη σημαντικό. Γι' αυτούς τους λόγους χρησιμοποιείται ως επίπεδο σημαντικότητας το $\alpha=10\%$ και όχι το συνηθισμένο $\alpha=5\%$ [Fleiss,1981].

Όσο πιο ουσιαστική είναι η ετερογένεια, τόσο περισσότερο επηρεάζονται τα αποτελέσματα της μετανάλυσης. Επομένως είναι αναγκαίο να μπορέσουμε να μετρήσουμε την έκταση της ετερογένειας σε σχέση με το σύνολο των μελετών. Οι Higgins και Thompson[2002] πρότειναν δυο σχετικούς δείκτες:

$$H^2 = \frac{Q}{k-1} \quad \text{και} \quad I^2 = \frac{H^2 - 1}{H^2} \times 100\%$$

όπου Q είναι το στατιστικό του ελέγχου ετερογένειας και k είναι το πλήθος των μελετών.

Ο δείκτης H^2 εκφράζει το πόσες φορές είναι μεγαλύτερο το στατιστικό ετερογένειας Q από τους αντίστοιχους βαθμούς ελευθερίας του. Επειδή όταν δεν υπάρχει ετερογένεια ισχύει $E[Q] = k - 1$ (βλέπε παρ. 4.3.1), αν είναι $H = 1$ τότε αυτό σημαίνει ότι υπάρχει ομοιογένεια μεταξύ των επιδράσεων της θεραπείας στις k μελέτες. Τιμές μεγαλύτερες του 1.5 θα πρέπει να μας επιστήσουν την προσοχή για σημαντική ετερογένεια, ενώ τιμές μικρότερες του 1.2 δεν θα πρέπει να μας ανησυχούν.

Ο δείκτης I^2 περιγράφει το ποσοστό της μεταβλητότητας μεταξύ των μελετών, το οποίο οφείλεται στην ετερογένεια και όχι στο σφάλμα δειγματοληψίας. Ένα ποσοστό μεγαλύτερο του 50% μπορεί να θεωρηθεί ουσιαστική ετερογένεια, ενώ ποσοστά μικρότερα του 30% δηλώνουν ήπια ετερογένεια.

5.4 Πώς θα χειριστούμε την ετερογένεια;

Μερικές φορές η ετερογένεια μπορεί να οφείλεται μόνο στην τύχη ή να είναι προϊόν αυταπάτης λόγω της επιλογής του χρησιμοποιούμενου μέτρου επίδρασης της θεραπείας. Έτσι αν για παράδειγμα η μεταβλητή απόκρισης ενός συνόλου μελετών είναι δίτιμη, μπορεί χρησιμοποιώντας ως μέτρο την διαφορά των ποσοστών να εντοπίζουμε ετερογένεια μεταξύ των μελετών, ενώ για τα ίδια δεδομένα, αν χρησιμοποιήσουμε ως μέτρο τον λογάριθμο του λόγου σχετικών πιθανοτήτων να μην εντοπίζεται ετερογένεια. Στην συγκεκριμένη περίπτωση αν θεωρήσουμε ως μέτρο σύγκρισης των θεραπειών την διαφορά των ποσοστών, υπάρχει

πιθανότητα να εμφανιστεί ετερογένεια αν τα ποσοστά των ομάδων *control* κυμαίνονται σε ευρύ φάσμα, ή είναι κοντά στο μηδέν ή στη μονάδα [Engels, 2000].

Αν οι μελέτες είναι σημαντικά ετερογενείς και ιδιαίτερα όταν υπάρχει ποιοτική ετερογένεια, τότε είναι παραπλανητικό να εκτιμήσουμε την ολική επίδραση της θεραπείας και επομένως δεν είναι σωστό να συνδυάσουμε αυτές τις μελέτες.

Στην περίπτωση που η ετερογένεια είναι ασήμαντη, μπορούμε να την αγνοήσουμε και να χρησιμοποιήσουμε το μοντέλο σταθερών επιδράσεων. Στο μοντέλο αυτό υποθέτουμε ότι οι παρατηρηθείσες διαφορές μεταξύ των μελετών οφείλονται μόνο στο σφάλμα δειγματοληψίας. Η προέκταση των αποτελεσμάτων του μοντέλου σταθερών επιδράσεων στο γενικό πληθυσμό, κάνει την υπόθεση ότι τα χαρακτηριστικά των ασθενών που συμμετέχουν στις μελέτες είναι τα ίδια με τα χαρακτηριστικά των ασθενών του γενικού πληθυσμού.

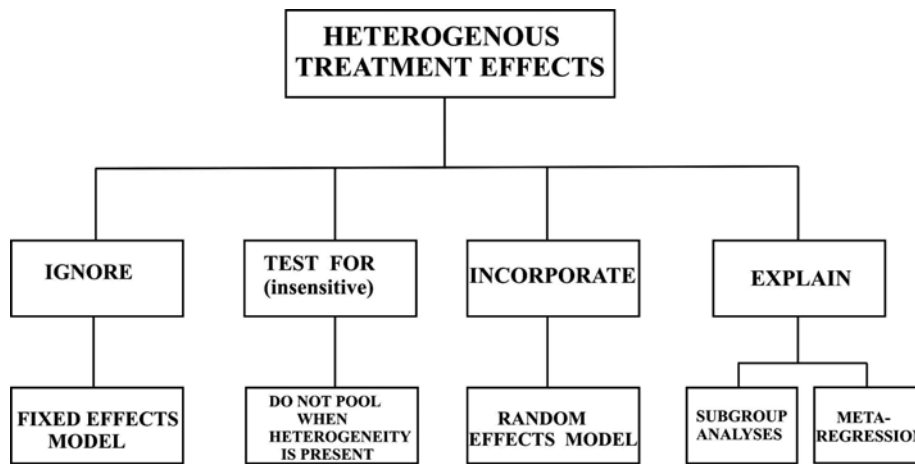
Όταν όμως η ετερογένεια μεταξύ των μελετών είναι στατιστικά σημαντική, τότε δεν θα μπορούμε να χρησιμοποιήσουμε το μοντέλο σταθερών επιδράσεων. Αν διαπιστώσουμε ότι η ετερογένεια οφείλεται σε τυχαίες διαφορές μεταξύ των μελετών και οι πηγές της δεν μπορούν να προσδιοριστούν, τότε χρησιμοποιούμε το μοντέλο τυχαίων επιδράσεων για να την ενσωματώσουμε στην ανάλυσή μας. Όπως έχουμε αναφέρει στην παράγραφο 4.3 στο μοντέλο αυτό οι επιδράσεις της θεραπείας στις k μελέτες θεωρούνται τυχαίες. Το μοντέλο τυχαίων επιδράσεων παράγει πιο γενικεύσιμα συμπεράσματα και στηρίζεται στην υπόθεση ότι οι μελέτες που λαμβάνουν μέρος στην μετανάλυση είναι αντιπροσωπευτικές όλου του πληθυσμού των μελετών που έχουν πραγματοποιηθεί ή πρόκειται να πραγματοποιηθούν στο μέλλον.

Αν η ετερογένεια είναι συστηματική, οφείλεται δηλαδή σε αναγνωρίσιμες διαφορές μεταξύ των μελετών, τότε έχει ενδιαφέρον να εξετάσουμε ποιοι από αυτούς τους παράγοντες προκαλούν την ετερογένεια. Η ετερογένεια επομένως μπορεί να εξηγηθεί με χρήση των μεθόδων της μετα-παλινδρόμησης (*meta-regression*) ή με ανάλυση υποομάδων (*subgroup analysis*). Οι μεθοδολογίες αυτές θα παρουσιαστούν στις παραγράφους 5.4.1 και 5.4.2 αντίστοιχα με περισσότερες λεπτομέρειες.

Τέλος, η ετερογένεια μπορεί να οφείλεται στην παρουσία μιας ή δύο απομακρυσμένων μελετών (*outlying trials*) των οποίων τα αποτελέσματα συγκρούονται με το υπόλοιπο πλήθος των μελετών. Γενικά είναι παράλογο να αποκλειστούν μελέτες από μια μετανάλυση βάσει των αποτελεσμάτων τους, γιατί αυτό μπορεί να εισαγάγει μεροληψία. Εντούτοις, αν υπάρχει ένας προφανής λόγος για τα τόσο διαφορετικά αποτελέσματα αυτών των μελετών τότε η

μελέτη μπορεί να αφαιρεθεί με περισσότερη εμπιστοσύνη. Δεδομένου όμως ότι τουλάχιστον ένα χαρακτηριστικό μιας μελέτης μπορεί να βρεθεί διαφορετικό από τις υπόλοιπες, το κριτήριο αυτό είναι αναξιόπιστο αφού μπορεί να προκύψει για οποιαδήποτε μελέτη. Αυτό που ενδείκνυται σε αυτή την περίπτωση είναι να εκτελεστούν αναλύσεις ευαισθησίας (*sensitivity analysis*) με ή χωρίς αυτές τις απομακρυσμένες μελέτες. Το θέμα την ανάλυσης ευαισθησίας παρουσιάζεται με περισσότερες λεπτομέρειες στην παράγραφο 8.2.

Συνοψίζοντας παραθέτουμε το επόμενο σχήμα το οποίο παρουσιάζει το μεθοδολογικό πλαίσιο χειρισμού της ετερογένειας :



Σχήμα 5.3 Χειρισμός της ετερογένειας (Πηγή: Joseph Lau [1997])

5.4.1 Μετα-παλινδρόμηση (*Meta-regression*)

Η εξάρτηση της επίδρασης της θεραπείας θ_i από ένα ή περισσότερα χαρακτηριστικά μιας μελέτης μπορεί να μελετηθεί με τα μοντέλα της μετα-παλινδρόμησης (*meta-regression*). Ουσιαστικά η μέθοδος αυτή είναι παλινδρόμηση με ανεξάρτητες μεταβλητές κάποιους παράγοντες (*trial-level covariates*) των οποίων οι τιμές δίνονται σε κάθε μελέτη. Με την βοήθεια της μετα-παλινδρόμησης μπορούμε να εξετάσουμε αν κάποιοι από αυτούς τους παράγοντες εξηγούν ένα ποσό της ετερογένειας η οποία εντοπίζεται μεταξύ των μελετών. Αν οι μελέτες είναι λίγες, ακόμα κι αν το πλήθος των ασθενών είναι μεγάλο, η μετα-παλινδρόμηση δεν μπορεί να παράγει επιστημονικά χρήσιμα αποτελέσματα [Thompson, 2002]. Η παλινδρόμηση θα πρέπει να είναι σταθμισμένη έτσι ώστε οι πιο ακριβείς μελέτες να έχουν μεγαλύτερη επίδραση στην ανάλυση. Η μετα-παλινδρόμηση μπορεί να

πραγματοποιηθεί ακόμα κι αν το τεστ ετερογένειας καταλήξει σε στατιστικά μη σημαντικό αποτέλεσμα. Επειδή όπως αναφέρθηκε η ισχύς του τεστ είναι χαμηλή, ένα στατιστικά μη σημαντικό αποτέλεσμα δεν αποδεικνύει με σιγουριά ότι δεν υπάρχει ετερογένεια μεταξύ των μελετών.

Ένα πρακτικό πρόβλημα που περιορίζει τη χρήση της μετα-παλινδρόμησης είναι η μη διαθεσιμότητα των αναγκαίων δεδομένων από τις μελέτες. Για να μπορέσουμε να εκτελέσουμε την μετα-παλινδρόμηση θα πρέπει σε κάθε μελέτη να γνωρίζουμε την εκτιμημένη επίδραση της θεραπείας, τη διακύμανσή της και τις τιμές των υπό εξέταση συμμεταβλητών. Ο πιο συνηθισμένος λόγος για τον οποίο δεν μπορεί να πραγματοποιηθεί η μετα-παλινδρόμηση είναι ότι δεν είναι διαθέσιμα όλα τα παραπάνω στοιχεία. Για περισσότερες λεπτομέρειες σχετικά με την εφαρμογή της μετα-παλινδρόμησης ο αναγνώστης παραπέμπεται στο άρθρο των *Thompson* και *Higgins* [2002].

Στη συνέχεια θα παρουσιάσουμε την μέθοδο της μετα-παλινδρόμησης στο πλαίσιο του μοντέλου σταθερών επιδράσεων (παρ. 5.4.1.1) και του μοντέλου τυχαίων επιδράσεων (παρ. 5.4.1.2). Αν και θα θεωρήσουμε ότι υπάρχει ένας επεξηγηματικός παράγοντας της ετερογένειας, εύκολα μπορεί να γίνει επέκταση της μεθοδολογίας όταν υπάρχουν περισσότεροι παράγοντες.

Στο άρθρο των *Thompson* και *Sharp* [1999] δίνονται περισσότερες πληροφορίες σχετικά με τα μοντέλα της μετα-παλινδρόμησης.

5.4.1.1 Μέθοδος Σταθερών Επιδράσεων

Στο μοντέλο σταθερών επιδράσεων υποθέσαμε ότι η επίδραση της θεραπείας θ_i είναι η ίδια σε όλες τις μελέτες. Έστω ακόμα ότι x_i είναι η τιμή του επεξηγηματικού παράγοντα στην i μελέτη. Τότε σε αντιστοιχία με το μοντέλο σταθερών επιδράσεων θα είναι:

$$\hat{\theta}_i = \beta_1 + \beta_2 x_i + e_i, \quad i = 1, 2, \dots, k$$

Τα σφάλματα e_i ακολουθούν κανονική κατανομή με μέση τιμή 0 και διακύμανση σ_i^2 , δηλαδή $e_i \sim N(0, \sigma_i^2)$.

Το προηγούμενο μοντέλο μπορεί να γραφεί ισοδύναμα στην μορφή:

$$\hat{\theta}_i \sim N(\beta_1 + \beta_2 x_i, \sigma_i^2)$$

όπου το x_i συμβολίζει τον επεξηγηματικό παράγοντα, το β_2 εκφράζει την αλλαγή του $\hat{\theta}_i$ για κάθε μοναδιαία αλλαγή του x_i και το β_1 είναι η ολική επίδραση της θεραπείας όταν $x_i = 0$.

Οι εκτιμητές μέγιστης πιθανοφάνειας των συντελεστών του μοντέλου μπορούν να υπολογιστούν με την μέθοδο σταθμισμένων ελαχίστων τετραγώνων, όπου τα βάρη w_i είναι ίσα με τον αντίστροφο της διακύμανσης των $\hat{\theta}_i$, δηλαδή $w_i = \frac{1}{\text{Var}(\hat{\theta}_i)}$. Συνήθως στα

στατιστικά πακέτα γίνεται η υπόθεση $\sigma_i^2 = \sigma^2 / w_i$ αντί της υπόθεσης $\sigma_i^2 = 1 / w_i$, όπου η τιμή του σ^2 υπολογίζεται από τα δεδομένα. Επομένως για να υπολογιστούν σωστά τα τυπικά σφάλματα των συντελεστών της μετα-παλινδρόμησης θα πρέπει να διαιρέσουμε τις τιμές που δίνει το υπολογιστικό πακέτο με την ρίζα του μέσου τετραγωνικού σφάλματος (MSE) [Thompson και Sharp, 1999].

5.4.1.2 Μέθοδος Τυχαίων Επιδράσεων

Στο μοντέλο των σταθερών επιδράσεων δεν επιτράπηκε η μεταβλητότητα στην επίδραση της θεραπείας μεταξύ των μελετών. Για να ενσωματώσουμε την ετερογένεια των μελετών θα πρέπει να συμπεριλάβουμε στο προηγούμενο μοντέλο έναν τυχαίο παράγοντα ως εξής:

$$\hat{\theta}_i = \beta_1 + \beta_2 x_i + v_i + e_i, \quad i = 1, 2, \dots, k$$

Τα e_i και v_i είναι ανεξάρτητες τυχαίες μεταβλητές με αντίστοιχες κατανομές:

$$e_i \sim N(0, \sigma_i^2) \quad \text{και} \quad v_i \sim N(0, \tau^2).$$

Το τ^2 εκφράζει την μεταξύ των μελετών μεταβλητότητα (*between study variability*).

Το προηγούμενο μοντέλο μπορεί να γραφεί ισοδύναμα στην μορφή:

$$\hat{\theta}_i \sim N\left(\beta_1 + \beta_2 x_i, \sigma_i^2 + \tau^2\right)$$

Οι εκτιμητές των συντελεστών β_1 και β_2 της μετα-παλινδρόμησης προκύπτουν με την μέθοδο σταθμισμένων ελαχίστων τετραγώνων με βάρη τα $w_i^* = \frac{1}{\sigma_i^2 + \tau^2}$. Επομένως για να

μπορέσουμε να εφαρμόσουμε αυτή την μέθοδο, θα πρέπει αρχικά να εκτιμήσουμε την μεταβλητότητα τ^2 .

Στις επόμενες παραγράφους θα παρουσιαστούν οι εκτιμητές που έχουν προταθεί για την εκτίμηση του τ^2 .

5.4.1.2.1 Εκτιμητής του τ^2 με τη Μέθοδο των Ροπών

Ο εκτιμητής του τ^2 με τη μέθοδο των ροπών όπως αναφέρθηκε στην παράγραφο 4.3.1, μπορεί να επεκταθεί και στην περίπτωση που υπάρχουν συμμεταβλητές. Εντούτοις θα πρέπει να τονίσουμε ότι αυτή η επέκταση αυτής της μεθόδου δεν είναι εύκολη για περισσότερες από μια συμμεταβλητές.

Οι *Thompson* και *Sharp* [1999] παρουσίασαν την μέθοδο εκτίμησης του τ^2 με την μέθοδο των ροπών στην περίπτωση που υπάρχει μόνο ένας επεξηγηματικός παράγοντας. Ο εκτιμητής του τ^2 δίνεται από τη σχέση:

$$\hat{\tau}^2 = \begin{cases} \frac{Q - (\kappa - 2)}{F(w, x)}, & \text{αν } Q > \kappa - 2 \\ 0, & \text{αλλιώς} \end{cases}$$

$$\text{όπου } F(w, x) = \sum_i w_i - \frac{\sum_i w_i^2 \sum_i w_i X_i^2 - 2 \sum_i w_i^2 X_i \sum_i w_i X_i + \sum_i w_i \sum_i w_i^2 X_i^2}{\sum_i w_i \sum_i w_i X_i^2 - \left(\sum_i w_i X_i\right)^2}$$

Στην συνέχεια, για να υπολογίσουμε τους συντελεστές β_1 και β_2 της μετα-παλινδρόμησης χρησιμοποιούμε την μέθοδο των σταθμισμένων ελαχίστων τετραγώνων με βάρη τα

$w_i^* = \left(\sigma_i^2 + \hat{\tau}^2 \right)^{-1}$. Ο εκτιμητής του τ^2 είναι αμερόληπτος αν αφαιρεθεί ο περιορισμός $\hat{\tau}^2 \geq 0$.

5.4.1.2.2 Εκτιμητής Μέγιστης Πιθανοφάνειας του τ^2 (MLE)

Ο εκτιμητής μέγιστης πιθανοφάνειας της παραμέτρου τ^2 μπορεί να υπολογιστεί χρησιμοποιώντας επαναληπτικά την ακόλουθη εξίσωση, [Thompson & Sharp, 1999] :

$$\tau^2 = \sum_i w_i^{*2} \left\{ \left(\hat{\theta}_i - \hat{\beta}_1 - \hat{\beta}_2 x_i \right)^2 - \sigma_i^2 \right\} / \sum_i w_i^{*2}, \quad \text{όπου} \quad w_i^* = \left(\sigma_i^2 + \hat{\tau}^2 \right)^{-1}.$$

Ξεκινώντας με $\hat{\tau}^2 = 0$ εκτιμάμε με την παλινδρόμηση τις παραμέτρους β_1 , β_2 και στην συνέχεια με την βοήθεια της προηγούμενης εξίσωσης υπολογίζουμε τη νέα την τιμή του τ^2 κ.ο.κ. μέχρι να επιτευχθεί η σύγκλιση. Στην πράξη η σύγκλιση είναι αργή και χρειάζεται μεγάλη προσοχή.

5.4.1.2.3 Εκτιμητής Περιορισμένης Μέγιστης Πιθανοφάνειας του τ^2 (REML)

Ο εκτιμητής περιορισμένης μέγιστης πιθανοφάνειας, είναι προτιμότερος από τον εκτιμητή μέγιστης πιθανοφάνειας γιατί παράγει αμερόληπτους εκτιμητές των διακυμάνσεων.

Ο εκτιμητής περιορισμένης πιθανοφάνειας της παραμέτρου τ^2 μπορεί να υπολογιστεί χρησιμοποιώντας επαναληπτικά την ακόλουθη εξίσωση [Thompson & Sharp, 1999]:

$$\tau^2 = \sum_i w_i^{*2} \left\{ \left[k / (k - 2) \right] \left(\hat{\theta}_i - \hat{\beta}_1 - \hat{\beta}_2 x_i \right)^2 - \sigma_i^2 \right\} / \sum_i w_i^{*2}$$

όπου τα βάρη δίνονται και πάλι από τη σχέση $w_i^* = \left(\sigma_i^2 + \hat{\tau}^2 \right)^{-1}$.

Ξεκινώντας με $\hat{\tau}^2 = 0$ εκτιμάμε με την παλινδρόμηση τις παραμέτρους β_1 , β_2 και στην συνέχεια με την βοήθεια της προηγούμενης εξίσωσης υπολογίζουμε τη νέα την τιμή του τ^2 κ.ο.κ. μέχρι να επιτευχθεί η σύγκλιση.

5.4.2 Ανάλυση υποομάδων (*Subgroup analysis*)

Η ανάλυση των υποομάδων (*subgroup analysis*) περιλαμβάνει το διαχωρισμό των στοιχείων όλων των συμμετεχόντων ασθενών σε υποομάδες, ώστε να γίνουν συγκρίσεις μεταξύ αυτών των ομάδων. Οι αναλύσεις αυτές γίνονται συνήθως σε υποσύνολα ασθενών (για παράδειγμα σε άνδρες ή γυναίκες), σε υποσύνολα μελετών (π.χ. σε σχέση με την γεωγραφική θέση της μελέτης) κ.λ.π. Αποτελούν μέσο έρευνας των ετερογενών αποτελεσμάτων ή γίνονται για να απαντήσουν σε συγκεκριμένα ερωτήματα που αφορούν τις ομάδες ασθενών, το είδος της θεραπευτικής μεθόδου ή το είδος της μελέτης. Βασική προϋπόθεση για την διεξαγωγή αυτών των αναλύσεων είναι να είναι γνωστά τα αντίστοιχα στοιχεία των ασθενών.

Οι ερευνητές θα πρέπει να είναι ιδιαίτερα προσεκτικοί όταν παρουσιάζουν τα συμπεράσματα αυτών των αναλύσεων. Η ανάλυση των υποομάδων είναι από τη φύση της παρατηρητική μέθοδος. Αν και τα άτομα τυχαιοποιούνται μέσα σε κάθε μελέτη, δεν τυχαιοποιούνται στο σύνολο όλων των μελετών. Επομένως η ανάλυση αυτή υφίσταται τους περιορισμούς μιας οποιασδήποτε παρατηρητικής έρευνας ενώ επιπλέον συγχρητικοί παράγοντες (*confounding*) σε σχέση με τα χαρακτηριστικά της μελέτης μπορούν να αλλοιώσουν τα αποτελέσματά της [Cochrane Reviewers' Handbook, 2004].

Αν η επίδραση της θεραπείας ή ο έλεγχος ετερογένειας σε μια υποομάδα είναι στατιστικά σημαντικός και σε μια άλλη ομάδα δεν είναι, τότε δεν μπορούμε να πούμε ότι ο παράγοντας με τον οποίο δημιουργήσαμε τις υποομάδες εξηγεί την ετερογένεια. Δεδομένου ότι οι υποομάδες περιέχουν διαφορετικά ποσά πληροφορίας, έχουν διαφορετικές δυνατότητες να ανιχνεύσουν μια επίδραση. Επομένως είναι παραπλανητικό να συγκρίνουμε απλά την στατιστική σημαντικότητα των αποτελεσμάτων. Παρ' όλα αυτά, οι αναλύσεις αυτές μπορούν

να χρησιμοποιηθούν σαν εργαλεία για την δημιουργία υποθέσεων ή για να παράγουν σημαντικές παρατηρήσεις σχετικά με το σύνολο των δεδομένων.

Τα άρθρα των *Oxman* [1992] και *Yusuf et al* [1991] παρέχουν χρήσιμες οδηγίες για την ανάλυση των υποομάδων.

ΚΕΦΑΛΑΙΟ 6

Απλές μέθοδοι για τον συνδυασμό μελετών

6.1 Μέθοδοι “*Vote-counting*”

Οι μέθοδοι “*vote-counting*” αποτελούν έναν από τους απλούστερους τρόπους για τον συνδυασμό των αποτελεσμάτων πολλών μελετών. Το μόνο που εξετάζεται από την κάθε μελέτη είναι η κατεύθυνση του αποτελέσματος (θετική ή αρνητική επίδραση της θεραπείας σε σχέση με το *control*). Επομένως αυτές οι μέθοδοι αγνοούν το μέγεθος της επίδρασης της θεραπείας σε σχέση με το *control* καθώς και την ακρίβεια του αποτελέσματος. Γι’ αυτούς τους λόγους οι μέθοδοι “*vote-counting*” προτείνονται μόνο στη χειρίστη περίπτωση που όλες οι μελέτες ή κάποιες από αυτές δεν αναφέρουν ούτε την επίδραση της θεραπείας, ούτε το *p-value* του αντίστοιχου ελέγχου.

6.1.1 Conventional “*vote-counting*”

Οι *Light* και *Smith* το 1971 ήταν οι πρώτοι που επίσημα περιέγραψαν τη μέθοδο “*vote-counting*”. Πολύ απλά, κάθε μελέτη καταγράφεται σε μια από τις τρεις κατηγορίες : a) θετική επίδραση, b) αρνητική επίδραση ή c) μηδενική επίδραση. Η κατηγορία με τις περισσότερες ψήφους-μελέτες θα δίνει και το τελικό αποτέλεσμα της επίδρασης της θεραπείας σε σχέση με το *control*. Όπως είναι φανερό η μέθοδος αυτή δεν θα μπορούσε να είναι πιο απλή. Τα μειονεκτήματά της είναι πολλά και αναφέρουμε ενδεικτικά τα εξής:

- i. Το μέγεθος δείγματος της κάθε μελέτης και η ακρίβεια του αντίστοιχου εκτιμητή δεν λαμβάνονται υπόψη στην ψηφοφορία.
- ii. Δεν παρέχει εκτιμητή της ολικής επίδρασης της θεραπείας σε σχέση με το *control*.

Συμπερασματικά μπορούμε να πούμε ότι η μέθοδος αυτή είναι αρκετά απλοϊκή, χωρίς στατιστική λογική και επειδή μπορεί να οδηγήσει σε εσφαλμένα συμπεράσματα δεν συνίσταται.

6.1.2 Προσημικό τεστ (*Sign test*)

Το προσημικό τεστ αποτελεί άλλη μια απλή μεθοδολογία με την οποία μπορούμε να διαπιστώσουμε αν η θεραπεία που εξετάζουμε είναι αποτελεσματική σε σχέση με το *control*. Σύμφωνα με τη μέθοδο, οι μελέτες με θετική επίδραση της θεραπείας παίρνουν το πρόσημο “+” ενώ οι μελέτες με αρνητική επίδραση παίρνουν το πρόσημο “-”.

Επειδή η πιθανότητα μιας μελέτης να έχει θετική επίδραση είναι ίση με την πιθανότητα να έχει αρνητική επίδραση, μπορούμε να θεωρήσουμε ότι το σύνολο των σημείων “+” προέρχεται από διωνυμικό πληθυσμό $B(k, p = 1/2)$ όπου k είναι το πλήθος των μελετών. Επομένως θα ελέγξουμε την υπόθεση $H_0 : p = 0$ έναντι της μονόπλευρης εναλλακτικής $H_a : p > 1/2$. Αν U είναι το πλήθος των θετικών πρόσημων, τότε με βάση τους πίνακες της διωνυμικής κατανομής θα αποφασίσουμε αν θα δεχτούμε ή αν θα απορρίψουμε την μηδενική υπόθεση.

Όπως και η προηγούμενη μέθοδος έτσι και αυτή, δεν λαμβάνει υπόψη το μέγεθος του δείγματος της κάθε μελέτης, την ακρίβεια του αντίστοιχου εκτιμητή και δεν παρέχει εκτιμητή της ολικής επίδρασης της θεραπείας σε σχέση με το *control*.

6.2 Συνδυασμός των *p-values*

Στις επόμενες παραγράφους θα παρουσιάσουμε μεθόδους συνδυασμού των *p-values* των στατιστικών τεστ τα οποία υπολογίζονται στις επιμέρους μελέτες. Οι μέθοδοι αυτές προτείνονται στις εξής δυο πιθανές περιπτώσεις: α) όταν κάποιες μελέτες δεν αναφέρουν τον εκτιμητή της επίδρασης της θεραπείας ούτε δίνουν στοιχεία για τον υπολογισμό του, αλλά αναφέρουν το *p-value* του αντίστοιχου ελέγχου και β) όταν οι μελέτες διαφέρουν πολύ στο σχεδιασμό που ο συνδυασμός των εκτιμητών της επίδρασης της θεραπείας δεν είναι

κατάλληλος. Όπως συμβαίνει και με τις μεθόδους “*vote-counting*”, έτσι και αυτές οι μεθοδολογίες έχουν το μειονέκτημα ότι δεν παρέχουν εκτιμητή της ολικής επίδρασης της θεραπείας.

Οι παρακάτω μέθοδοι αναφέρονται στον συνδυασμό *p-values* τα οποία προέρχονται από τον έλεγχο μονόπλευρων υποθέσεων. Έστω λοιπόν θ η παράμετρος που εκφράζει την διαφορά της θεραπείας σε σχέση με το *control*. Υποθέτουμε ότι αν οι θεραπείες είναι ισοδύναμες θα ισχύει $\theta = 0$, ενώ αν η εξεταζόμενη θεραπεία είναι πιο αποτελεσματική από την θεραπεία ελέγχου (*control*) θα ισχύει $\theta > 0$. Το ενδιαφέρον μας επομένως βρίσκεται στον έλεγχο της υπόθεσης $H_0 : \theta = 0$ έναντι της εναλλακτικής $H_a : \theta > 0$. Σε αυτόν τον έλεγχο υποθέσεων, το *p-value* ορίζεται ως η πιθανότητα:

$$p\text{-value} = P(T(X) \geq T(x_0) / H_0)$$

όπου $T(X)$ είναι το αντίστοιχο στατιστικό ελέγχου της δοσμένης υπόθεσης. Το *p-value* επομένως η είναι η πιθανότητα εμφάνισης μιας τιμής του στατιστικού τουλάχιστον τόσο ακραίας, όσο αυτή που παρατηρήσαμε.

Έστω p_{1i} το *p-value* του αντίστοιχου μονόπλευρου ελέγχου της i μελέτης. Αν σε μία μελέτη δίνεται το *p-value* του δίπλευρου ελέγχου, έστω p_{2i} , τότε μπορούμε να υπολογίσουμε το *p-value* του μονόπλευρου ελέγχου. Αν ο εκτιμητής της παραμέτρου θ είναι θετικός τότε είναι γνωστό ότι $p_{1i} = \frac{p_{2i}}{2}$, ενώ αν ο εκτιμητής είναι αρνητικός θα είναι

$$p_{1i} = 1 - \frac{p_{2i}}{2}.$$

Αν πρόκειται να συνδυαστούν *p-values* τα οποία προέρχονται από διαφορετικές μεταβλητές απόκρισης, τότε θα πρέπει να προσέξουμε οι τιμές p_{1i} να προκύπτουν από ελέγχους στους οποίους η εναλλακτική υποθέτει ότι η νέα θεραπεία είναι καλύτερη από την θεραπεία ελέγχου. Στην συνέχεια θα παρουσιάσουμε τέσσερις μεθοδολογίες για τον συνδυασμό των *p-values*.

6.2.1 Μέθοδος του Tippett (*Minimum p method*)

Πρώτος ο *Tippett* [1931] με το τεστ γνωστό ως “*minimum p*”, απορρίπτει την μηδενική υπόθεση ότι οι δυο θεραπείες είναι ισοδύναμες σε όλες τις μελέτες, έναντι της εναλλακτικής υπόθεσης ότι σε τουλάχιστον μια μελέτη η νέα θεραπεία είναι καλύτερη από το *control*, αν τουλάχιστον ένα από τα p_{1i} είναι μικρότερο της τιμής:

$$a^* = 1 - (1 - a)^{1/k}$$

όπου a είναι το επίπεδο σημαντικότητας για τον έλεγχο της συνδυασμένης μηδενικής υπόθεσης (συνήθως $\alpha=5\%$) και k είναι το πλήθος των μελετών .

Δίνοντας μια πιο ακριβή περιγραφή της μεθόδου μπορούμε να πούμε ότι η μηδενική υπόθεση $H_0 : \theta_i = 0 \forall i$ απορρίπτεται αν:

$$\text{Min}(p_1, p_2, \dots, p_k) = p_{[1]} < a^* = 1 - (1 - a)^{1/k}$$

6.2.2 Μέθοδος του Fisher (*Sum of logs method*)

Ο *Fisher* [1932] κατασκεύασε ένα στατιστικό το οποίο στηρίζεται στα p_{1i} , για τον έλεγχο της υπόθεσης ότι οι δυο θεραπείες είναι ισοδύναμες σε όλες τις μελέτες, $H_0 : \theta_i = 0, \forall i = 1, 2, \dots, k$, έναντι της εναλλακτικής ότι σε τουλάχιστον μια μελέτη η θεραπεία είναι καλύτερη από το *control*, $H_a : \exists i = 1, 2, \dots, k : \theta_i > 0$. Τα p_{1i} κατανέμονται ομοιόμορφα στο διάστημα $U(0,1)$ επομένως το στατιστικό $T_i = -2 \log(p_{1i})$ ακολουθεί εκθετική κατανομή με παράμετρο $1/2$.

Επειδή οι μελέτες είναι ανεξάρτητες, το στατιστικό:

$$P = \sum_{i=1}^k T_i = \sum_{i=1}^k (-2 \log p_{1i}) = -2 \log \prod_{i=1}^k p_{1i}$$

κάτω από την μηδενική υπόθεση θα ακολουθεί κατανομή $\text{Gamma}\left(k, \frac{1}{2}\right)$ ή ισοδύναμα χ^2 κατανομή με $2k$ βαθμούς ελευθερίας.

Αν θέλουμε να ελέγξουμε την μηδενική υπόθεση ότι οι θεραπείες είναι ισοδύναμες σε όλες τις μελέτες, έναντι της εναλλακτικής ότι σε τουλάχιστον μια μελέτη η υπό εξέταση θεραπεία είναι χειρότερη από την θεραπεία *control*, τότε ανάλογα θα εργαζόμασταν με το στατιστικό:

$$P^- = -2 \sum_{i=1}^k \log(1 - p_{1i})$$

το οποίο πάλι κάτω από την μηδενική υπόθεση ακολουθεί χ^2 κατανομή με $2k$ βαθμούς ελευθερίας.

6.2.3 Μέθοδος του Stouffer (*Sum of z's method*)

Η μέθοδος του *Stouffer et al.* [1949] βασίζεται στο στατιστικό, $U_k = \frac{\sum_{i=1}^k z(p_{1i})}{\sqrt{k}}$

όπου το $z(p_{1i})$ είναι το p_{1i} -ποσοστημόριο της τυπικής κανονικής κατανομής, το οποίο δίνεται από την σχέση $z(p_{1i}) = 1 - \Phi^{-1}(1 - p_{1i})$. Το στατιστικό U_k κάτω από την μηδενική υπόθεση ακολουθεί τυπική κανονική κατανομή. Επομένως αν είναι $U_k > z(a)$, η μηδενική υπόθεση απορρίπτεται και συμπεραίνουμε ότι η νέα θεραπεία είναι καλύτερη του *control* σε τουλάχιστον μια μελέτη.

Το βασικό μειονέκτημα των τριών προηγούμενων μεθοδολογιών είναι ότι σε κάθε μελέτη δίνεται το ίδιο βάρος. Το πιο λογικό όμως είναι ότι επειδή οι μεγαλύτερες μελέτες δίνουν πιο ακριβή αποτελέσματα, θα πρέπει να τους δοθούν μεγαλύτερα βάρη.

6.2.4 Μέθοδος των Mosteller και Bush

Οι *Mosteller* και *Bush* [1954] πρότειναν μια γενίκευση της μεθόδου του *Stouffer et al* με την οποία δίνονται βάρη στα αντίστοιχα $z(p_{1i})$. Η νέα αυτή προσέγγιση βασίζεται στο στατιστικό,

$$U_{wk} = \frac{\sum_{i=1}^k g_i \cdot z(p_{1i})}{\sqrt{\sum_{i=1}^k g_i^2}}$$

Το στατιστικό U_{wk} κάτω από την μηδενική υπόθεση ακολουθεί τυπική κανονική κατανομή.

Αν τα βάρη $g_i, i = 1, \dots, k$ θεωρηθούν όλα ίσα με $1/\sqrt{k}$, τότε η μεθοδολογία των *Mosteller* και *Bush* ταυτίζεται με την μεθοδολογία του *Stouffer et al*, αφού θα είναι $U_{wk} = U_k$. Αν δεν έχουμε κάποια γνώση για τα βάρη g_i , τότε μια πρακτική επιλογή θα ήταν η $g_i = \sqrt{n_i}$, όπου n_i είναι ο συνολικός αριθμός των ασθενών που λαμβάνουν μέρος στην μελέτη $i = 1, 2, \dots, k$.

Έστω ότι σε κάθε μελέτη έχουμε την ίδια μεταβλητή απόκρισης και εκτιμάμε την ίδια παράμετρο η οποία εκφράζει την διαφορά των θεραπειών, $\theta_i, i = 1, 2, \dots, k$. Σύμφωνα με όσα έχουμε πει στο μοντέλο σταθερών επιδράσεων (παρ. 4.2.3) ισχύει:

$$\hat{\theta}_i \sim N(\theta_i, w_i^{-1}), \text{ όπου } w_i = 1/\text{Var}(\hat{\theta}_i).$$

Αν τα p_{1i} υπολογιστούν κάτω από αυτή την υπόθεση, τότε θα ισχύουν οι σχέσεις

$$p_{1i} = 1 - \Phi\left(\hat{\theta}_i \sqrt{w_i}\right) \text{ και } z(p_{1i}) = \hat{\theta}_i \sqrt{w_i}.$$

Αν επιλέξουμε ως βάρη τα $g_i = \sqrt{w_i}$, τότε θα είναι :

$$U_{wr} = \frac{\sum_{i=1}^k \hat{\theta}_i \cdot \sqrt{w_i} \cdot \sqrt{w_i}}{\sqrt{\sum_{i=1}^k w_i}} = \frac{\sum_{i=1}^k \hat{\theta}_i \cdot w_i}{\sqrt{\sum_{i=1}^k w_i}}$$

Το U_{wk}^2 επομένως κάτω από αυτές τις υποθέσεις ταυτίζεται με το στατιστικό U του μοντέλου σταθερών επιδράσεων.

ΚΕΦΑΛΑΙΟ 7

Μετανάλυση με τα ατομικά δεδομένα των ασθενών

7.1 Εισαγωγή

Οι περισσότερες μεταanalύσεις στηρίζονται σε συνοπτικά στατιστικά στοιχεία τα οποία αναφέρονται στις δημοσιευμένες κλινικές δοκιμές. Τελευταία όμως ο αριθμός των μεταanalύσεων οι οποίες βασίζονται στα ατομικά δεδομένα των ασθενών (*Individual Patient Data*) ολοένα και αυξάνει. Η γνώση των πρωτογενών δεδομένων επιτρέπει, όχι μόνο τον έλεγχο των αποτελεσμάτων στα οποία καταλήγουν οι μεμονωμένες μελέτες, αλλά και την δυνατότητα εφαρμογής πολυπλοκότερων μοντέλων από τα οποία προκύπτουν πιο αξιόπιστα αποτελέσματα. Γι' αυτούς τους λόγους η σύνθεση κλινικών μελετών με χρήση των ατομικών δεδομένων των ασθενών αποτελούν το “*gold standard*” των μεταanalύσεων.

Η στατιστική ανάλυση των δεδομένων μιας μετανάλυσης με I.P.D μοιάζει πολύ με την ανάλυση που εφαρμόζεται σε μια πολυκεντρική μελέτη, μόνο που σε αυτή τη περίπτωση αντί για “κέντρα” υπάρχουν “μελέτες”. Από την κάθε μελέτη θα πρέπει για κάθε ασθενή να συλλεχθούν τουλάχιστον τα εξής δεδομένα: η ταυτότητα του ασθενή, η θεραπεία που εφαρμόστηκε, η τιμή της απόκρισης, η ημέρα τυχαιοποίησης και η ημέρα που καταγράφηκε το αναμενόμενο γεγονός (αν πρόκειται να γίνει ανάλυση επιβίωσης). Στην συνέχεια τα δεδομένα αυτά θα πρέπει να ελεγχθούν και να κωδικοποιηθούν προκειμένου να εφαρμοστεί η κατάλληλη στατιστική ανάλυση.

Εντούτοις δεν είναι πάντα εύκολο να πραγματοποιηθούν μεταanalύσεις με χρήση των πρωτογενών δεδομένων. Τα στοιχεία των ασθενών είναι σπάνια διαθέσιμα στις δημοσιευμένες μελέτες λόγω του απόρρητου των φορέων που διεξάγουν τις κλινικές μελέτες. Επιπλέον η ανάλυση ενός τόσο μεγάλου όγκου πληροφοριών είναι αρκετά χρονοβόρα και δαπανηρή. Συνήθως λόγω όλων των προηγούμενων προβλημάτων, οι μεταanalύσεις αυτού

του τύπου πραγματοποιούνται συνήθως από φαρμακευτικές εταιρείες οι οποίες έχουν πρόσβαση στα αρχεία των ασθενών.

Στον πίνακα 7.1 παρουσιάζονται τα πλεονεκτήματα και τα μειονεκτήματα της μετανάλυσης με I.P.D σε σχέση με την μετανάλυση με χρήση συνοπτικών στατιστικών.

Advantages

Can use common definitions, coding and cut-points for variables
 Allows adjustment for the same variables
 Allows estimation of effects in identically defined subgroups
 Increases power in subgroup analyses
 Can address questions not assessed in the original publication

Disadvantages

More expensive
 More time-consuming
 Subject to bias due to exclusion because of data unavailability

Πίνακας 7.1 Πλεονεκτήματα και μειονεκτήματα της I.P.D μετανάλυσης σε σχέση με την μετανάλυση συνοπτικών στατιστικών, Peititi [2000]

Μια ερώτηση που έχει απασχολήσει πολλούς ερευνητές είναι αν η μετανάλυση των συνοπτικών δεδομένων (πολλές φορές αναφέρεται ως MAL: MetAnalysis of Literature) καταλήγει σε διαφορετικά συμπεράσματα από την μετανάλυση με I.P.D (αντίστοιχα αναφέρεται ως MAP: MetAnalysis of Patient data). Οι *Jeng, Scott* και *Burmeister* [1995], και οι *Stewart* και *Parmar* [1993] διαπίστωσαν ότι τα αποτελέσματα μπορεί να διαφέρουν ανάλογα με το ποιες μελέτες συμπεριλήφθηκαν σε κάθε ανάλυση ή ανάλογα με το ποια μεταβλητή χρησιμοποιήθηκε. Ο *Steinberg* [1997] έδειξε ότι όταν οι μελέτες είναι παρόμοιες και στις δυο προσεγγίσεις χρησιμοποιούνται οι ίδιες μεταβλητές τα συμπεράσματα της MAP είναι σε συμφωνία με τα συμπεράσματα της MAL. Οι *Ingram* και *Sampson* [1998] κατέληξαν στο συμπέρασμα ότι όταν οι μελέτες είναι ομοιογενείς, τότε ο εκτιμητής της ολικής επίδρασης της θεραπείας (παρ. 4.2.3) ταυτίζεται με τον εκτιμητή ελαχίστων τετραγώνων του two-way ANOVA στο μοντέλο χωρίς αλληλεπίδραση (βλέπε παρ 7.2.1).

Στην συνέχεια θα παρουσιάσουμε τα απλούστερα μοντέλα τα οποία μπορούμε να χρησιμοποιήσουμε στην περίπτωση που τα δεδομένα μας είναι συνεχή ή δίτιμα. Στην παράγραφο 7.2 θα παρουσιαστούν τα μοντέλα σταθερών επιδράσεων για συνεχή και δίτιμη μεταβλητή απόκρισης, ενώ στην παράγραφο 7.3 θα παρουσιαστούν τα αντίστοιχα μοντέλα τυχαίων επιδράσεων.

7.2 Μοντέλο σταθερών επιδράσεων (*Fixed Effects Model*)

7.2.1 Συνεχής μεταβλητή απόκρισης

Έστω y_{ij} η απόκριση του j ασθενή της i μελέτης όπου $j=1,2,\dots,n_i$ και $i=1,2,\dots,k$. Έστω $n = \sum_{i=1}^k n_i$ το πλήθος των ασθενών σε όλες τις μελέτες. Υποθέτουμε ότι η τυχαία μεταβλητή Y_{ij} ακολουθεί κανονική κατανομή με μέση τιμή μ_{ij} και διακύμανση σ^2 , δηλαδή είναι $y_{ij} \sim N(\mu_{ij}, \sigma^2)$. Ακόμα και αν τα δεδομένα δεν είναι κανονικά κατανεμημένα, μπορούμε με κατάλληλο μετασχηματισμό να τα τροποποιήσουμε ώστε να ικανοποιούν την υπόθεση της κανονικότητας.

Το αντίστοιχο γραμμικό μοντέλο μπορεί να γραφεί με τη μορφή $y_{ij} = \mu_{ij} + \varepsilon_{ij}$ όπου τα σφάλματα ε_{ij} είναι ασυσχέτιστα μεταξύ τους και ακολουθούν κανονική κατανομή με μέση τιμή 0 και διακύμανση σ^2 , δηλαδή $\varepsilon_{ij} \sim N(0, \sigma^2)$. Στο μοντέλο αυτό υποθέτουμε ότι η μεταβλητότητα μεταξύ των ασθενών μιας μελέτης (*between patient variability*) σ^2 είναι η ίδια για όλες τις μελέτες.

Όπως αναφέρθηκε στο 4^ο κεφάλαιο, το μοντέλο σταθερών επιδράσεων υποθέτει ότι η επίδραση της θεραπείας είναι η ίδια σε όλες τις μελέτες. Το συστηματικό μέρος του μοντέλου μπορεί να γραφεί στην μορφή:

$$\mu_{ij} = \alpha + \beta_{0i} + \beta_1 x_{1ij}, \text{ όπου } i=1,2,\dots,k \text{ και } j=1,2,\dots,n_i$$

Ο παράγοντας β_{0i} εκφράζει την επίδραση της i μελέτης και μπορεί να γραφεί στην μορφή $\beta_{0i} = \beta_{01}x_{01ij} + \beta_{02}x_{02ij} + \dots + \beta_{0(k-1)}x_{0(k-1)ij}$, όπου $x_{0hij} = 1$ αν ο ασθενής ανήκει στην h μελέτη και $x_{0hij} = 0$ διαφορετικά. Ο παράγοντας x_{1ij} παίρνει την τιμή “1” αν ο ασθενής ανήκει στην ομάδα “*treatment*” αλλιώς παίρνει τη τιμή “0” αν ο ασθενής ανήκει στην ομάδα

“control”. Ο συντελεστής β_1 είναι ίσος με την επίδραση της θεραπείας σε σχέση με το *control*.

Το προηγούμενο μοντέλο σταθερών επιδράσεων είναι ένα κλασσικό μοντέλο ANOVA, επομένως οι εκτιμητές των παραμέτρων και της διακύμανσης σ^2 μπορούν να υπολογιστούν με τη μέθοδο ελαχίστων τετραγώνων.

Για τον έλεγχο της υπόθεσης $H_0 : \beta_1 = 0$ (ότι δηλαδή η επίδραση της θεραπείας σε σχέση με το *control* είναι μηδέν ή αντίστοιχα οι δυο θεραπείες είναι ισοδύναμες), συγκρίνουμε το αρχικό μοντέλο με το μοντέλο $\mu_{ij} = \alpha + \beta_{0i}$ το οποίο περιέχει μόνο τον παράγοντα β_{0i} που δηλώνει την επίδραση της i μελέτης. Η σύγκριση των δυο μοντέλων θα γίνει με βάση τη διαφορά των αθροισμάτων των τετραγώνων των μοντέλων (F - τεστ). Το στατιστικό που προκύπτει ακολουθεί κάτω από την μηδενική υπόθεση κατανομή $F_{1, n-k-1}$, αφού το αρχικό μοντέλο έχει $k+1$ βαθμούς ελευθερίας ενώ το δεύτερο μοντέλο έχει k βαθμούς ελευθερίας.

Στο σημείο αυτό θα πρέπει να αναφέρουμε ότι σύμφωνα με τους *Olkin* και *Sampson*[1998] ο εκτιμητής $\hat{\beta}_1$ της επίδρασης της θεραπείας με βάση το προηγούμενο γενικευμένο γραμμικό μοντέλο, είναι ο ίδιος με τον αντίστοιχο εκτιμητή $\hat{\theta}$ που προκύπτει με τον συνδυασμό των εκτιμητών (παράγραφος 4.2).

Για να ελέγξουμε την υπόθεση της ετερογένειας των μελετών, συγκρίνουμε το αρχικό μοντέλο με το μοντέλο $\mu_{ij} = \alpha + \beta_{0i} + \beta_{1i}x_{1ij}$, το οποίο περιέχει επιπλέον τον παράγοντα αλληλεπίδρασης της μελέτης με τη θεραπεία. Το στατιστικό που προκύπτει ακολουθεί κάτω από την μηδενική υπόθεση κατανομή $F_{k-1, n-2k}$ αφού το αρχικό μοντέλο έχει $k+1$ βαθμούς ελευθερίας ενώ το δεύτερο μοντέλο έχει $2k$ βαθμούς ελευθερίας.

Όπως αναφέρθηκε στη αρχή της παραγράφου, στο μοντέλο σταθερών επιδράσεων υποθέσαμε ότι τα σφάλματα ε_{ij} ακολουθούν κανονική κατανομή με μέση τιμή 0 και κοινή διακύμανση σ^2 για όλους τους ασθενείς. Η υπόθεση αυτή μπορεί να ελεγχθεί χρησιμοποιώντας το *Bartlett's test*, το οποίο αναφέρθηκε στην παράγραφο 3.2.3. Στην

περίπτωση που η υπόθεση της κοινής διακύμανσης απορριφθεί μπορούμε να ακολουθήσουμε έναν από τους παρακάτω εναλλακτικούς τρόπους ανάλυσης:

Η πρώτη προσέγγιση είναι να θεωρήσουμε το ίδιο μοντέλο $Y_{ij} = \alpha + \beta_{0i} + \beta_1 x_{1ij} + \varepsilon_{ij}$ όπου όμως τα σφάλματα ε_{ij} θα ακολουθούν κανονική κατανομή με μέση τιμή 0 και διακύμανση σ_i^2 . Η εκτίμηση των παραμέτρων και ο έλεγχος υποθέσεων θα γίνει με βάση την ανάλυση των γενικευμένων γραμμικών μικτών μοντέλων (*general linear mixed models*). Η δεύτερη προσέγγιση είναι να θεωρήσουμε ότι το μέτρο που εκφράζει την διαφορά των θεραπειών στις δυο ομάδες είναι η τυποποιημένη διαφορά των μέσων τιμών. Στην συνέχεια μπορούμε να ακολουθήσουμε την αντίστοιχη μη παραμετρική μεθοδολογία του 4^{ου} κεφαλαίου, για τον συνδυασμό των εκτιμητών που θα προκύψουν από τις μελέτες.

7.2.2 Δίτιμη μεταβλητή απόκρισης

Έστω y_{ij} η απόκριση του j ασθενή της i μελέτης όπου $j=1,2,\dots,n_i$ και $i=1,2,\dots,k$. Επειδή η μεταβλητή απόκρισης είναι δίτιμη, θεωρούμε ότι $y_{ij}=1$ αν ο j ασθενής της i μελέτης είχε επιτυχία και $y_{ij}=0$ αν ο j ασθενής της i μελέτης είχε αποτυχία στη θεραπεία. Η τυχαία μεταβλητή Y_{ij} ακολουθεί κατανομή *Bernoulli* με πιθανότητα επιτυχίας έστω π_{ij} , αναμενόμενη τιμή ίση με π_{ij} και διακύμανση ίση με $\pi_{ij}(1-\pi_{ij})$.

Για να εξετάσουμε την εξάρτηση της πιθανότητας π_{ij} με τις επεξηγηματικές μεταβλητές, θα πρέπει να χρησιμοποιήσουμε μετασχηματισμό ο οποίος να μετατρέπει τιμές του διαστήματος $(0,1)$ σε τιμές της πραγματικής ευθείας $(-\infty, +\infty)$. Για αυτό το λόγο χρησιμοποιούμε *logit* σύνδεσμο και καταλήγουμε στο αντίστοιχο μοντέλο σταθερών επιδράσεων:

$$\log it(\pi_{ij}) = \log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \alpha + \beta_{0i} + \beta_1 x_{1ij}$$

Ο παράγοντας β_{0i} είναι η επίδραση της i μελέτης και μπορεί να γραφεί στην μορφή $\beta_{0i} = \beta_{01}x_{01ij} + \beta_{02}x_{02ij} + \dots + \beta_{0(k-1)}x_{0(k-1)ij}$, όπου $x_{0hij} = 1$ αν ο ασθενής ανήκει στην h μελέτη και $x_{0hij} = 0$ διαφορετικά.. Ο παράγοντας x_{1ij} παίρνει την τιμή “1” αν ο ασθενής ανήκει στην ομάδα “*treatment*” αλλιώς παίρνει τη τιμή “0” αν ο ασθενής ανήκει στην ομάδα “*control*”. Η παράμετρος β_1 είναι ίση με τον λογάριθμο του ολικού λόγου των σχετικών πιθανοτήτων (*log-odds ratio*) μεταξύ των δυο ομάδων.

Για τον έλεγχο της υπόθεσης $H_0 : \beta_1 = 0$ (η επίδραση της θεραπείας σε σχέση με το *control* είναι μηδέν ή αντίστοιχα οι δυο θεραπείες είναι ισοδύναμες) συγκρίνουμε το αρχικό μοντέλο με το μοντέλο $\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \alpha + \beta_{0i}$ το οποίο περιέχει μόνο τον παράγοντα β_{0i} ο οποίος δηλώνει την επίδραση της i μελέτης. Επειδή το αρχικό μοντέλο έχει $k+1$ βαθμούς ελευθερίας και το δεύτερο μοντέλο έχει k βαθμούς ελευθερίας, η διαφορά της απόκλισης των δυο μοντέλων ($-2\ddot{a}$ *log-likelihood*) θα συγκριθεί με την χ^2 κατανομή με 1 βαθμό ελευθερίας.

Σύμφωνα με την *Whitehead* [2002], ο εκτιμητής της ολικής επίδρασης της θεραπείας και το αντίστοιχο τυπικό σφάλμα που προκύπτουν με την MAP προσέγγιση, είναι ελαφρώς διαφορετικά σε σχέση με τα αποτελέσματα της MAL προσέγγισης.

Για να ελέγξουμε την ετερογένεια μεταξύ των μελετών, θα συγκρίνουμε το αρχικό μοντέλο με το μοντέλο $\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \alpha + \beta_{0i} + \beta_{1i}x_{1ij}$, το οποίο περιέχει επιπλέον τον παράγοντα αλληλεπίδρασης της μελέτης με τη θεραπεία. Το μοντέλο αυτό έχει $2k$ βαθμούς ελευθερίας επομένως η διαφορά της απόκλισης των δυο μοντέλων ($-2\ddot{a}$ *log-likelihood*) συγκρίνεται με την χ^2 κατανομή με $k-1$ βαθμούς ελευθερίας.

Περαιτέρω για την ανάλυση των δίτιμων δεδομένων με άλλες συναρτήσεις σύνδεσης (*link function*) αναφέρονται στους *Collett* [1991] και *McCullagh* [1989].

7.3 Μοντέλο τυχαίων επιδράσεων (*Random Effects Model*)

7.3.1 Συνεχής μεταβλητή απόκρισης

Στα μοντέλα που παρουσιάστηκαν στις προηγούμενες παραγράφους για την ανάλυση I.P.D θεωρήσαμε ότι η επίδραση της θεραπείας είναι η ίδια σε όλες τις μελέτες που συμπεριλαμβάνονται στην μετανάλυση. Σε μερικές περιπτώσεις η υπόθεση ότι η επίδραση της θεραπείας διαφέρει από μελέτη σε μελέτη μπορεί όμως να είναι πιο κατάλληλη.

Η ανάλυσή μας περιλαμβάνει δυο παραμέτρους: τους ασθενείς και τις μελέτες. Επειδή οι ασθενείς περιέχονται στις μελέτες (*nested model*) το μοντέλο που θα εξετάσουμε θα έχει ιεραρχικό χαρακτήρα με δυο επίπεδα (*multilevel study*): οι ασθενείς θα ανήκουν στο πρώτο επίπεδο και οι μελέτες στο δεύτερο επίπεδο. Η συγκεκριμένη δομή του μοντέλου μας επιτρέπει να ληφθεί υπόψη αλλά και να αναλυθεί η φυσική ιεραρχία των μονάδων που περιλαμβάνονται στην στατιστική ανάλυση.

Έστω y_{ij} η απόκριση του j ασθενή της i μελέτης όπου $j=1,2,\dots,n_i$ και $i=1,2,\dots,k$. Το μοντέλο τυχαίων επιδράσεων θα έχει την μορφή:

$$y_{ij} = \alpha + \beta_{0i} + \beta_{1i}x_{1ij} + \varepsilon_{ij}$$

$$\beta_{1i} = \beta_1 + \nu_{1i}$$

στο οποίο υποθέτουμε ότι:

$$\varepsilon_{ij} \sim N(0, \sigma_i^2) \quad \text{και} \quad \nu_{1i} \sim N(0, \tau^2)$$

Αν συνδυάσουμε τις προηγούμενες υποθέσεις, το μοντέλο θα έχει την μορφή:

$$y_{ij} = \alpha + \beta_{0i} + \beta_1 x_{1ij} + \nu_{1i} x_{1ij} + \varepsilon_{ij}, \quad \text{όπου} \quad \varepsilon_{ij} \sim N(0, \sigma_i^2) \quad \text{και} \quad \nu_{1i} \sim N(0, \tau^2)$$

και θα ανήκει στην κατηγορία των μικτών μοντέλων αφού περιέχει σταθερές και τυχαίες επιδράσεις.

Ο παράγοντας β_{0i} εκφράζει την επίδραση της i μελέτης και μπορεί να γραφεί στην μορφή $\beta_{0i} = \beta_{01}x_{01ij} + \beta_{02}x_{02ij} + \dots + \beta_{0(r-1)}x_{0(r-1)ij}$, όπου $x_{0hij} = 1$ αν ο ασθενής ανήκει

στην h μελέτη και $x_{0hij} = 0$ διαφορετικά. Ο παράγοντας x_{1ij} παίρνει την τιμή “1” αν ο ασθενής ανήκει στην ομάδα “*treatment*” αλλιώς παίρνει τη τιμή “0” αν ο ασθενής ανήκει στην ομάδα “*control*”. Ο συντελεστής β_1 αντιπροσωπεύει την διαφορά των μέσων τιμών των δυο ομάδων, δηλαδή της επίδραση της θεραπείας σε σχέση με το *control*. Επιπλέον έχουμε υποθέσει ότι η μεταβλητότητα μεταξύ των ασθενών (*between-patient variance*) διαφέρει από μελέτη σε μελέτη. Οι τυχαίοι παράγοντες ε_{ij} και ν_i θεωρούνται ασυσχέτιστοι. Αν θέλουμε να αντιστοιχίσουμε το προηγούμενο μοντέλο με το μοντέλο της παραγράφου 4.3 αρκεί να θέσουμε $\beta_1 = \theta$ και $\beta_{1i} = \theta_i$.

Η εκτίμηση των παραμέτρων a , β_{0i} και β_1 αλλά και των διακυμάνσεων σ_i^2 και τ^2 γίνεται με την μέθοδο μέγιστης πιθανοφάνειας. Οι εκτιμητές περιορισμένης πιθανοφάνειας (REML) είναι προτιμότεροι από τους εκτιμητές μέγιστης πιθανοφάνειας (MLE), γιατί προκύπτουν αμερόληπτοι εκτιμητές των παραμέτρων μεταβλητότητας.

Για περισσότερες πληροφορίες σχετικά με τα μικτά γραμμικά μοντέλα, την εκτίμηση των παραμέτρων και τον έλεγχο υποθέσεων ο αναγνώστης παραπέμπεται στην *Whitehead* [2002].

7.3.2 Δίτιμη μεταβλητή απόκρισης

Έστω y_{ij} η απόκριση του j ασθενή της i μελέτης όπου $j=1,2,\dots,n_i$ και $i=1,2,\dots,k$. Επειδή η μεταβλητή απόκρισης είναι δίτιμη, θεωρούμε ότι $y_{ij}=1$ αν ο j ασθενής της i μελέτης είχε επιτυχία και $y_{ij}=0$ αν ο j ασθενής της i μελέτης είχε αποτυχία στη θεραπεία. Η τυχαία μεταβλητή Y_{ij} ακολουθεί κατανομή *Bernoulli* με πιθανότητα επιτυχίας έστω π_{ij} , αναμενόμενη τιμή ίση με π_{ij} και διακύμανση ίση με $\pi_{ij}(1-\pi_{ij})$.

Για να εξετάσουμε την εξάρτηση της πιθανότητας π_{ij} με τις επεξηγηματικές μεταβλητές, θα πρέπει να χρησιμοποιήσουμε μετασχηματισμό ο οποίος να μετατρέπει τιμές του διαστήματος $(0,1)$ σε τιμές της πραγματικής ευθείας $(-\infty,+\infty)$. Αυτό το πετυχαίνουμε

χρησιμοποιώντας *logit* σύνδεσμο. Επειδή στο μοντέλο τυχαίων επιδράσεων η επίδραση της θεραπείας διαφέρει από μελέτη σε μελέτη, προκύπτει το εξής μοντέλο:

$$\log it(\pi_{ij}) = \log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \alpha + \beta_{0i} + \beta_{1i}x_{1ij}$$

όπου $\beta_{1i} = \beta_1 + \nu_{1i}$ και $\nu_{1i} \sim N(0, \tau^2)$

Αν συνδυάσουμε τις προηγούμενες υποθέσεις, το *logit* μοντέλο θα έχει την μορφή:

$$\log it(\pi_{ij}) = \log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \alpha + \beta_{0i} + \beta_1x_{1ij} + \nu_{1i}x_{1ij}$$

και θα ανήκει στην κατηγορία των μικτών μοντέλων αφού περιέχει σταθερές και τυχαίες επιδράσεις.

Για περισσότερες πληροφορίες σχετικά με την ανάλυση του προηγούμενου μοντέλου ο αναγνώστης παραπέμπεται στο άρθρο της *Turner* [2000].

ΚΕΦΑΛΑΙΟ 8

Αξιολόγηση του μοντέλου

8.1 Εισαγωγή

Όταν τα δεδομένα συλλεχθούν και αναλυθούν ο ερευνητής θα πρέπει να ελέγξει την καταλληλότητα των υποθέσεων που έχουν γίνει. Σύμφωνα με τον *Normand* [1999] υπάρχουν δυο κριτήρια με τα οποία μπορούμε να ελέγξουμε την εγκυρότητα των αποτελεσμάτων της μετανάλυσης: η ανάλυση ευαισθησίας (*sensitivity analysis*) και ο έλεγχος για την ανίχνευση και την διόρθωση της μεροληψίας.

8.2 Ανάλυση Ευαισθησίας (*Sensitivity Analysis*)

Επειδή υπάρχουν διαφορετικοί τρόποι για να πραγματοποιηθεί μια μετανάλυση, οι ερευνητές θα πρέπει να αναρωτηθούν: πόσο ευαίσθητα είναι τα αποτελέσματα της ανάλυσης σε αλλαγές της αρχικής μεθοδολογικής προσέγγισης; Η ανάλυση ευαισθησίας είναι μια προσέγγιση με την οποία μπορούμε να ελέγξουμε το πόσο ανθεκτικά (*robust*) είναι τα αποτελέσματα σε σχέση με τα κριτήρια εισαγωγής των κλινικών μελετών και τις βασικές υποθέσεις που λήφθηκαν στην πορεία της ανάλυσης.

Στην ανάλυση ευαισθησίας θα πρέπει να εξεταστούν ένα ή περισσότερα από τα επόμενα σενάρια τα οποία προτείνονται στο *Cochrane Reviewer's Handbook* [2004]:

- i. Αλλαγή των κριτηρίων εισαγωγής των μελετών, των χαρακτηριστικών των ασθενών, του είδους της θεραπείας ή της μεταβλητής απόκρισης.
- ii. Εισαγωγή ή εξαίρεση μελετών για τις οποίες υπάρχει σχετικά μικρή ασάφεια για το αν πληρούν τα κριτήρια εισαγωγής.
- iii. Επανάληψη της ανάλυσης χρησιμοποιώντας ένα λογικό εύρος για τα αποτελέσματα των μελετών, για όσες από αυτές υπάρχει αβεβαιότητα σε σχέση με το πώς έγινε η καταγραφή των στοιχείων τους.

- iv. Επανάληψη της ανάλυσης χρησιμοποιώντας ένα λογικό εύρος τιμών για τα ελλιπή δεδομένα (*missing data*).
- v. Επανάληψη της ανάλυσης χρησιμοποιώντας διαφορετική στατιστική προσέγγιση.

Αν επομένως υπάρχουν μελέτες οι οποίες επηρεάζουν σε μεγάλο βαθμό το ολικό αποτέλεσμα της μετανάλυσης, τότε με την ανάλυση ευαισθησίας μπορούμε να δούμε πόσο αλλάζουν τα συμπεράσματα όταν αυτές οι μελέτες αποκλειστούν. Επιπλέον η σύγκριση των αποτελεσμάτων του μοντέλου σταθερών επιδράσεων και του μοντέλου τυχαίων επιδράσεων είναι μια ανάλυση ευαισθησίας, αφού η εφαρμογή του μοντέλου τυχαίων επιδράσεων μπορεί να θεωρηθεί συγκριτικός έλεγχος στην περίπτωση που η υπόθεση ομοιογένειας αποτύχει.

Αν μετά τις αναλύσεις ευαισθησίας τα αποτελέσματα δεν αλλάζουν, τότε αυτό δυναμώνει την εμπιστοσύνη των συμπερασμάτων της μετανάλυσης. Αν τα αποτελέσματα αλλάζουν και οδηγούν σε αντιφατικά συμπεράσματα τότε αυτό θα πρέπει να μας προειδοποιήσει για να είμαστε πιο προσεκτικοί στην εξαγωγή και την ερμηνεία των αποτελεσμάτων της ανάλυσης. Επιπλέον αν προκύψουν διαφορές, οι ερευνητές μπορούν να κάνουν υποθέσεις για το ποιοι παράγοντες επηρεάζουν την αποτελεσματικότητα της θεραπείας.

Ενώ αυτές οι αναλύσεις είναι χρήσιμες για την εξέταση των επιπτώσεων που έχουν οι υποκειμενικές αποφάσεις που λαμβάνονται στην πορεία μιας μετανάλυσης, πολλές φορές είναι δύσκολο να διαμορφωθεί μια γενική ερμηνεία από τα αποτελέσματα πολλών αναλύσεων ευαισθησίας, ειδικά όταν το πλήθος των μελετών που συμπεριλαμβάνονται στη μετανάλυση είναι μικρό.

8.3 Μεροληψία (*Bias*)

Όπως αναφέρθηκε στο 2^ο κεφάλαιο ένα από τα σημαντικότερα κομμάτια της μετανάλυσης είναι ο προσδιορισμός όλων των κλινικών μελετών οι οποίες είναι σχετικές με το θέμα που εξετάζουμε. Αν οι μελέτες αναζητηθούν μόνο στην δημοσιευμένη βιβλιογραφία τότε οδηγούμαστε σε μεροληψία η οποία ονομάζεται “σφάλμα δημοσίευσης” (*publication bias*). Η μεροληψία αυτή οφείλεται στο γεγονός ότι η πιθανότητα δημοσίευσης μιας μελέτης εξαρτάται σε μεγάλο ποσοστό από το τελικό συμπέρασμά της. Είναι γνωστό ότι η υποβολή μιας μελέτης σε ένα ιατρικό επιστημονικό περιοδικό με σκοπό την δημοσίευση, αλλά και η

αποδοχή της μελέτης από το περιοδικό εξαρτώνται από το αν η επίδραση της θεραπείας είναι θετική ή αρνητική. Για παράδειγμα μελέτες οι οποίες καταλήγουν σε στατιστικά σημαντικά συμπεράσματα υπέρ της νέας θεραπείας, έχουν περισσότερες πιθανότητες να δημοσιευτούν απ' ότι μελέτες οι οποίες πιθανόν λόγω του μικρού δείγματος καταλήγουν σε στατιστικά μη σημαντικά συμπεράσματα ή σε μικρή επίδραση της θεραπείας. Αν επομένως η μετανάλυση βασιστεί μόνο σε μελέτες οι οποίες έχουν δημοσιευτεί τότε ο εκτιμητής της ολικής επίδρασης της θεραπείας είναι μεροληπτικός υπέρ της θετικής επίδρασης.

Ακόμα και αν μια μελέτη με αρνητικό συμπέρασμα δημοσιευτεί, η πιθανότητα εντοπισμού της εξαρτάται και από άλλους παράγοντες. Η γλώσσα στην οποία είναι γραμμένη μια μελέτη είναι σημαντικό εμπόδιο για την ανεύρεσή της (*language bias*). Επιπλέον οι μελέτες με αρνητικό συμπέρασμα δεν αναφέρονται τόσο συχνά στις παραπομπές όσο οι υπόλοιπες (*citation bias*).

Ένα σημείο στο οποίο θα πρέπει να είμαστε προσεκτικοί είναι ότι πολλές φορές συγχέονται και άλλοι παράγοντες με το σφάλμα δημοσίευσης. Όπως παρατήρησαν οι *Sterne et al* [2001a] οι μελέτες με την χαμηλότερη μεθοδολογική ποιότητα είναι συνήθως οι μελέτες με μικρό μέγεθος δείγματος, οι οποίες έχουν την τάση να δίνουν μεγαλύτερη επίδραση στην θεραπεία.

Επειδή ο προσδιορισμός όλων των σχετικών μελετών δεν είναι εύκολη υπόθεση, χρειαζόμαστε μεθόδους για να προσδιορίσουμε την μεροληψία βασιζόμενοι μόνο στα δεδομένα που είναι διαθέσιμα. Στην συνέχεια θα παρουσιάσουμε γραφικές και στατιστικές μεθόδους οι οποίες μπορούν να μας βοηθήσουν στην ανίχνευση και την διόρθωση του σφάλματος δημοσίευσης.

8.3.1 Γράφημα “χοάνη” (*Funnel plot*)

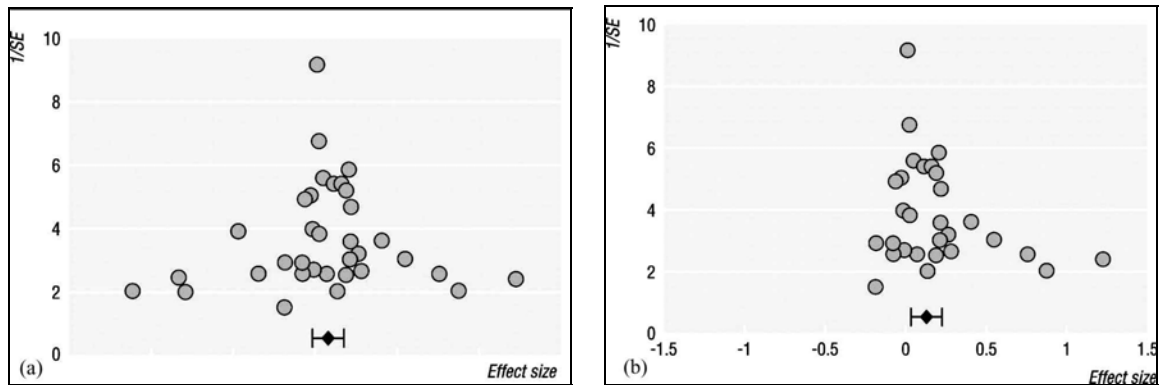
Οι *Light* και *Pillemer* το 1984 παρουσίασαν το “γράφημα χοάνη” (*funnel plot*), για την γραφική ανίχνευση του σφάλματος δημοσίευσης. Πρόκειται για ένα διμεταβλητό διάγραμμα διασποράς (x, y) , της εκτιμημένης τιμής της επίδρασης της θεραπείας (άξονας x) σε σχέση με το αντίστοιχο μέγεθος δείγματος της μελέτης (άξονας y). Το γράφημα βασίζεται στο γεγονός ότι η ακρίβεια στην εκτίμηση της επίδρασης της θεραπείας αυξάνει όσο το μέγεθος του δείγματος της μελέτης αυξάνει. Εναλλακτικά στον άξονα y μπορούμε να

αντιστοιχίσουμε την ακρίβεια (*precision*) της μελέτης η οποία είναι ίση με το βάρος $\sqrt{w_i} = s.e.(\hat{\theta}_i)^{-1}$. Οι *Sterne* και *Egger* [2001] εξέτασαν τις πιθανές επιλογές που έχουμε για τους άξονες x' και y' του γραφήματος “χοάνη” και παρουσίασαν τον πίνακα 8.1 δίνοντας κάποιες συστάσεις για την κάθε μια επιλογή.

| Possible choices of axis in funnel plots: advantages, disadvantages and recommendations | | |
|---|---|--|
| Axis / measure | Advantages and disadvantages | Recommendations |
| Vertical axis | | |
| Standard error | Funnel shape with straight 95% confidence lines. Emphasis of the plot is on smaller studies where bias is more likely. Axis has to be inverted to place the largest trials at the top of the graph. | The best choice in most cases |
| Precision (1/SE) | Plot is not funnel shaped; 95% confidence lines are curved. Emphasis of the plot is on larger studies; smaller studies are compressed at the bottom. | An option in studies which focus on a comparison of meta-analyses of small trials with subsequent large trials |
| Variance (SE ²) | Plot is not funnel shaped; 95% confidence lines are curved. Emphasis of the plot is on smaller studies where bias is more likely. | Not recommended |
| Inverse variance (1/variance) | Plot is not funnel shaped; 95% confidence lines are curved. Emphasis of plot is on larger studies, smaller studies are compressed at the bottom. | An option in studies which focus on a comparison of meta-analyses of small trials with subsequent large trials, but precision would usually be better. |
| Sample size or log sample size | Expected shape of plot in absence of bias is unpredictable. | Invalid choice |
| Horizontal axis | | |
| Log odds ratio | Scale is not constrained and plots have the same shape whether the outcome is defined as occurrence or non-occurrence of the disease. Odds ratios may be misinterpreted as risk ratios. | The best choice in most cases |
| Log risk ratio | Readily understood measure. Scale is naturally constrained so that heterogeneity may be introduced if the event rate is high. | Valid choice in many cases but not recommended if the event rate is high. Can give different conclusions depending on outcome definition. |
| Risk difference | Readily understood measure. Often associated with increased heterogeneity which may result in funnel plot asymmetry which is not apparent when ratio measures are used. | Not recommended in most cases |

Πίνακας 8.1: Επιλογή αξόνων στο γράφημα “χοάνη” (Πηγή: *Sterne, Egger* [2001])

Οι μικρές μελέτες δίνουν διασκορπισμένα σημεία στο κάτω μέρος του γραφήματος ενώ οι μεγαλύτερες μελέτες δίνουν σημεία με μικρότερη διασπορά γύρω από το σημείο στο οποίο αντιστοιχεί η ολική επίδραση της θεραπείας. Αν δεν υπάρχει μεροληψία το γράφημα έχει το σχήμα μιας ανεστραμμένης χοάνης, με άξονα συμμετρίας την νοητή κάθετη ευθεία η οποία αντιστοιχεί στην ολική επίδραση της θεραπείας (σχήμα 8.1a). Όταν υπάρχει μεροληψία το γράφημα είναι μη συμμετρικό (λοξό) αφού λείπουν τα σημεία που αντιστοιχούν σε μικρές μελέτες με στατιστικά μη σημαντικό συμπέρασμα (σχήμα 8.1 b). Όταν η επίδραση της θεραπείας υπολογίζεται με βάση το *relative risk* ή το *odds ratio* τότε στο γράφημα χρησιμοποιούμε τους λογαρίθμους αυτών των ποσοτήτων. Αυτό μας εξασφαλίζει ότι δυο επιδράσεις του ίδιου μέγεθος αλλά αντίθετης κατεύθυνσης (για παράδειγμα όταν ο σχετικός κίνδυνος είναι 0.5 και 2 αντίστοιχα) θα ισαπέχουν από την μονάδα [*Galbraith*, 1988].



Σχήμα 8.1 Υποθετικά γραφήματα “χοάνης”: Αριστερά: συμμετρικό γράφημα όταν δεν υπάρχει μεροληψία Δεξιά: μη συμμετρικό γράφημα όταν υπάρχει μεροληψία (Πηγή: Sutton, Duval, Tweedie [2000])

Όσο πιο έντονη είναι η ασυμμετρία του γραφήματος, τόσο πιθανότερο είναι ότι το ποσό της μεροληψίας που υπεισέρχεται στην μετανάλυση να είναι ουσιαστικό. Αν και η ασυμμετρία του γραφήματος έχει συσχετιστεί με το συστηματικό σφάλμα δημοσίευσης, το γράφημα αυτό θα πρέπει να θεωρηθεί ως ένα μέσο για τον έλεγχο της τάσης των μικρότερων μελετών να παρουσιάζουν μεγαλύτερη επίδραση της θεραπείας [Sterne, 2001a]. Η τάση αυτή όμως μπορεί να οφείλεται και σε άλλους λόγους εκτός του σφάλματος δημοσίευσης.

Στην συνέχεια θα αναφέρουμε τις πιθανές πηγές ασυμμετρίας του γραφήματος οι οποίες αναφέρονται στο άρθρο του Egger [1997]:

1. Μεροληψία επιλογής των μελετών (*selection bias*)
 - i. Σφάλμα δημοσίευσης (*publication bias*)
 - ii. Μεροληψία θέσης (*location bias*)
 - Μεροληψία γλώσσας (*language bias*)
 - Μεροληψία αναφορών (*citation bias*)
 - Πολλαπλό σφάλμα δημοσίευσης (*multiple publication bias*)
2. Χαμηλή ποιότητα των δεδομένων που προέρχονται από μικρές μελέτες
 - Φτωχός μεθοδολογικός σχεδιασμός
 - Ανεπαρκής ανάλυση
 - Ανακριβείς μετρήσεις-απάτη
3. Ετερογένεια μεταξύ των μελετών
4. Επιλογή του μέτρου επίδρασης της θεραπείας
5. Τυχαίοι παράγοντες

Επειδή η συμμετρία ή η ασυμμετρία του γραφήματος “χοάνη” καθορίζεται ανεπίσημα μετά από οπτική εξέταση, η ερμηνεία μπορεί να είναι διαφορετική μεταξύ των παρατηρητών. Αν επιπλέον στη μετανάλυση συμπεριλαμβάνονται λίγες μόνο μελέτες τότε δεν είναι εύκολο να διαπιστώσουμε την ασυμμετρία του γραφήματος. Στις επόμενες παραγράφους θα παρουσιαστούν πιο επίσημες στατιστικές μεθοδολογίες οι οποίες εξετάζουν τη σχέση μεταξύ του αποτελέσματος μιας μελέτης με το αντίστοιχο μέγεθος δείγματος.

Οι ερευνητές παρ’ όλα αυτά θα πρέπει να κατασκευάζουν και να εξετάζουν το γράφημα “χοάνη” κάθε φορά που εκτελούν μετα-ανάλυση. Αν υπάρχει πρόβλημα ασυμμετρίας τότε σίγουρα θα πρέπει να ερευνηθούν οι λόγοι που την προκαλούν. Αυτό που θα πρέπει όμως να κατανοήσουμε είναι ότι αν και το γράφημα “χοάνη” μπορεί να προειδοποιήσει τους ερευνητές όταν υπάρχει πρόβλημα μεροληψίας, δεν παρέχει λύση.

8.3.2 Μη παραμετρική μέθοδος των Begg και Mazumdar

Οι *Begg* και *Mazumdar* [1994] πρότειναν ένα μη παραμετρικό έλεγχο συσχέτισης (*rank correlation method*) ο οποίος βασίζεται στο στατιστικό *Kendall’s tau*, για την ανίχνευση του σφάλματος δημοσίευσης σε μια μετανάλυση.

Έστω $\hat{\theta}_i$ και $v_i = Var(\hat{\theta}_i)$ η επίδραση και η αντίστοιχη διακύμανση στην i μελέτη. Το τεστ ελέγχει την συσχέτιση της τυποποιημένης επίδρασης $\hat{\theta}_i^*$ σε σχέση με την διακύμανση

v_i , όπου $\hat{\theta}_i^* = \frac{\hat{\theta}_i - \hat{\theta}}{\sqrt{v_i^*}}$, $\hat{\theta} = \frac{\sum_i \hat{\theta}_i / v_i}{\sum_i 1 / v_i}$ και $v_i^* = v_i - \left(\sum_j 1 / v_j \right)^{-1}$. Εναλλακτικά ο έλεγχος

μπορεί να βασιστεί στη συσχέτιση του $\hat{\theta}_i^*$ με το αντίστοιχο βάρος της μελέτης n_i .

Στην συνέχεια σε κάθε μελέτη δίνονται *ranks* για κάθε $\hat{\theta}_i^*$ και v_i (ή n_i) και κατασκευάζονται τα ζευγάρια $(\hat{\theta}_i^*, v_i)$. Αν P είναι ο αριθμός των ζευγαριών που διατάσσονται με τον ίδιο τρόπο στους παράγοντες $\hat{\theta}_i^*$ και v_i , και Q είναι ο αριθμός των ζευγαριών που η διάταξη αντιστρέφεται, τότε έχουμε το κανονικοποιημένο στατιστικό:

$$z = \frac{P-Q}{[k(k-1)(2k+5)/18]^{1/2}}$$

στην περίπτωση που δεν υπάρχουν δεσμοί (*ties*).

Αν υπάρχει μεροληψία, τότε οι μικρότερες μελέτες θα δίνουν μεγαλύτερες εκτιμήσεις για την επίδραση της θεραπείας. Αυτό σημαίνει ότι θα υπάρχει θετική συσχέτιση μεταξύ της τυποποιημένης επίδρασης $\hat{\theta}_i^*$ και της διακύμανσης v_i , αφού η διακυμάνσεις των εκτιμητών που προέρχονται από τις μικρές μελέτες έχουν μεγάλες τιμές.

Η μέθοδος αυτή έχει χαμηλή ισχύ όταν τα δεδομένα είναι συνεχή (κανονικά κατανομημένα) και ειδικά όταν το πλήθος των μελετών είναι μικρό. Για περισσότερες πληροφορίες ο αναγνώστης παραπέμπεται στο άρθρο των *Begg και Mazumdar* [1994].

8.3.3 Μέθοδος παλινδρόμησης του Egger

Ο *Egger* [1997] χρησιμοποίησε την μέθοδο της απλής γραμμικής παλινδρόμησης για να μετρήσει την ασυμμετρία του γραφήματος “χοάνη”. Η εξαρτημένη μεταβλητή είναι ο τυποποιημένος εκτιμητής της επίδρασης της θεραπείας $y_i = \hat{\theta}_i / s.e.(\hat{\theta}_i)$ και ανεξάρτητη μεταβλητή είναι η αντίστοιχη ακρίβεια η οποία δίνεται από τον τύπο $x_i = w_i = 1 / s.e.(\hat{\theta}_i)$.

Το μοντέλο της παλινδρόμησης δίνεται από τη σχέση:

$$y_i = a + \beta x_i + e_i, \quad i = 1, 2, \dots, k$$

Τα σφάλματα e_i είναι κανονικά κατανομημένα με $e_i \sim N(0,1)$. Το μοντέλο παλινδρόμησης για το γράφημα “χοάνη” είναι παρόμοιο με το μοντέλο παλινδρόμησης που παρουσιάσαμε στο ακτινικό διάγραμμα (*radial plot*), μόνο που σε αυτή την περίπτωση η ευθεία δεν είναι απαραίτητο να διέρχεται από την αρχή των αξόνων.

Όταν το μέγεθος της μελέτης είναι μικρό τότε το τυπικό σφάλμα του εκτιμητή θα είναι μεγάλο. Επομένως τόσο η ακρίβεια όσο και ο τυποποιημένος εκτιμητής θα παίρνουν τιμές

και στους δυο άξονες κοντά στο μηδέν. Αυτό σημαίνει ότι οι μικρές μελέτες θα δίνουν σημεία κοντά στην αρχή των αξόνων. Στην αντίθετη περίπτωση, οι μεγάλες μελέτες δίνουν πιο ακριβείς εκτιμητές και επομένως αν η θεραπεία είναι ευεργετική οι τιμές των αντίστοιχων τυποποιημένων εκτιμητών θα είναι μεγάλες. Αν δεν υπάρχει μεροληψία, τα σημεία θα είναι διασκορπισμένα πάνω σε μια ευθεία γραμμή η οποία θα διέρχεται από την αρχή των αξόνων, δηλαδή θα είναι $a=0$, με κλίση β . Η κλίση της ευθείας είναι ίση με την ολική επίδραση της θεραπείας $\hat{\theta}$.

Αν όμως υπάρχει μεροληψία τότε η γραμμή παλινδρόμησης δεν θα διέρχεται από την αρχή των αξόνων. Συνεπώς η σταθερά a δίνει ένα μέτρο για την ασυμμετρία του γραφήματος: όσο πιο μεγάλη είναι η τιμή της τόσο πιο μεγάλη είναι η ασυμμετρία του γραφήματος.

Οι εκτιμητές ελαχίστων τετραγώνων των a και β δίνονται από τις σχέσεις:

$$\hat{a} = \frac{1}{k} \sum_{i=1}^k (\hat{\theta}_i - \hat{\theta}) \sqrt{w_i} \quad \text{και} \quad \hat{\beta} = \frac{k \sum_{i=1}^k \hat{\theta}_i w_i - \left(\sum_{i=1}^k \sqrt{w_i} \right) \left(\sum_{i=1}^k \hat{\theta}_i \sqrt{w_i} \right)}{k \sum_{i=1}^k w_i - \left(\sum_{i=1}^k \sqrt{w_i} \right)^2}.$$

Κάτω από το μοντέλο των σταθερών επιδράσεων, η διακύμανση του εκτιμητή \hat{a} δίνεται από τη σχέση:

$$\text{Var}(\hat{a}) = \frac{\sum_{i=1}^k w_i}{k \sum_{i=1}^k w_i - \left(\sum_{i=1}^k \sqrt{w_i} \right)^2}.$$

Στα στατιστικά πακέτα γίνεται η υπόθεση ότι $e_i \sim N(0, \sigma^2)$, με το σ^2 να εκτιμάται από τα δεδομένα. Επομένως για να υπολογίσουμε την διακύμανση του εκτιμητή \hat{a} θα πρέπει να διαιρέσουμε την τιμή που δίνει το πακέτο με το μέσο τετραγωνικό σφάλμα (MSE).

Ο έλεγχος της μηδενικής υπόθεσης $H_0 : a = 0$ γίνεται με το στατιστικό $z = \frac{\hat{a}}{\sqrt{\text{Var}(\hat{a})}}$, το

οποίο κάτω από την μηδενική υπόθεση ακολουθεί τυπική κανονική κατανομή.

Όταν στην μετανάλυση συμπεριλαμβάνονται πολλές μικρές μελέτες μαζί με μια μελέτη μεγαλύτερης έκτασης, τότε θα πρέπει να χρησιμοποιήσουμε σταθμισμένη παλινδρόμηση για να υπολογίσουμε τις παραμέτρους α και β . Έτσι δίνοντας ως βάρους τις αντίστροφες τιμές της διακύμανσης των εκτιμητών, $1/Var(\hat{\theta}_i)$ η ανάλυσή μας αποκτά μεγαλύτερη ισχύ.

Στο άρθρο των *Macaskill* και *Walter* [2001] αναφέρεται ότι η παλινδρόμηση του *Egger* είναι μεροληπτική μέθοδος. Αυτό συμβαίνει γιατί αφού το τυπικό σφάλμα $s.e.(\hat{\theta}_i)$ εκτιμάται από τα δεδομένα, η ανεξάρτητη μεταβλητή εμπεριέχει και το σφάλμα μέτρησης.

8.3.4 Εναλλακτική μέθοδος παλινδρόμησης

Εναλλακτικά με τη μέθοδο του *Egger*, στο άρθρο των *Macaskill* και *Walter* [2001] προτείνεται η σταθμισμένη παλινδρόμηση της επίδρασης της θεραπείας $\hat{\theta}_i$ (εξαρτημένη μεταβλητή) σε σχέση με το μέγεθος της μελέτης n_i (ανεξάρτητη μεταβλητή), με αντίστοιχα βάρη τα $w_i = 1/Var(\hat{\theta}_i)$. Όταν δεν υπάρχει σφάλμα δημοσίευσης η κλίση της γραμμής παλινδρόμησης αναμένεται να είναι μηδέν. Επειδή τα βάρη εκτιμώνται από τα δεδομένα, ο εκτιμητής της κλίσης της ευθείας δεν είναι αμερόληπτος.

Για περισσότερες πληροφορίες ο αναγνώστης παραπέμπεται στο αντίστοιχο άρθρο.

8.3.5 Μέθοδος “Trim and Fill”

Οι *Taylor* και *Tweedie* [1998] παρουσίασαν μια μέθοδο διόρθωσης της ασυμμετρίας του γραφήματος “χοάνη”. Αρχικά εκτιμάται το πλήθος των μελετών οι οποίες λείπουν από το αριστερό κομμάτι του γραφήματος (σχήμα 8.1b) και που η παρουσία τους θα έκανε το γράφημα συμμετρικό. Στη συνέχεια παραλείπονται τόσες μελέτες από το δεξιό μη συμμετρικό μέρος του γραφήματος, ώστε το γράφημα να γίνει συμμετρικό. Στο επόμενο

στάδιο από το συμμετρικό γράφημα που απομένει υπολογίζουμε την ολική επίδραση της θεραπείας, το σημείο δηλαδή που αντιστοιχεί στον άξονα συμμετρίας του γραφήματος. Στην συνέχεια επανατοποθετούμε τις μελέτες που είχαμε αφαιρέσει μαζί με τις αντίστοιχες συμμετρικές μελέτες ως προς την ολική επίδραση της θεραπείας, την οποία υπολογίσαμε στο προηγούμενο βήμα. Στο τελευταίο στάδιο υπολογίζουμε ξανά την ολική επίδραση της θεραπείας από το συμπληρωμένο γράφημα.

Το μειονέκτημα της μεθόδου όπως διαπιστώθηκε με μεθόδους προσομοίωσης είναι ότι προσδιορίζει σαν ελλείπουσες κάποιες μελέτες οι οποίες στην πραγματικότητα δεν υπάρχουν [Sterne και Egger, 2001]. Κατά συνέπεια η εφαρμογή της μεθόδου σε κάποιες μεταanalύσεις θα μπορούσε να εντοπίσει και ανύπαρκτες μελέτες, προκειμένου να διορθωθεί ακόμα και μια τυχαία ασυμμετρία του γραφήματος “χοάνη”.

Στην επόμενη παράγραφο θα παρουσιάσουμε τον επαναληπτικό αλγόριθμο της μεθόδου “Trim and Fill”.

8.3.5.1 Επαναληπτική μέθοδος “Trim and Fill”

Έστω ότι η μεταanάλυση περιλαμβάνει n μελέτες οι οποίες αναφέρονται στο ίδιο πρόβλημα. Ας υποθέσουμε ότι k_0 είναι ο αριθμός των σχετικών μελετών οι οποίες δεν έχουν εντοπιστεί λόγω του σφάλματος δημοσίευσης. Στο πρώτο στάδιο θα εκτιμήσουμε την τιμή του k_0 .

Έστω ότι η εκτιμημένη τιμή της ολικής επίδρασης της θεραπείας $\hat{\theta}$ είναι γνωστή. Αρχικά ταξινομούμε τις μελέτες σε αύξουσα σειρά με βάση την επίδραση της θεραπείας. Για κάθε μελέτη υπολογίζουμε την διαφορά $X_i = \hat{\theta}_i - \hat{\theta}$ και στην συνέχεια δίνουμε ranks r_i στις απόλυτες τιμές των X_i . Τα ranks επομένως παίρνουν τιμές από 1 έως n .

Έστω $\gamma^* \geq 0$ το μήκος της τελευταίας ακολουθίας στην οποία όλα τα X_i έχουν θετική τιμή και T_n το άθροισμα των ranks τα οποία αντιστοιχούν σε θετικά X_i (δηλαδή $T_n = \sum_{X_i > 0} r_i$).

Βασιζόμενοι σε αυτές τις ποσότητες προκύπτουν τρεις εκτιμητές του k_0 :

$$R_0 = \gamma^* - 1$$

$$L_0 = [4T_n - n(n+1)] / [2n - 1]$$

$$Q_0 = n - 1/2 - \sqrt{2n^2 - 4T_n + 1/4}$$

Επειδή το k_0 αντιπροσωπεύει το πλήθος των μελετών που λείπουν, οι τιμές που δίνουν οι προηγούμενοι εκτιμητές στρογγυλοποιούνται στο πλησιέστερο ακέραιο. Με προσομοιώσεις που έγιναν, οι εκτιμητές L_0 και R_0 έχουν μικρό ποσό μεροληψίας και συμπεριφέρονται καλύτερα από τον εκτιμητή Q_0 . Όσο αυξάνει το n τόσο ο εκτιμητής R_0 είναι προτιμότερος από τον εκτιμητή L_0 . Στην πράξη θα πρέπει να υπολογιστούν και οι δύο εκτιμητές (L_0 και R_0) πριν καταλήξουμε σε κάποιο συμπέρασμα για το k_0 . Αν και οι δύο συμφωνούν τότε το συμπέρασμα είναι προφανές, ενώ αν διαφωνούν τότε μπορούμε με αυτές τις τιμές να εκτελέσουμε μια ανάλυση ευαισθησίας. Για περισσότερες πληροφορίες σχετικά με τους εκτιμητές του k_0 και τις ιδιότητες τους ο αναγνώστης παραπέμπεται στο άρθρο των *Duval, Tweedie* [2000].

Όπως είδαμε οι εκτιμητές του k_0 εξαρτώνται από την ολική επίδραση της θεραπείας, $\hat{\theta}$, γιατί θα πρέπει να γνωρίζουμε αν μια μελέτη-σημείο του γραφήματος “χοάνη” βρίσκεται δεξιά ή αριστερά του $\hat{\theta}$. Στη συνέχεια θα παρουσιάσουμε τα βασικά σημεία της επαναληπτικής μεθόδου “*Trim and Fill*”.

Αρχικά εκτιμάμε τη τιμή της ολικής επίδρασης της θεραπείας θ χρησιμοποιώντας ανάλογα το μοντέλο σταθερών ή τυχαίων επιδράσεων. Στη συνέχεια βρίσκουμε μια πρώτη εκτίμηση της τιμής του k_0 και αφαιρούμε τις k_0 πιο ακραίες θετικές μελέτες. Στο επόμενο βήμα υπολογίζουμε ξανά την ολική επίδραση της θεραπείας και επαναλαμβάνουμε την ίδια διαδικασία για τον νέο υπολογισμό του k_0 . Η διαδικασία επαναλαμβάνεται μέχρις ότου δεν υπάρχει αλλαγή στην εκτίμηση του k_0 . Πειραματικά η μέθοδος σταματά μετά από 2-3 επαναλήψεις.

Στο σημείο αυτό έχουμε υπολογίσει τους τελικούς εκτιμητές $\hat{\theta}^*$ και k_0^* . Στην συνέχεια υπολογίζουμε τις επιδράσεις της θεραπείας $\hat{\theta}_j^*$ και το αντίστοιχο τυπικό σφάλμα s_j^* για κάθε νέα μελέτη η οποία θα προστεθεί στην μετανάλυση για να προκύψει συμμετρικό το γράφημα “χοάνη”, με βάση τις σχέσεις:

$$\hat{\theta}_j^* = 2\hat{\theta}^* - \hat{\theta}_{n-j+1}, \quad s_j^* = s_{n-j+1} \quad \text{όπου } j = 1, 2, \dots, k_0$$

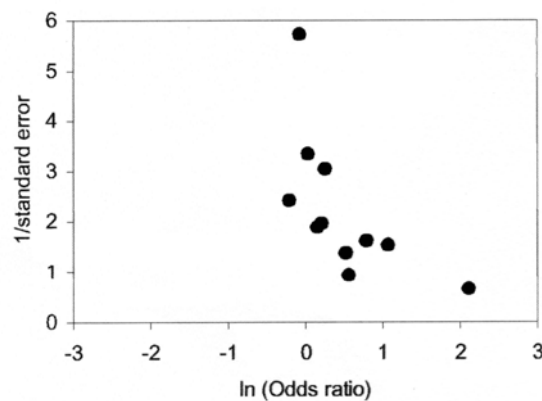
8.3.5.2 Εφαρμογή της μεθόδου “Trim and Fill”

Για την καλύτερη κατανόηση της μεθοδολογίας “Trim and Fill” θα εργαστούμε με ένα παράδειγμα το οποίο παρουσιάζεται από τους Taylor και Tweedie [1998]. Τα δεδομένα προέρχονται από την βάση δεδομένων Cochrane. Πρόκειται για 11 μελέτες στις οποίες μελετήθηκε η επίδραση των gangliosides (είδος λιπιδίων) στην πρόληψη του μοιραίου συμβάντος μετά από οξύ ισχαιμικό έμφραγμα. Στον πίνακα 8.2 δίνονται τα δεδομένα των 11 μελετών (log-odds ratio εκτιμητές) με τα αντίστοιχα τυπικά σφάλματα.

| Μελέτη | Log-odds ratio | s.e | Μελέτη | Log-odds ratio | s.e |
|--------|----------------|------|--------|----------------|------|
| 1. | -0.20 | 0.41 | 7. | 0.53 | 0.74 |
| 2. | -0.07 | 0.18 | 8. | 0.56 | 1.08 |
| 3. | 0.04 | 0.30 | 9. | 0.80 | 0.62 |
| 4. | 0.16 | 0.53 | 10. | 1.08 | 0.66 |
| 5. | 0.21 | 0.51 | 11 | 2.11 | 1.55 |
| 6. | 0.27 | 0.33 | | | |

Πίνακας 8.2: Effect of gagliosides in acute ischemic stroke

Το γράφημα “χοάνη” για τις 11 μελέτες (Σχήμα 8.2) παρουσιάζει έντονη ασυμμετρία, γεγονός που μας προειδοποιεί ότι υπάρχει πρόβλημα μεροληψίας.



Σχήμα 8.2: Funnel plot of the effect of gagliosides in acute ischemic stroke

Με βάση το μοντέλο τυχαίων επιδράσεων υπολογίζουμε ότι η ολική επίδραση της θεραπείας είναι ίση με $\hat{\theta} = \Delta_1 = 0.105$. Στον πίνακα 8.3 παρουσιάζουμε τα αποτελέσματα της επαναληπτικής μεθόδου “Trim and Fill”.

Στην τρίτη στήλη του πίνακα υπολογίζουμε τις διαφορές $X_i = \hat{\theta}_i - \Delta_1$ ενώ στην τέταρτη στήλη δίνουμε *ranks* στις τιμών των $|X_i|$. Όπως είναι φανερό είναι $\gamma^* = 5$ και $T_n = 1 + 3 + 4 + 7 + 8 + 9 + 10 + 11 = 53$. Επομένως $L_0 = 3.81$ και $R_0 = 4$. Και από τους δυο εκτιμητές βρίσκουμε ότι $\hat{k}_0 = 4$. Παραλείποντας τις 4 τελευταίες μελέτες υπολογίζουμε ξανά την ολική επίδραση της θεραπείας και βρίσκουμε ότι $\hat{\theta} = \Delta_2 = 0.028$. Υπολογίζουμε ξανά τις διαφορές $X_i = \hat{\theta}_i - \Delta_2$ και βρίσκουμε τα νέα *ranks* των $|X_i|$ (6^η στήλη). Με βάση τα νέα *ranks* βρίσκουμε ότι $\gamma^* = 6$ και $T_n = 59$, επομένως θα είναι $L_0 = 4.95$ και $R_0 = 5$. Και από τους δυο εκτιμητές βρίσκουμε ότι $\hat{k}_0 = 5$. Στο επόμενο βήμα επομένως θα διαγράψουμε τις 5 τελευταίες μελέτες και η ολική επίδραση της θεραπείας θα είναι $\hat{\theta} = \Delta_3 = 0.014$. Υπολογίζοντας τα νέα *ranks* των $|X_i| = \left| \hat{\theta}_i - \Delta_3 \right|$ βλέπουμε ότι δεν υπάρχει διαφορά μεταξύ των στηλών 6 και 7. Άρα η επαναληπτική μέθοδος υπολογισμού του k_0 θα σταματήσει για

$\hat{k}_0 = 5$. Στο επόμενο βήμα θα υπολογιστούν οι επιδράσεις και τα τυπικά σφάλματα των 5 νέων μελετών που θα προστεθούν στο γράφημα με βάση τους τύπους της προηγούμενης παραγράφου.

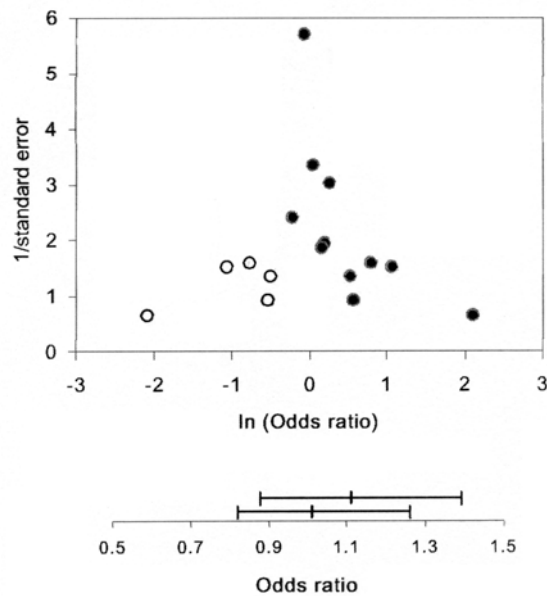
Για παράδειγμα, για $j = 1$ θα είναι:

$$\hat{\theta}_1^* = 2\hat{\theta}^* - \hat{\theta}_{11} = 2 \times 0.014 - 2.11 = -2.082 \quad \text{και} \quad s_1^* = s_{11} = 1.55$$

| Αρχικά δεδομένα | | | | Διαγράφουμε τις 4 πιο ακραίες μελέτες | | Διαγράφουμε τις 5 πιο ακραίες μελέτες | | Συμπληρωμένα δεδομένα | |
|-----------------|------|-----------------------|--------------|---------------------------------------|--------------|---------------------------------------|--------------|-----------------------|------|
| InO R | SE | InOR - Δ ₁ | Signed ranks | InOR - Δ ₂ | Signed ranks | InOR-Δ ₃ | Signed ranks | InOR | SE |
| -0.20 | 0.41 | -0.31 | -6 | -0.23 | -5 | -0.22 | -5 | -0.20 | 0.41 |
| -0.07 | 0.18 | -0.17 | -5 | -0.10 | -2 | -0.08 | -2 | -0.07 | 0.18 |
| 0.04 | 0.30 | -0.06 | -2 | 0.01 | 1 | 0.03 | 1 | 0.04 | 0.30 |
| 0.16 | 0.53 | 0.06 | 1 | 0.14 | 3 | 0.15 | 3 | 0.16 | 0.53 |
| 0.21 | 0.51 | 0.11 | 3 | 0.18 | 4 | 0.20 | 4 | 0.21 | 0.51 |
| 0.27 | 0.33 | 0.16 | 4 | 0.24 | 6 | 0.25 | 6 | 0.27 | 0.33 |
| 0.53 | 0.74 | 0.42 | 7 | 0.50 | 7 | 0.51 | 7 | 0.53 | 0.74 |
| 0.56 | 1.08 | 0.46 | 8 | 0.54 | 8 | 0.55 | 8 | 0.56 | 1.08 |
| 0.80 | 0.62 | 0.69 | 9 | 0.77 | 9 | 0.78 | 9 | 0.80 | 0.62 |
| 1.08 | 0.66 | 0.97 | 10 | 1.05 | 10 | 1.06 | 10 | 1.08 | 0.66 |
| 2.11 | 1.55 | 2.01 | 11 | 2.09 | 11 | 2.10 | 11 | 2.11 | 1.55 |
| | | | | | | | | -0.50 | 0.74 |
| | | | | | | | | -0.54 | 1.08 |
| | | | | | | | | -0.77 | 0.62 |
| | | | | | | | | -1.05 | 0.66 |
| | | | | | | | | -2.09 | 1.55 |

Πίνακας 8.3: Συμπλήρωση των δεδομένων με τη μέθοδο “Trim and Fill”

Στο επόμενο σχήμα παρουσιάζουμε το συμμετρικό πλέον γράφημα “χοάνη” μετά την συμπλήρωση των 5 μελετών.



Σχήμα 8.3: Filled funnel plot of the effect of gagliosides in acute ischemic stroke

Μετα την διόρθωση του συστηματικού σφάλματος δημοσίευσης με την μέθοδο “*Trim and Fill*” βρίσκουμε ότι ο ολικός εκτιμητής της επίδρασης της θεραπείας είναι ίσος με $\hat{\theta} = 1.01$ με αντίστοιχο 95% διάστημα εμπιστοσύνης ίσο με $(0.82, 1.26)$. Ο ολικός εκτιμητής της επίδρασης της θεραπείας στα αρχικά δεδομένα ήταν ίσος με $\hat{\theta} = 1.11$ ενώ το αντίστοιχο 95% διάστημα εμπιστοσύνης ήταν ίσο με $(0.88, 1.39)$. Είναι φανερό επομένως ότι στα αρχικά δεδομένα η εκτίμηση της επίδρασης της θεραπείας είναι σαφώς μεγαλύτερη λόγω της ύπαρξης του σφάλματος δημοσίευσης.

8.3.6 Rosenthal’s file-drawer method

Έστω ότι η μετανάλυση συμπεριλαμβάνει k μελέτες με τις οποίες θέλουμε να ελέγξουμε την μηδενική υπόθεση $H_0 : \theta = 0$, ότι δηλαδή η επίδραση της θεραπείας είναι μηδέν, έναντι της μονόπλευρης εναλλακτικής ότι η επίδραση της θεραπείας είναι θετική $H_a : \theta > 0$. Ας υποθέσουμε ακόμα ότι με βάση τις k μελέτες καταλήξαμε στο συμπέρασμα ότι η επίδραση

της θεραπείας είναι θετική. Ο *Rosenthal* το 1979 παρουσίασε μια απλή μέθοδο με την οποία μπορούμε να υπολογίσουμε το πλήθος των μελετών, που έχουν κατά μέσο όρο μηδενική επίδραση της θεραπείας, οι οποίες αν συμπεριληφθούν στη μετανάλυση θα οδηγήσουν στην αποδοχή της μηδενικής υπόθεσης ($H_0 : \theta = 0$). Οι μελέτες αυτές δεν συμπεριλαμβάνονται στην μετανάλυση επειδή δεν έχουν δημοσιευτεί και κατά κάποιο τρόπο φυλάσσονται σε ένα αρχείο ή σε κάποιο συρτάρι (γι' αυτό και η μέθοδος ονομάζεται “*file-drawer method*”).

Ας υποθέσουμε ότι για την ανάλυση των k μελετών χρησιμοποιούμε την μέθοδο σταθερών επιδράσεων. Ανάλογη μεθοδολογία μπορεί να εφαρμοστεί και στην περίπτωση του μοντέλου τυχαίων επιδράσεων. Όπως αναφέρθηκε στο 4^ο κεφάλαιο αν $\hat{\theta}_i$ είναι ο εκτιμητής της ολικής επίδρασης της θεραπείας θ ο οποίος προέρχεται από την i μελέτη τότε θα ισχύει:

$$\hat{\theta}_i \sim N(\theta, w_i^{-1}), \quad \text{όπου } w_i = 1/\text{Var}(\hat{\theta}_i)$$

Κάτω από την μηδενική υπόθεση ότι η επίδραση της θεραπείας είναι μηδέν η προηγούμενη σχέση δίνει $\hat{\theta}_i \sim N(0, w_i^{-1})$ και επομένως $\hat{\theta}_i \sqrt{w_i} \sim N(0, 1)$.

Συνδυάζοντας τις k ανεξάρτητες μελέτες, το στατιστικό $Z_k = \frac{\sum_{i=1}^k \hat{\theta}_i \sqrt{w_i}}{\sqrt{k}}$ θα ακολουθεί τυπική κανονική κατανομή. Επομένως αν το επίπεδο σημαντικότητας είναι α , η μηδενική υπόθεση $H_0 : \theta = 0$ απορρίπτεται έναντι της μονόπλευρης εναλλακτικής $H_a : \theta > 0$ αν $Z_k > z_\alpha$.

Αν r είναι το πλήθος των μη δημοσιευμένων μελετών στις οποίες η επίδραση της θεραπείας είναι κατά μέσο όρο μηδέν, τότε για να μην μπορούμε να απορρίψουμε την μηδενική υπόθεση θα πρέπει να ισχύει:

$$Z_{k,r} = \frac{\sum_{i=1}^k \hat{\theta}_i \sqrt{w_i}}{\sqrt{k+r}} < z_\alpha$$

Επομένως για το πλήθος των μελετών r θα ισχύει:

$$r > -k + \frac{\left\{ \sum_{i=1}^k \hat{\theta}_i \sqrt{w_i} \right\}^2}{z_\alpha^2}$$

Αν το πλήθος r των μελετών είναι σχετικά μεγάλο τότε μπορούμε να υποστηρίξουμε ότι δεν είναι πιθανό να υπάρχουν τόσες μελέτες αδημοσίευτες και επομένως μπορούμε να έχουμε εμπιστοσύνη στα αποτελέσματά μας.

Αν και η μέθοδος είναι απλή, η υπόθεση στην οποία στηρίζεται δεν είναι ρεαλιστική, αφού υποθέτει ότι η επίδραση της θεραπείας στις αδημοσίευτες μελέτες είναι κατά μέσο όρο μηδέν. Άλλο ένα μειονέκτημα της μεθόδου είναι ότι αγνοεί το μέγεθος των μη δημοσιευμένων μελετών και επιπλέον δεν επηρεάζεται από τη διαφορά των εκτιμητών η οποία προκύπτει μεταξύ των μικρών και των μεγάλων μελετών [Whitehead, 2002].

8.3.7 Weighted distributions approach

Στην συνέχεια θα αναφέρουμε μεθοδολογίες με τις οποίες μοντελοποιούμε το συστηματικό σφάλμα δημοσίευσης, χρησιμοποιώντας σταθμισμένες κατανομές (*weighted distributions*). Επειδή οι μεθοδολογίες αυτές είναι αρκετά πολύπλοκες, θα τις παρουσιάσουμε περιληπτικά, δίνοντας ωστόσο τις απαραίτητες αναφορές. Η βασική υπόθεση στην οποία στηρίζονται αυτές οι μέθοδοι είναι ότι η πιθανότητα να συμπεριληφθεί μια μελέτη στη μετανάλυση εξαρτάται από το p -value του στατιστικού τεστ, για τον έλεγχο της υπόθεσης $H_0 : \theta_i = 0$.

Η πιο απλή προσέγγιση δόθηκε από τον Hedges' [1984], ο οποίος ασχολήθηκε με συνεχή δεδομένα εκφράζοντας την επίδραση της θεραπείας με βάση την τυποποιημένη διαφορά των μέσων τιμών. Η κατανομή κατασκευάστηκε απλά δίνοντας την τιμή "1" στις στατιστικά σημαντικές μελέτες και την τιμή "0" στις υπόλοιπες. Η μέθοδος αυτή υποθέτει ότι όλες οι στατιστικά σημαντικές μελέτες δημοσιεύονται, ενώ όλες οι μη σημαντικές όχι. Αυτό όμως δεν περιγράφει την πραγματικότητα αφού κάποιες μελέτες με τιμή p -value λίγο μικρότερη του 0.05 μπορεί να μην δημοσιευτούν. Σε αυτή την περίπτωση η μέθοδος υποεκτιμά το σφάλμα μεροληψίας.

Οι Iyengar και Greenhouse [1988] παρουσίασαν δυο διαφορετικές προσεγγίσεις στη χρήση των σταθμισμένων κατανομών. Και στις δυο θεώρησαν ότι όλες οι μελέτες που είναι στατιστικά σημαντικές δημοσιεύονται και επομένως η συνάρτηση σε αυτή τη περίπτωση θα παίρνει την τιμή "1". Για τις μη σημαντικές μελέτες στην πρώτη προσέγγιση έδιναν μια

σταθερή τιμή (διαφορετική του 0), ενώ στη δεύτερη προσέγγιση θεώρησαν ότι η πιθανότητα έκδοσης αυξάνει εκθετικά όσο η απόκριση πλησιάζει την στατιστική σημαντικότητα.

Αργότερα ο *Hedges* [1992] παρουσίασε μια πιο γενικευμένη μορφή της σταθμισμένης συνάρτησης. Το μοντέλο που πρότεινε επέτρεπε στην συνάρτηση να παίρνει διαφορετικές τιμές σε διαφορετικές περιοχές της κλίμακας του *p-value*, μετατρέποντας την σε κλιμακωτή συνάρτηση. Έτσι επιτρέπεται ο συλλογισμός: μια μελέτη με *p-value* ίσο με 0.01 είναι πιθανότερο να εκδοθεί σε σχέση με μια μελέτη με *p-value* 0.05, ενώ η τελευταία είναι πιθανότερη να εκδοθεί από μια μελέτη με *p-value* ίσο με 0.10 κ.λ.π. Το μοντέλο μπορεί να χρησιμοποιηθεί ακόμα και αν οι επιδράσεις της θεραπείας θεωρηθούν τυχαίες.

Οι *Dear* και *Begg* [1992] παρατήρησαν ότι η προηγούμενη συνάρτηση έχει πρόβλημα με την μονοτονία της. Έτσι παρουσίασαν μια προσέγγιση η οποία επιτρέπει την αλλαγή του σχήματος της σταθμισμένης συνάρτησης, χρησιμοποιώντας ένα ημιπαραμετρικό μοντέλο. Και αυτό το μοντέλο, όπως το μοντέλο του *Hedges*, επιτρέπει την χρήση τυχαίων επιδράσεων της θεραπείας. Η μέθοδος των *Dear* και *Begg* είναι συμπληρωματική του γραφήματος “χοάνη”, αλλά είναι αρκετά ευαίσθητη στο να εντοπίσει το σφάλμα δημοσίευσης ακόμα και όταν τα μεγέθη των μελετών είναι παρόμοια.

ΚΕΦΑΛΑΙΟ 9

Λογισμικό Μετανάλυσης (*Software*)

9.1 Εισαγωγή

Η αυξανόμενη χρήση των τεχνικών της μετανάλυσης είχε ως άμεση συνέπεια τη δημιουργία αντίστοιχου λογισμικού για υπολογιστές, τα οποία είτε είναι αφιερωμένα αποκλειστικά στην μετανάλυση είτε αποτελούν ρουτίνες (*macros*) των γενικών στατιστικών προγραμμάτων. Στο κεφάλαιο αυτό θα παρουσιάσουμε εμπορικά (*commercial*) προγράμματα μετανάλυσης, προγράμματα που διατίθενται δωρεάν (*free-ware*) καθώς και ρουτίνες μετανάλυσης τις οποίες παρέχουν μεγάλα στατιστικά πακέτα. Τέλος θα αναφερθεί σχετική βιβλιογραφία στην οποία γίνεται σύγκριση των στατιστικών πακέτων που εκτελούν μετανάλυση.

9.2 Εμπορικά προγράμματα (*Commercial software*)

- **Comprehensive Meta-Analysis** (*Operating system: windows*)

Input data: 2 x 2 tables, means, p-values, correlation coefficients

Statistical models: Fixed effects (Mantel Haenszel, Inverse variance, Peto), Random effects (Mantel Haenszel, Inverse Variance), Heterogeneity test, Subgroup Analysis

Output effect measures: odds ratio, relative risk, rate difference, standardized mean difference and correlation

Plots: Funnel plot, Forest plot

Biostat, Inc. 14 North Dean Street Englewood, New Jersey 07631 USA

<http://www.MetaAnalysis.com>

- **MetaWin 2.0** (*Operating system: windows*)

Input data: 2 x 2 tables, means, standard deviations, correlation coefficients

Statistical models: Fixed effects, random effects, mixed model regression, heterogeneity test, cumulative analysis, Fail-safe calculations, Correlation methods for detecting publication bias

Output effect measures: Hedges' d for standardized mean difference, response ratio, odds ratio, rate difference, relative rate and Fisher's z-transform

Plots: Funnel plots, radial plots, histograms, quantile plots, regression plots

Sinauer Associates, Inc. P.O Box 407, Sunderland, Massachusetts USA

<http://www.metawinsoft.com/>

- **WeasyMa** (*Operating system: windows*)

Input data: 2 x 2 tables

Statistical models: Fixed effects, random effects, heterogeneity test

Output effect measures: odds ratio (or relative odds reduction), risk difference (or absolute risk reduction), NNT, rate ratio, rate difference

Plots: Forest plot, cumulative meta-analysis graph¹, Funnel plots, Radial plots (Galbraith's plot), L'Abbe plot

ClinInfo SA and EA643 UFR RTH LAËNNEC, Rue Guillaume Paradin BP 8071 - 69376 Lyon Cedex 08

FRANCE

<http://www.weasyrna.com/>

¹*Cumulative meta-analysis:* Τεχνική για την πραγματοποίηση νέων μεταanalύσεων κάθε φορά που είναι διαθέσιμα τα αποτελέσματα νέων κλινικών μελετών

9.3 Δωρεάν προγράμματα (*Free-ware software*)

- **Review Manager (RevMan) 5.0** (*Operating system: windows*)

Το λογισμικό πακέτο RevMan 5.0 σχεδιάστηκε για την εισαγωγή πρωτοκόλλων αναθεώρησης μελετών, σύμφωνα με τις αρχές της *Cochrane Collaboration*. Το πρόγραμμα περιλαμβάνει ένα δομημένο κείμενο σύμφωνα με το οποίο θα πρέπει να γίνει η ανασκόπηση των πινάκων των συμπεριλαμβανόμενων καθώς επίσης και αποκλεισμένων μελετών. Το πρόγραμμα είναι βασισμένο στο λειτουργικό σύστημα Windows και είναι εύκολο στη χρήση. Μπορούν να εισαχθούν τόσο δίτιμα όσο και συνεχή δεδομένα, ενώ στην ανάλυση υπάρχει η δυνατότητα και των δυο μοντέλων επίδρασης (*fixed* και *random effects*). Ελέγχει για ετερογένεια, ενώ η επίδραση της θεραπείας μπορεί να εκφραστεί σε *odds ratio*, *relative risk*, *risk difference* και *mean difference*. Παρέχεται το *forest plot* όχι όμως και το γράφημα *funnel plot*.

Το πρόγραμμα RevMan είναι διαθέσιμο στη διεύθυνση:

<http://cochrane.org/cochrane/revman.htm/>

- **EasyMA** (*Operating system: MS-DOS*)

Το πρόγραμμα EasyMA κατασκευάστηκε από τον *Michel Cucherat* στο University of Lyon. Το EasyMA αναπτύχθηκε για τη μετανάλυση κλινικών δοκιμών με δυο ομάδες θεραπείας οι οποίες έχουν μια ή περισσότερες δίτιμες μεταβλητές απόκρισης. Για την εκτίμηση της ολικής επίδρασης (με *odds ratio* ή *relative risk* ή *risk difference*) διαθέτει τα μοντέλα σταθερών επιδράσεων (*Peto*, *Mantel-Haenszel*) και το μοντέλο τυχαίων επιδράσεων (*DerSimonian and Laird*). Επιπλέον δίνεται ο αριθμός των ασθενών NNT (*Number Needed to Treat*) που χρειάζονται για να αποφευχθεί ένα γεγονός. Μπορεί επίσης να εκτελέσει σταθμισμένη ή μη-σταθμισμένη παλινδρόμηση. Τα γραφήματα που είναι διαθέσιμα είναι το *forest plot*, το *funnel plot*, το *radial plot* καθώς και γραφήματα για *cumulative* μετανάλυση. Παρέχει τον έλεγχο ετερογένειας, το τεστ των *Begg and*

Mazumdar's για τον έλεγχο του *publication bias*, ενώ υπολογίζει και τον αριθμό μη δημοσιευμένων μελετών σύμφωνα με τη μέθοδο του *Rosenthal*.

Το πρόγραμμα EasyMa είναι διαθέσιμο στη διεύθυνση:

<http://www.spc.univlyon.fr/~mceu/easyma>

- **Meta-Analyst** (*Operating system: MS-DOS*)

Το πρόγραμμα Meta-analyst κατασκευάστηκε από τον *Joseph Lau* για το New England Medical Center. Όπως και το EasyMA, αυτό το πρόγραμμα μπορεί να εκτελέσει κλασσική ή *cumulative* μετανάλυση σε κλινικές μελέτες με δυο ομάδες θεραπείας που έχουν δίτιμη μεταβλητή απόκρισης. Το πρόγραμμα είναι εύκολο στη χρήση, ενώ μπορεί να χρησιμοποιηθεί το μοντέλο σταθερών ή τυχαίων επιδράσεων για τον συνδυασμό *odds ratios*, *relative risks* και *risk differences*. Παρέχει τον έλεγχο ετερογένειας ενώ προσφέρονται ανάλογα γραφήματα με το EasyMa. Το πρόγραμμα περιλαμβάνει ανάλυση παλινδρόμησης και εκτίμηση μεγέθους δείγματος.

Για το πρόγραμμα MetaAnalyst μπορείτε να απευθυνθείτε στον κατασκευαστή του:

joseph.lau@es.nemc.org

- **EpiMeta** (*Operating system: MS-DOS*)

Το πρόγραμμα EpiMeta κατασκευάστηκε για τα Centers of Disease Control, προκειμένου να εκτελέσει μεθόδους μετανάλυσης για τον συνδυασμό μελετών. Εκτελεί *fixed* και *random effects* μεθόδους. Αν και πρόκειται για ένα πρόγραμμα που “τρέχει” σε MS-DOS, η εισαγωγή δεδομένων, η διαχείριση των αρχείων και η *subgroup* ανάλυση γίνονται σε παραθυρικό περιβάλλον κάνοντας το πρόγραμμα εύκολο στη χρήση.

Το πρόγραμμα EpiMeta βρίσκεται στη διεύθυνση:

<http://www.cdc.gov/epo/dpram/epimeta/epimeta.htm>

- **Meta-Analysis Easy To Answer-META** (*Operating system: MS-DOS*)

Το πρόγραμμα META κατασκευάστηκε από τον *David Kenny* για το τμήμα Ψυχολογίας του University of Connecticut. Το META υπολογίζει την επίδραση της κάθε μελέτης, συνδυάζει αυτά τα αποτελέσματα και κάνει τον έλεγχο ετερογένειας. Οι εκτιμητές μπορούν να σταθμιστούν ως προς το μέγεθος της κάθε μελέτης, την διακύμανση του εκτιμητή ή οποιεσδήποτε άλλες τιμές θελήσουμε. Με το META δεν μπορεί να ελεγχθεί η συσχέτιση μεταξύ των επιδράσεων των μελετών. Η επίδραση της μελέτης μπορεί να εκφραστεί ως *standardized mean difference* ή ως διαφορά ποσοστών.

Το πρόγραμμα MetaAnalyst είναι διαθέσιμο στη διεύθυνση:

<http://users.rcn.com/dakenny/meta.htm>

- **MetaStat** (*Operating system: MS-DOS*)

Το πρόγραμμα MetaStat κατασκευάστηκε με χρηματοδότηση του National Institute of Health από την LMP Associates Inc. Το πρόγραμμα υπολογίζει διάφορα μέτρα επίδρασης της θεραπείας, ελέγχει για ετερογένεια, εκτελεί παλινδρόμηση και υπολογίζει τον πίνακα ANOVA, χρησιμοποιεί μεθόδους μετανάλυσης που περιγράφονται στο βιβλίο των *Hedges* και *Olkin*, υπολογίζει το πλήθος των μη-δημοσιευμένων μελετών (*Rosenthal's N*) και δίνει σχετικές γραφικές παραστάσεις.

Το πρόγραμμα MetaStast είναι διαθέσιμο στη διεύθυνση:

<http://www.edres.org/meta/metastat.htm>

- **Meta-Analysis 5.3** (*Operating system: MS-DOS*)

Το πρόγραμμα Meta-Analysis κατασκευάστηκε από τον *Ralf Schwarzer*. Είναι ένα από τα πιο πολυχρησιμοποιημένα προγράμματα μετανάλυσης. Μπορεί να συνδυάσει είτε τις επιδράσεις της θεραπείας είτε τα *p-values* των στατιστικών ελέγχων. Ερευνητές οι οποίοι

δεν ενδιαφέρονται για τη μετανάλυση, αλλά που ψάχνουν ένα κατάλληλο εργαλείο για να υπολογίσουν *t-tests* ή να εξετάσουν τη συσχέτιση μπορούν επίσης να ωφεληθούν από αυτό το πρόγραμμα.

Το πρόγραμμα Meta-Analyst 5.3 είναι διαθέσιμο στη διεύθυνση:

http://www.fu-berlin.de/gesund/publicat/ehps_cd/health/meta_e.htm

9.4 Ρουτίνες μετανάλυσης διαθέσιμες σε γενικά στατιστικά πακέτα (*macros*).

1. SAS. www.sas.com
2. Stata. <http://www.stata.com/>
3. StatsDirect. <http://www.camcode.com/>
4. StatXact. <http://www.cytel.com/products/statxact/statxact1.html/>
5. SPSS. <http://www.wam.umd.edu/~wilsondb/ma.html>.
6. S-Plus. <http://www.mathsoft.com/splus/>
<http://www.prw.le.ac.uk/epidemiology/personal/ajs22/meta/routine.html/>
<http://www.research.att.com/~dumouchel/bsoft.html/>
7. R. <http://cran.r-project.org/>
8. WinBUGS. <http://www.mrc-bsu.cam.ac.uk/bugs>
 - Το πρόγραμμα αυτό εκτελεί Μπεϋζιανές αναλύσεις

9.5 Σχετική Βιβλιογραφία

Είναι φανερό ότι υπάρχει πληθώρα προγραμμάτων που εκτελούν μετανάλυση. Αν και για κάποια από αυτά τα προγράμματα είναι σαφής η μεθοδολογία που χρησιμοποιούν, για κάποια άλλα δεν έχουμε επαρκή στοιχεία. Τα τελευταία χρόνια έχουν δημοσιευτεί άρθρα τα οποία συγκρίνουν προγράμματα μετανάλυσης έτσι ώστε να υπάρχει σαφής εικόνα για το τι μπορεί να κάνει το καθένα ή για να αναδείξει τα σημεία στα οποία ένα πρόγραμμα υστερεί.

Στην συνέχεια παραθέτουμε μερικά από αυτά τα άρθρα στα οποία μπορεί να απευθυνθεί ο αναγνώστης για περισσότερες πληροφορίες.

1. Arthur, W., Bennett, W. & Huffcutt, A. (1994). Choice of software and programs in meta-analysis research: Does it make a difference? *Educational & Psychological Measurement*, 54, 776-787.
2. Normand, S.L. (1995). Meta-analysis software: A comparative review. *American Statistician*, 49, 298-309.
3. Cucherat M, Boissel J-P, Leizorovicz A, Haugh MC. (1997) EasyMA: a program for the meta- analysis of clinical trials. *Computer Methods and Programs in Biomedicine*; 53:187- 190.
4. Egger, M., Sterne, J.A.C. & Davey Smith, G. (1998). Meta-analysis software. *British Medical Journal*, 316. <http://www.bmj.com/archive/7126/7126ed9.htm/>
5. Sutton, A.J., Lambert, P.C., Hellmich, M., et al. (2000). Meta-analysis in practice: A critical review of available software. In D.A. Berry & D.K. Stangl (Eds.), *Meta-analysis in Medicine and Health Policy*. New York: Marcel Dekker.
6. Sterne, J.A.C., Egger, M. & Sutton, A.J. (2001). Meta-analysis software. In M. Egger, G. Davey Smith & D.G. Altman (Eds.), *Systematic Reviews in Health Care: Meta-analysis in Context* (pp. 336-346). London: BMJ Pub. Group.
7. Sterne, J.A.C., Bradburn, M.J. & Egger, M. (2001). Meta-analysis in StataTM. In M. Egger, G. Davey Smith & D.G. Altman (Eds.), *Systematic Reviews in Health Care: Meta-analysis in Context* (pp. 347-369). London, UK: BMJ Pub. Group.

ΚΕΦΑΛΑΙΟ 10

Εφαρμογή

10.1 Εισαγωγή

Σε αυτό το κεφάλαιο θα παρουσιάσουμε ένα παράδειγμα μετανάλυσης. Η μετανάλυση θα βασιστεί σε 16 κλινικές μελέτες στις οποίες εξετάστηκε αν η ενδοφλέβια χορήγηση μαγνησίου σε ασθενείς που υπέστησαν οξύ έμφραγμα του μυοκαρδίου, μειώνει την πιθανότητα θανάτου. Τα δεδομένα των μελετών αυτών αναφέρονται στην *Whitehead* [2002], ενώ η ανάλυση θα γίνει με χρήση του προγράμματος “*Comprehensive Meta-Analysis*”. Σκοπός αυτής της εφαρμογής δεν είναι να πραγματοποιήσουμε όλα τα στάδια της μετανάλυσης, αλλά να παρουσιάσουμε τις δυνατότητες που μας δίνει ένα ειδικό πακέτο μετανάλυσης στην ανάλυση των δεδομένων.

10.2 Ανάλυση

Στον πίνακα 10.1 δίνονται τα δεδομένα των 16 μελετών. Σε κάθε μελέτη υπάρχουν δυο ομάδες θεραπείας (“*treated*” και “*control*”), ενώ για την κάθε ομάδα δίνεται ο συνολικός αριθμός των ασθενών και ο αντίστοιχος αριθμός των θανάτων που σημειώθηκαν σε κάθε ομάδα.

| Study | TREATED GROUP | | CONTROL GROUP | |
|----------------------------------|---------------|-------|---------------|-------|
| | Dead | Total | Dead | Total |
| 1 Morton (1984) | 1 | 40 | 2 | 36 |
| 2 Rasmussen (1986) | 9 | 135 | 23 | 135 |
| 3 Smith (1986) | 2 | 200 | 7 | 200 |
| 4 Abraham (1987) | 1 | 48 | 1 | 46 |
| 5 Feldstedt (1988) | 10 | 150 | 8 | 148 |
| 6 Shechter (1989) | 1 | 59 | 9 | 56 |
| 7 Ceremuzynski (1989) | 1 | 25 | 3 | 23 |
| 8 Singh (1990) | 6 | 76 | 11 | 75 |
| 9 Shechter and Hid (1991) | 2 | 89 | 12 | 80 |
| 10 Woods et al. (LIMIT-2) (1992) | 90 | 1159 | 118 | 1157 |
| 11 Bertscat (1989) | 0 | 22 | 1 | 21 |
| 12 Pereira (1990) | 1 | 27 | 7 | 27 |
| 13 Golf (1991) | 5 | 23 | 13 | 33 |
| 14 Thogersen (1991) | 4 | 130 | 8 | 122 |
| 15 Schechter and Hod (1995) | 4 | 107 | 17 | 108 |
| 16 ISIS-4 (1995) | 2216 | 29011 | 2103 | 29039 |

Πίνακας 10.1 Intravenous magnesium following acute myocardial infraction

[Πηγή: A. Whitehead, 2002]

Στην συνέχεια, για κάθε μελέτη υπολογίζουμε την επίδραση της θεραπείας σε σχέση με το *control*. Όπως έχουμε αναφέρει στο 3^ο κεφάλαιο έχουμε τρεις επιλογές για να υπολογίσουμε αυτή την επίδραση, χρησιμοποιώντας τη διαφορά των ποσοστών των δυο ομάδων, τον σχετικό κίνδυνο (*relative risk*) ή τον λόγο των σχετικών ποσοστών (*odds ratio*). Στο παράδειγμα αυτό θα εκτιμήσουμε την ολική επίδραση της θεραπείας μόνο με το *odds ratio*.

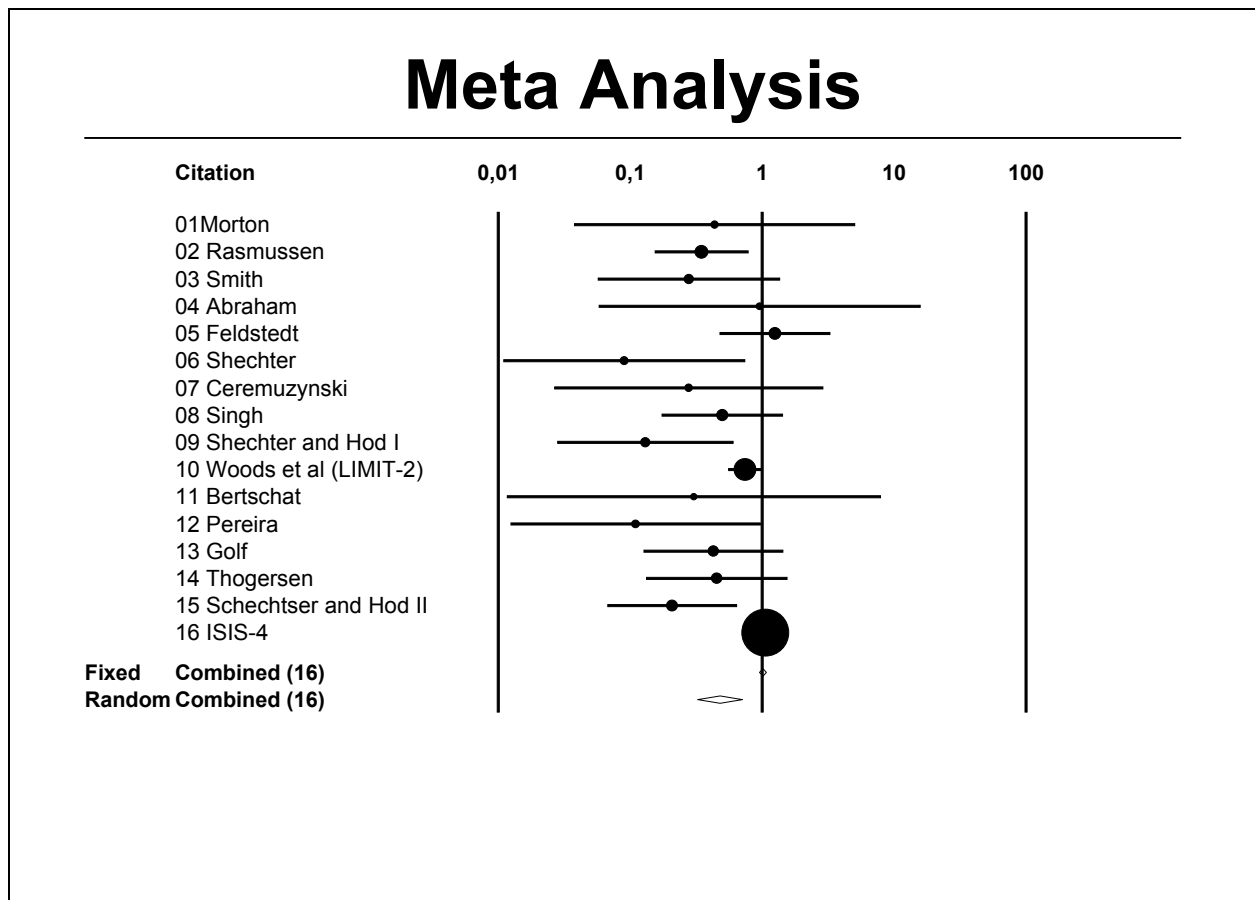
Στο *output* 10.1 δίνονται οι εκτιμητές του *odds ratio* για κάθε μελέτη και τα αντίστοιχα βάρη στην περίπτωση των δυο μοντέλων μετανάλυσης (*fixed effect* και *random effect*). Μπορούμε να παρατηρήσουμε ότι στις περισσότερες μελέτες (εκτός της 5^{ης} και της 16^{ης} μελέτης), η θεραπεία προλαμβάνει το θανατηφόρο συμβάν, αφού ο εκτιμητής του *odds ratio* για κάθε μια από τις μελέτες αυτές έχει τιμή μικρότερη της μονάδας. Στις δυο τελευταίες γραμμές δίνεται η ολική επίδραση της θεραπείας με βάση και τα δυο μοντέλα μετανάλυσης.

Meta Analysis

| Citation | N1 | Events1 | N2 | Events2 | Effect | FixedWt | RandomWt |
|-----------------------------|--------------|-------------|--------------|-------------|--------------|---------|----------|
| 01 Morton | 40 | 1 | 36 | 2 | ,436 | ,643 | ,559 |
| 02 Rasmussen | 135 | 9 | 135 | 23 | ,348 | 5,832 | 2,474 |
| 03 Smith | 200 | 2 | 200 | 7 | ,278 | 1,531 | 1,129 |
| 04 Abraham | 48 | 1 | 46 | 1 | ,957 | ,489 | ,439 |
| 05 Feldstedt | 150 | 10 | 148 | 8 | 1,250 | 4,179 | 2,118 |
| 06 Shechter | 59 | 1 | 56 | 9 | ,090 | ,870 | ,723 |
| 07 Ceremuzynski | 25 | 1 | 23 | 3 | ,278 | ,702 | ,603 |
| 08 Singh | 76 | 6 | 75 | 11 | ,499 | 3,478 | 1,922 |
| 09 Shechter and Hod I | 89 | 2 | 80 | 12 | ,130 | 1,641 | 1,187 |
| 10 Woods et al (LIMIT-2) | 1159 | 90 | 1157 | 118 | ,741 | 46,547 | 3,933 |
| 11 Bertschat | 22 | | 21 | 1 | ,304 | ,362 | ,334 |
| 12 Pereira | 27 | 1 | 27 | 7 | ,110 | ,812 | ,683 |
| 13 Golf | 23 | 5 | 33 | 13 | ,427 | 2,615 | 1,625 |
| 14 Thogersen | 130 | 4 | 122 | 8 | ,452 | 2,553 | 1,601 |
| 15 Schechtser and Hod II | 107 | 4 | 108 | 17 | ,208 | 3,035 | 1,778 |
| 16 ISIS-4 | 29011 | 2216 | 29039 | 2103 | 1,059 | 998,781 | 4,278 |
| Fixed Combined (16) | 31301 | 2353 | 31306 | 2343 | 1,015 | | |
| Random Combined (16) | 31301 | 2353 | 31306 | 2343 | ,480 | | |

Output 10.1 Υπολογισμός του odds ratio για τις 16 μελέτες μαζί με τα αντίστοιχα βάρη με fixed effect και random effect μοντέλο.

Το γράφημα *forest plot* δίνεται στο *output 10.2* όπου τα διαστήματα εμπιστοσύνης της επίδρασης της θεραπείας σε κάθε μια από τις 16 μελέτες έχουν υπολογιστεί σε επίπεδο σημαντικότητας $\alpha=5\%$. Όπως μπορούμε να παρατηρήσουμε η ακρίβεια της εκτίμησης του *odds ratio* είναι ανάλογη με το μέγεθος του κύκλου που σημειώνεται σε κάθε διάστημα. Γι' αυτό άλλωστε την μεγαλύτερη ακρίβεια στην εκτίμηση την δίνουν οι μελέτες 10 και 16, οι οποίες έχουν χρησιμοποιήσει πολύ μεγάλο δείγμα ασθενών. Επίσης παρατηρούμε ότι ενώ οι μικρότερες σε μέγεθος μελέτες δίνουν θετική επίδραση της θεραπείας (με εξαίρεση την 5^η μελέτη), η 16^η μελέτη η οποία είναι η μεγαλύτερη σε μέγεθος δίνει αρνητική επίδραση της θεραπείας. Από το γράφημα αυτό, αλλά και από τις τιμές της ολικής επίδρασης τη θεραπείας κάτω από τα δυο μοντέλα μετανάλυσης, είναι φανερό το πρόβλημα της ετερογένειας των μελετών.



Output 10.2 Forest plot των 16 μελετών χρησιμοποιώντας το odds ratio

Στα επόμενα *output* δίνονται τα αποτελέσματα του ελέγχου ετερογένειας και οι εκτιμητές της ολικής επίδρασης της θεραπείας. Το πακέτο παρέχει τον υπολογισμό και των τριών προσεγγίσεων που έχουμε αναφέρει στο 4^ο κεφάλαιο: *Mantel-Haenszel*, *Peto* και *Inverse Variance*.

Meta Analysis for all studies

Basic statistics

Number of outcomes = 16
Number of individual cases = 62607

| Index | Odds ratio and 95% confidence interval | | | Test of null (2 tail) | | Test of heterogeneity | | |
|-------------------|--|-------------|-------------|-----------------------|--------|-----------------------|---------|--------|
| | Point estimate | Lower limit | Upper limit | Z-Value | PValue | Q-Value | df (Q) | PValue |
| Odds ratio | | | | | | | | |
| Mantel-Haenszel | | | | | | | | |
| Fixed Effects | 1,0061 | ,9480 | 1,0678 | ,2000 | ,8414 | 47,1403 | 15,0000 | ,0000 |
| Random Effects | ,4829 | ,3286 | ,7097 | -3,7055 | ,0002 | | | |
| Inverse variance | | | | | | | | |
| Fixed Effects | 1,0149 | ,9560 | 1,0774 | ,4835 | ,6287 | 47,0593 | 15,0000 | ,0000 |
| Random Effects | ,4803 | ,3255 | ,7086 | -3,6955 | ,0002 | | | |
| Peto one-step | | | | | | | | |
| Fixed Effects | 1,0061 | ,9480 | 1,0678 | ,2001 | ,8414 | 55,3128 | 15,0000 | ,0000 |

Output 10.3 Αποτελέσματα της μετανάλυσης των 16 μελετών χρησιμοποιώντας το odds ratio

Στην τελευταία στήλη του *output* 10.3, το *p-value* του τεστ ετερογένειας είναι μικρότερο του επιπέδου σημαντικότητας $\alpha=10\%$, επομένως οι μελέτες είναι σημαντικά ετερογενείς. Οι εκτιμητές της ολικής επίδρασης της θεραπείας παρέχονται και για τα δυο μοντέλα, όμως πιο αξιόπιστος είναι ο εκτιμητής του μοντέλου τυχαίων επιδράσεων λόγω της ύπαρξης ετερογένειας. Παρατηρούμε ότι ο *Mantel-Haenszel* εκτιμητής δίνεται και στην περίπτωση του μοντέλου τυχαίων επιδράσεων. Η προσέγγιση αυτή δεν είναι συνήθης και δεν προτείνεται στη σχετική βιβλιογραφία.

Στην συνέχεια δίνουμε κάποιες λεπτομέρειες για να μπορέσουμε να κατανοήσουμε το πώς προέκυψαν τα αποτελέσματα του τελευταίου *output*:

Για το μοντέλο σταθερών επιδράσεων, η εκτίμηση της ολικής επίδρασης της θεραπείας στην περίπτωση της MH προσέγγισης, της *Inverse Variance* και της μεθόδου του *Peto* προκύπτει με βάση τους τύπους των παραγράφων 4.2.1.1, 4.2.2 και 4.2.3.

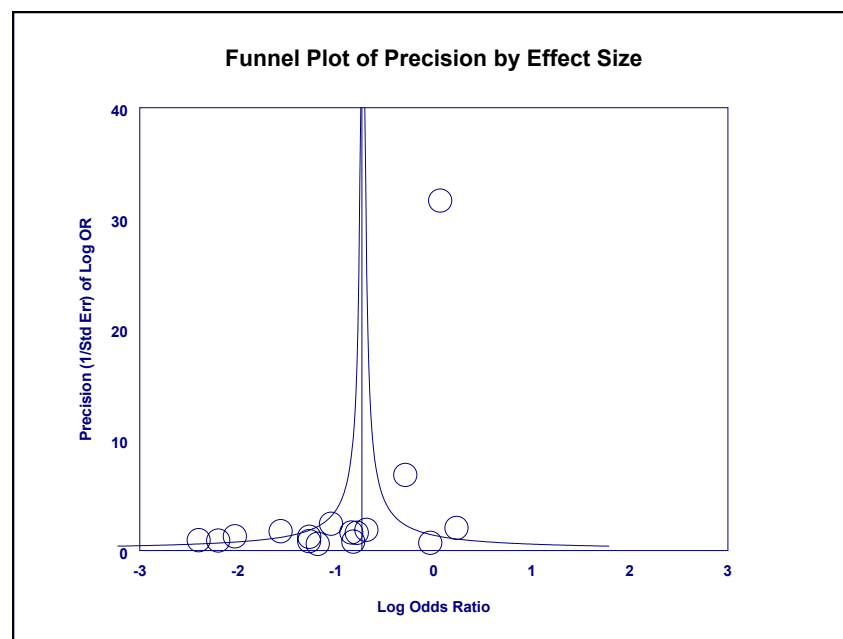
Για τις δυο πρώτες προσεγγίσεις (MH και *Inverse Variance*), το τεστ ετερογένειας προκύπτει με βάση τις αντίστροφες τιμές των διακυμάνσεων (παράγραφος 5.3). Επειδή ο

εκτιμητής της ολικής επίδρασης είναι διαφορετικός στις δυο προσεγγίσεις, το τεστ καταλήγει σε ελαφρώς διαφορετικές τιμές. Στην προσέγγιση του *Peto* τα βάρη στο τεστ ετερογένειας είναι ίσα με την υπεργεωμετρική διακύμανση v_i . Και στις τρεις προσεγγίσεις, χρησιμοποιούνται οι λογαριθμικές τιμές των εκτιμητών του *odds ratio*.

Στο μοντέλο τυχαίων επιδράσεων, χρησιμοποιούνται ως βάρη οι αντίστροφές τιμές των $Var(\hat{\theta}_i)$, αλλά επειδή η εκτίμηση του τ^2 εξαρτάται από την τιμή του τεστ ετερογένειας, τα αποτελέσματα για τις δυο προσεγγίσεις (MH, *Inverse Variance*) είναι ελαφρώς διαφορετικά. Ένας σημαντικός περιορισμός του προγράμματος είναι ότι η εκτίμηση του τ^2 γίνεται μόνο με την μέθοδο των ροπών.

Όπως είναι φανερό από τα αποτελέσματα του τεστ ετερογένειας, αλλά και από το γράφημα *forest plot*, οι μελέτες είναι σημαντικά ετερογενείς και επομένως θα πρέπει να εξεταστούν πιθανοί λόγοι που προκαλούν αυτή την ετερογένεια. Επειδή στο συγκεκριμένο παράδειγμα τα δεδομένα δεν είναι επαρκή δεν μπορούμε να εκτελέσουμε *subgroup* ανάλυση ή *meta-regression*. Ακόμα όμως κι αν είχαμε αυτή τη δυνατότητα, δεν μπορούμε να τις εκτελέσουμε με το συγκεκριμένο πρόγραμμα.

Σχετικά με το θέμα της μεροληψίας, και ειδικότερα με το σφάλμα δημοσίευσης, έχουμε τη δυνατότητα να κατασκευάσουμε το γράφημα “χοάνη”. Όπως βλέπουμε, το γράφημα αυτό είναι μη συμμετρικό, δίνοντας ένδειξη για πιθανή μεροληψία λόγω δημοσίευσης.



Output 10.4 Funnel plot για τις 16 μελέτες

Το πακέτο δεν προσφέρει άλλες μεθόδους για τον έλεγχο του σφάλματος μεροληψίας και αυτό είναι άλλο ένα σημαντικό μειονέκτημά του. Όπως αναφέραμε στο 8^ο κεφάλαιο, υπάρχουν προσεγγίσεις για την διόρθωση του σφάλματος λόγω της δημοσίευσης οι οποίες όμως δεν παρέχονται στο πρόγραμμα που χρησιμοποιήσαμε.

Συμπερασματικά, βλέπουμε ότι ακόμα και τα ειδικά προγράμματα μετανάλυσης δεν δίνουν ένα ολοκληρωμένο πλαίσιο για την σωστή διεξαγωγή της μετανάλυσης. Στο 9^ο κεφάλαιο είδαμε ότι τα περισσότερα από αυτά τα προγράμματα προσφέρουν ένα μόνο μέρος από το σύνολο των αναγκαίων αναλύσεων. Για παράδειγμα, αντιμετωπίζουν μόνο την περίπτωση που τα δεδομένα είναι δίτιμα ή συνεχή. Όπως αναφέρθηκε στην αρχή ο εκτιμητής της μεταβλητότητας των μελετών εκτιμάται μόνο με τη μέθοδο των ροπών, η οποία όμως δεν δίνει το ίδιο ακριβή αποτελέσματα όπως οι μέθοδοι ML ή REML. Η μετανάλυση δεν είναι μόνο η εύρεση του εκτιμητή της ολικής επίδρασης, όπως αφήνεται να εννοηθεί μέσα από τα αποτελέσματα του προγράμματος. Είναι ένα σύνολο αναλύσεων, οι οποίες λαμβάνονται ανάλογα υπόψη και δίνουν το τελικό συμπέρασμα. Επομένως, η επιλογή του προγράμματος με το οποίο θα πραγματοποιηθεί η μετανάλυση οφείλει να είναι προσεκτική, ενώ είναι αναγκαία η γνώση του συνόλου των μεθοδολογιών για να είναι η ανάλυση πλήρης και ακριβής.

ΚΕΦΑΛΑΙΟ 11

Συμπεράσματα

Όπως αναφέραμε στο 1^ο κεφάλαιο η μετανάλυση είναι ένα σύνολο στατιστικών τεχνικών, με τις οποίες συνδυάζονται και αναλύονται τα αποτελέσματα ανεξάρτητων κλινικών δοκιμών. Για να πραγματοποιηθεί όμως μια μετανάλυση, είναι απαραίτητη η συστηματική ανασκόπηση των μελετών (*systematic review*), προκειμένου να εντοπιστούν εκείνες οι μελέτες των οποίων τα αποτελέσματα θα εξεταστούν και εν συνεχεία θα συνδυαστούν. Ενώ η συστηματική ανασκόπηση είναι μόνο μια ποιοτική-περιγραφική μέθοδος συγκέντρωσης των αποτελεσμάτων, η μετανάλυση προσφέρει την ποσοτική ανασκόπηση των σχετικών μελετών υπολογίζοντας την ολική επίδραση της θεραπείας.

Τα οφέλη της μετανάλυσης είναι αρκετά. Είναι γνωστό ότι αρκετές μελέτες είναι περιορισμένες σε μέγεθος για να ανιχνεύσουν μια μικρή επίδραση της θεραπείας. Όταν όμως συνδυάσουμε όλες αυτές τις μελέτες το μέγεθος του δείγματος αυξάνεται και επομένως επιτυγχάνουμε μεγαλύτερη στατιστική ισχύ για την ανίχνευση θετικής επίδρασης της θεραπείας, εφόσον πραγματικά υπάρχει. Επιπλέον οι εκτιμητές της ολικής επίδρασης της θεραπείας που προκύπτουν είναι πιο ακριβείς, αφού βασίζονται σε μεγαλύτερο ποσό πληροφορίας. Όπως αναφέρθηκε στο 5^ο κεφάλαιο η μετανάλυση εξετάζει και την ετερογένεια των μελετών ως προς την επίδραση της θεραπείας, ενώ μπορεί ακόμα και να εξηγήσει τους λόγους που την προκαλούν χρησιμοποιώντας τη μέθοδο της μετα-παλινδρόμησης. Το γεγονός ότι η μετανάλυση καταλήγει σε ένα ολικό συμπέρασμα ενώ τα επιμέρους αποτελέσματα των μελετών μπορεί να είναι αντιφατικά, την κάνει ακόμα πιο απαραίτητη στον ιατρικό χώρο, αφού η ύπαρξη διαφορετικών συμπερασμάτων προκαλεί σύγχυση. Τέλος η μετανάλυση δίνει την δυνατότητα να εξεταστούν ερωτήματα τα οποία δεν μπορούν να απαντήσουν οι μεμονωμένες μελέτες. Έτσι μπορούμε να αξιολογήσουμε ζητήματα σχετικά με την επιλογή των ασθενών ή με την επιλογή της θεραπείας, ενώ παράλληλα τα αποτελέσματα γενικεύονται σε ευρύτερο πληθυσμό.

Παρόλα αυτά η μετανάλυση έχει γίνει στόχος αρκετών αρνητικών κριτικών. Είναι λογικό επακόλουθο ότι τα αποτελέσματα μιας μετανάλυσης είναι αξιόπιστα αν οι επιμέρους μελέτες

είναι καλής ποιότητας. Στην πράξη όμως αυτό δεν συμβαίνει. Αρκετές μελέτες είναι χαμηλής ποιότητας με σημαντικά ποσοστά μεροληψίας, με αποτέλεσμα η μεροληψία αυτή να μεταφέρεται στην μετανάλυση. Την μετανάλυση όμως την επηρεάζουν και άλλες μεροληψίες οι οποίες μπορεί να προκύψουν στην πορεία. Η μεροληψία επιλογής των μελετών και κυρίως το συστηματικό σφάλμα δημοσίευσης είναι το σημαντικότερο πρόβλημα της μετανάλυσης. Στο 8^ο κεφάλαιο παρουσιάσαμε τρόπους ανίχνευσης αυτής της μεροληψίας και τρόπους διόρθωσής της. Από αυτές τις μεθόδους, το γράφημα “χοάνη” μας δίνει μια καλή οπτική ανίχνευση του συστηματικού σφάλματος, δεδομένου όμως ότι το πλήθος των μελετών είναι μεγάλο. Από τις υπόλοιπες μεθόδους που έχουν προταθεί, αν και η χρήση σταθμισμένων κατανομών (*weighted distributions*) έχει σημαντικές προοπτικές, επειδή η εφαρμογή αυτών των μεθόδων είναι πολύπλοκη δεν είναι εύκολη η ευρεία χρήση τους στις μεταanalύσεις που πραγματοποιούνται.

Η “διαφορετικότητα” των μελετών είναι άλλο ένα πρόβλημα το οποίο επηρεάζει την μετανάλυση. Μερικές φορές δημιουργείται η εντύπωση ότι συνδυάζονται ετερόκλητες μελέτες ή, όπως έχει επικρατήσει να λέγεται, συνδυάζονται “*apples and oranges*”. Οι διαθέσιμες μελέτες διαφέρουν στο σχεδιασμό, στην ποιότητα, στον πληθυσμό που εξετάζεται ή στην απόκριση που μετριέται για κάθε ασθενή, έτσι είναι εύλογο το ερώτημα αν τα αποτελέσματά τους μπορούν να συνδυαστούν. Ο στατιστικός έλεγχος ετερογένειας έχει χαμηλή ισχύ (παράγραφος 5.3, σελ.56), επομένως δεν μπορούμε να βασιστούμε μόνο στο *p-value* του ελέγχου για να αποφασίσουμε αν οι μελέτες είναι όντως ετερογενείς. Επειδή οι μελέτες αναπόφευκτα διαφέρουν, η ερώτηση που τίθεται δεν είναι αν υπάρχει ετερογένεια, αλλά εάν ο βαθμός της επηρεάζει σημαντικά τα συμπεράσματα που εξάγονται. Καθώς η ετερογένεια είναι συνδεδεμένη και με την επιλογή του στατιστικού μοντέλου ανάλυσης των δεδομένων, το θέμα αυτό θα πρέπει να αντιμετωπιστεί με προσοχή, θεωρώντας πιθανό και το ενδεχόμενο να μην συνδυαστούν τελικά οι μελέτες λόγω της μεγάλης μεταξύ τους ετερογένειας.

Άλλο ένα σημείο, για το οποίο υπάρχει έντονη διαμάχη μεταξύ των ερευνητών, είναι η επιλογή του κατάλληλου στατιστικού μοντέλου για τον υπολογισμό της ολικής επίδρασης της θεραπείας. Το μοντέλο σταθερών επιδράσεων προτείνεται όταν οι μελέτες είναι ομοιογενείς ή όταν το πλήθος τους είναι μικρό. Πότε όμως οι μελέτες είναι πραγματικά ομοιογενείς; Ακόμα κι αν το πλήθος τους είναι μικρό, τι μας εξασφαλίζει ότι η ετερογένεια είναι ασήμαντη; Αντίστοιχα το μοντέλο τυχαίων επιδράσεων προτείνεται όταν οι μελέτες είναι ετερογενείς και

το πλήθος τους είναι σχετικά μεγάλο. Η ετερογένεια όμως αυτή θα πρέπει να είναι τυχαία, ενώ οι μελέτες πρέπει να είναι αντιπροσωπευτικές ενός υπερπληθυσμού μελετών. Πως άραγε μπορούν να ελεγχθούν αυτές οι υποθέσεις; Η απάντηση σε αυτές τις ερωτήσεις δεν είναι σαφής. Έτσι η επιλογή του μοντέλου αφήνεται στην κρίση και την εμπειρία του ερευνητή, ο οποίος θα πρέπει με επιχειρήματα να υποστηρίξει την οποιαδήποτε τελική απόφαση.

Υπάρχουν κι άλλα μικρότερα προβλήματα τα οποία σχετίζονται με την σωστή διεξαγωγή της μετανάλυσης. Επειδή για να πραγματοποιηθεί μια μετανάλυση απαιτείται ένα μέτρο κοινό σε όλες τις μελέτες, “χάνεται” κάποιο ποσό πληροφορίας που ενδέχεται να υπάρχει σε κάποιες από τις επιμέρους μελέτες. Επιπλέον η πραγματοποίηση πολλών αναλύσεων υποομάδων (*subgroup analyses*) ή πολλών αναλύσεων ευαισθησίας (*sensitivity analyses*) μπορεί να οδηγήσει σε παραπλανητικά συμπεράσματα. Τέλος, αν και έχουν κατασκευαστεί αρκετά προγράμματα που πραγματοποιούν μετανάλυση, οι μεθοδολογίες που είναι διαθέσιμες στο καθένα από αυτά είναι περιορισμένες. Εφαρμόζονται μόνο σε συνεχή ή δίτιμα δεδομένα ενώ για το πρόβλημα της μεροληψίας δεν πραγματοποιούνται όλες οι μέχρι τώρα γνωστές προσεγγίσεις για την διόρθωσή του.

Αυτό όμως που έχει δημιουργήσει έντονο προβληματισμό είναι τα αντιφατικά αποτελέσματα τα οποία μπορεί να υπάρξουν μεταξύ μιας μετανάλυσης και μιας μεταγενέστερης μεγάλης κλινικής μελέτης (*mega trial*). Πολλά άρθρα έχουν γραφεί σχετικά με αυτό το θέμα, ενδεικτικά όμως θα αναφέρουμε το άρθρο των *Lelorier et al.* [1997] όπου συγκρίθηκαν 12 μεγάλες τυχαιοποιημένες κλινικές μελέτες οι οποίες ήταν μεταγενέστερες 19 αντίστοιχων μετανάλυσεων μικρών μελετών. Η συμφωνία των αποτελεσμάτων μεταξύ της μετανάλυσης και των αντίστοιχών μεγάλων κλινικών μελετών ήταν ασθενής ($\kappa=0.35$). Σε ποια αποτελέσματα επομένως θα πρέπει να στηριζόμαστε; Οι περισσότεροι θα συμφωνήσουν ότι αν έχει πραγματοποιηθεί μια μεγάλη, υψηλής ποιότητας, τυχαιοποιημένη κλινική μελέτη, τότε τα αποτελέσματά της είναι ευρύτερα αποδεκτά στην επιστημονική κοινότητα. Το ερώτημα όμως προκύπτει αν τα μόνα διαθέσιμα στοιχεία έχουν προκύψει από κάποιες, μικρές σε έκταση, τυχαιοποιημένες μελέτες. Τότε στην περίπτωση αυτή η μετανάλυση είναι απαραίτητη για την διατύπωση ενός ολικού συμπεράσματος.

Συμπερασματικά, οι μετανάλυσεις προσφέρουν μια συστηματική και ποσοτική προσέγγιση στην ανασκόπηση των σημαντικών ερωτήσεων οι οποίες προκύπτουν από την εφαρμογή των θεραπευτικών μεθόδων. Εν τούτοις, αφθονούν οι παγίδες στην εκτέλεση των μετανάλυσεων οι οποίες σχετίζονται και με την ποιότητα των υπό εξέταση μελετών.

Βελτιώνοντας την ποιότητα των δημοσιευμένων κλινικών μελετών, βοηθάμε στην ενίσχυση της αξιοπιστίας των αποτελεσμάτων της μετανάλυσης. Επιπλέον μια προσεκτική ανάγνωση των δημοσιευμένων μεταanalύσεων καθώς και η ισορροπημένη αξιολόγησή τους, μπορούν να βοηθήσουν στην επίλυση αναπάντητων ερωτημάτων ή στη δημιουργία νέων σε σχέση με την αποτελεσματικότητα της θεραπείας. Η μετανάλυση είναι ένα σύνολο από κλασσικές στατιστικές τεχνικές, κατάλληλα προσαρμοσμένες, προκειμένου να ελεγχθεί η διαφορετικότητα των μελετών και στη συνέχεια να συνδυαστεί η πληροφορία που παρέχουν αυτές οι μελέτες, με στόχο την εκτίμηση της ολικής επίδρασης της θεραπείας. Δεν είναι όμως απλή διαδικασία εύρεσης του μέσου όρου, ούτε μπορεί να εφαρμοστεί σε όλες τις περιπτώσεις. Για την διεξαγωγή της επομένως, είναι αναγκαία η συνεργασία εξειδικευμένων ιατρών και βιοστατιστικών. Οι ιατροί είναι αυτοί που θα διατυπώσουν το ιατρικό πρόβλημα και θα επιλέξουν τις σχετικές μελέτες, ενώ οι βιοστατιστικοί είναι οι υπεύθυνοι για την σωστή εφαρμογή των στατιστικών μεθοδολογιών. Μια σωστή ομάδα επιστημόνων και μια καλά οργανωμένη μετανάλυση έχει πολλά να προσφέρει στον ιατρικό χώρο.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Ελληνική

Παπαϊωάννου, Τ. και Φερεντίνος, Κ. (1999). *Ιατρική Στατιστική και Στοιχεία Βιομαθηματικών*, Εκδόσεις Πανεπιστημίου Ιωαννίνων, Ιωάννινα.

Ξένη

- Agresti, A. [2002]. *Categorical Data Analysis* (2nd edn) Florida: Wiley
- Alderson P., Green S., Higgins J.P.T [2004] *Cochrane Reviewers' Handbook 4.2.2* The Cochrane Library, Issue 1
- Altman D.G [2000]. *Statistics in medical journals: some recent trends*. *Statistics in Medicine* 19: 3275-3289
- Bartlett, M.S. [1937]. *Properties of sufficiency and statistical test*. *Proceedings of the Royal Society of London, Series A*. 160: 268-282 (4,5)
- Begg C.B, Mazumdar M [1994]. *Operating characteristics of a rank correlation test for publication bias*. *Biometrics* 50 (4): 1088-1101
- Brockwell S.E, Gordon I.R [2001]. *A comparison of statistical methods for meta-analysis*. *Statistics in Medicine* 20: 825-840
- Brown H.K, Kempton R.A [1994] *The application of REML in clinical trials*. *Statistics in Medicine* 13: 1601-1617
- Chalmers TC, Smith H Jr, Blackburn B, Silverman B, Schroeder B, Reitman D [1981] *A method for assessing the quality of a randomized control trial*. *Controlled Clinical Trials* 2: 31-49
- Cochran W.G [1954]. *The combination of estimates from different experiments*. *Biometrics*, 10 (4): 101-129
- Collett, D. [1991]. *Modeling Binary Data* London: Chapman and Hall (5)
- Dear KBG, Begg CB [1992] *An approach for assessing publication bias prior to performing a meta-analysis*. *Statistical Science* 7: 237-245
- Deeks J.J, Altman D.G, Dooley G, Sackett D.L [1997]. *Choosing an appropriate dichotomous effect measure for meta-analysis: empirical evidence of the appropriateness of the odds ratio and relative risk*. (Abstract). *Controlled Clinical Trials* 18 (3S):84S-85S
- Deeks J [1999] *Statistical Methods Programmed in MetaView, version 4*. On behalf of the Statistical Methods Working Group of the Cochrane Collaboration.

- DerSimonian R, Laird N [1986] *Meta-Analysis in Clinical Trials*. *Controlled Clinical Trials* 7: 177-188
- Duval S, Tweedie R [2000]. *Trim and Fill: a simple funnel plot based method of testing and adjusting for publication bias in meta-analysis*. *Biometrics* 56: 455-463
- Egger M, Davey S.G, Schneider M, Minder C. [1997b]. *Bias in meta-analysis detected by a simple graphical test*. *British Medical Journal* 315 (8): 629-634
- Engels E., Schmid C.H, Terrin N. , Olkin I., Lau J. [2000]. *Heterogeneity and statistical significance in meta-analysis: an empirical study of 125 meta-analyses*. *Statistics in Medicine* 19: 1707-1728
- Fleiss J.L, [1981]. *Statistical Methods for Rates and Proportions* (2nd edn) New York: J Wiley
- Fleiss JL, Cooper H, Hedges LV. [1994]. *Measures of effect size for categorical data. The handbook of research synthesis*. New York:Russell Sage Foundation p. 245-260
- Galbraith R.F. [1988]. *A note on graphical presentation of estimated odds ratios from several trials*. *Statistics in Medicine* 7 (7): 889-894
- Greenland S, Salvan A [1990]. *Bias in the one-step method for pooling study results*. *Statistics in Medicine* 9: 247-252
- Greenland S. [1994] *Invited commentary: a critical look at some popular meta-analytic methods*. *Am J Epidemiol* 140: 290-296
- Hardy R.J, Thompson S.G [1988]. *Detecting and describing heterogeneity in meta-analysis*. *Statistics in Medicine* 17: 841-856
- Hedges LV [1984] *Estimation of effect size under nonrandom sampling: the effect of censoring studies yielding statistically insignificant mean differences*. *J Educ Stat* 9: 61-85
- Hedges LV [1992] *Modeling publication selection effects in meta-analysis*. *Statistical Science* 7: 246-255
- Higgins J.P, Thompson S.G. [2002]. *Quantifying heterogeneity in a meta-analysis*. *Statistics in Medicine* 21: 1539-1558
- Iyengar S, Greenhouse JB [1988] *Selection models and the file drawer problem*. *Statist Sci* 3:109-135
- Jeng G.T, Scott J.R, Burmeister L.E [1995]. *A comparison of meta-analysis results using literature vs. individual patient data*. *Journal of the American Medical Association* 274: 830-836
- L' Abbe K.A, Detsky A.S, O'Rourke K. [1987]. *Meta-analysis in clinical research*. *Annals of Internal Medicine* 107: 224-233
- Lau J., Ioannidis J, Schmid C.H, [1997]. *Quantitative synthesis in systematic reviews*. *Ann Intern Med* 127 (9): 820-826
- Leloirier J, Gregoire G, Benhaddad A, Lapierre J, Derderian F [1997]. *Discrepancies between meta-analyses and subsequent large randomized, controlled trials*. *The New England Journal of Medicine*. 337 (8): 536-542
- Li YZ, Shi L, Roth HD [1994] *The bias of the commonly used estimate of variance in meta-analysis*. *Commun Statist Theory Methods* 23: 1063-1085

- Light R.J, Pillemer D.B [1984]. *Summing Up: The Science of Reviewing Research*. Cambridge MA: Harvard University Press (8)
- Light R.J, Smith P.V [1971]. *Accumulating evidence: procedures for resolving contradictions among different research studies*. Harvard Educational Review. 41: 429-471
- Macaskill P, Walter S.D, Irwig L [2001]. *A comparison of methods to detect publication bias in meta-analysis*. Statistics in Medicine 20: 641-654
- McCullagh P., Nelder J.A., [1989]. *Generalized Linear Models* (2nd edn). London: Chapman and Hall (5, Appendix)
- Normand S.L [1999]. *Tutorial in Biostatistics. Meta-Analysis: Formulating, evaluating, combining and reporting*. Statistics in Medicine 18:321-359
- Olkin I., Sampson A. [1998]. *A Comparison Of Meta-Analysis Versus Analysis of Variance of Individual Patient Data*. Biometrics 54: 317-322
- Oxman A.P, Guyatt G.H [1992]. *A consumers guide to subgroup analyses*. Ann Intern Med 116: 78-84
- Patterson DV, Thompson R [1971]. *Recovery of inter-block information when the block sizes are unequal*. Biometrika 58: 545-554
- Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV [1977]. *Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II: Analysis and examples*. Br J Cancer 35: 1-39
- Robins J, Greenland S, Breslow NE [1986]. *A general estimator for the variance of the Mantel-Haenszel OR*. Am J Epidemiol 124: 719-723
- Rosenthal R [1979]. *The “file-drawer problem” and tolerance for null results*. Psychological Bulletin 86 (8): 638-641
- Sharp SJ, Thompson SG, Altman DG [1996]. *The relation between treatment benefit and underlying risk in meta-analysis*. BMJ 313: 735-738
- Steinberg K.K, Smith S.J, Stroup D.F, Olkin I., Lee N.C, Williamson G.D, Thacker S.B. [1997]. *A comparison of effect estimates from meta-analysis of summary data from published studies and from a meta-analysis using individual patient data for ovarian cancer studies*. American Journal of Epidemiology 145 (10): 917-925
- Sterne J.A.C, Egger M [2001]. *Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis*. Journal of Clinical Epidemiology 54: 1046-1055
- Sterne J.A.C, Egger M, Davey S.G [2001a]. *Investigating and dealing with publication and other biases*. In Egger, Smith and Altman (eds). Systematic Reviews in Health care: Meta-Analysis in Context (2nd edn) London: BMJ Books
- Sutton A.S, Duval S.J, Tweedie R.L, Abrams K.R, Jones D.R [2000] *Empirical assessment of affect of publication bias on meta-analyses*. BMJ 320: 1574-1577
- Taylor SJ, Tweedie RL [1998]. *Practical estimates of the effect of publication bias in meta-analysis* Aust Epidemiol 5: 14-17
- Thompson S.G, Sharp S.J. [1999]. *Explaining heterogeneity in meta-analysis: a comparison of methods*. Statistics in Medicine 18: 2693-2708

-
- Thompson S.G [1994]. *Why sources of heterogeneity in meta-analysis should be investigated*. British Medical Journal 309: 1351-1355
- Thompson S.G, Higgins J.P [2002]. *How should meta-regression analyses be undertaken and interpreted?* Statistics in Medicine 21: 1559-1573
- Turner R.M, Omar R.Z, Yang M., Goldstein H., Thompson S.G [2000]. *A multilevel model framework for meta-analysis of clinical trials with binary outcomes*. Statistics in Medicine 19: 3417-3432 (5)
- Whitehead A. [2002]. *Meta-Analysis of Controlled Clinical Trials*. England: J Wiley
- Yusuf S., Wittes J., Probstfield J., Tyroler H.A., [1991]. *Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials*. JAMA 266: 93-98
- Yusuf S., Peto R, Lewis J, Collins R, Sleight P. [1985]. *Beta blockade during and after myocardial infarction: an overview of the randomized trials*. Prog Cardiovasc Dis 27: 335-371