

ΕΦΑΡΜΟΣΜΕΝΗ ΙΑΤΡΙΚΗ ΕΡΕΥΝΑ APPLIED MEDICAL RESEARCH

Συστηματική ανασκόπηση και μετα-ανάλυση

1. Εισαγωγή
2. Συστηματική ανασκόπηση
 - 2.1. Διατύπωση επιστημονικής υπόθεσης
 - 2.2. Κριτήρια ένταξης και αποκλεισμού των μελετών
 - 2.3. Καθορισμός των μελετών
3. Μετα-ανάλυση
 - 3.1. Σφάλμα δημοσίευσης
 - 3.2. Ετερογένεια των μελετών
 - 3.3. Μαθηματική διαδικασία
4. Σύνοψη

1. ΕΙΣΑΓΩΓΗ

Η ανάγκη για έγκυρες και έγκαιρες αποφάσεις τόσο σε θέματα δημόσιας υγείας* όσο και στην καθημερινή κλινική πρακτική, καθώς και η ολοένα αυξανόμενη πληροφορία σχετικά με τις διάφορες επιστημονικές υποθέσεις, καθιστούν απαραίτητη τη σύνθεση των αποτελεσμάτων που προέρχονται από την πληθώρα των μελετών που διεξάγονται.¹⁻⁴ Για το σκοπό αυτόν, η πλέον αποδεκτή επιστημονική μέθοδος είναι η *συστηματική ανασκόπηση* (systematic review) σε συνδυασμό με την εφαρμογή της *μετα-ανάλυσης* (meta-analysis).

Η χρήση της *αφηγηματικής ανασκόπησης* (narrative review) έχει δεχθεί έντονη κριτική και έχει απορριφθεί από την πλειονότητα των ερευνητών, καθώς σε αρκετές περιπτώσεις οδήγησε στην εξαγωγή λανθασμένων συμπερασμάτων και στην εφαρμογή παρεμβάσεων με αρνητικές συνέπειες.^{2,4,5} Τα σημαντικότερα μειονεκτήματα των αφηγηματικών ανασκοπήσεων⁶ είναι (α) ο υποκειμενικός τρόπος με τον οποίο ο ερευνητής επιλέγει τις μελέτες που πρόκειται να συμπεριλάβει στην ανασκόπηση του, (β) η ανομοιογένεια στην ποιότητα και τον τρόπο λήψης των πρωτογενών δεδομένων, (γ) η μη ύπαρξη σε όλες τις μελέτες των

* Δημόσια υγεία ονομάζεται η φροντίδα υγείας που παρέχεται από το κράτος.

ΑΡΧΕΙΑ ΕΛΛΗΝΙΚΗΣ ΙΑΤΡΙΚΗΣ 2009, 26(6):826-841
ARCHIVES OF HELLENIC MEDICINE 2009, 26(6):826-841

Π. Γαλάνης

Εργαστήριο Οργάνωσης
και Αξιολόγησης Υπηρεσιών Υγείας,
Τμήμα Νοσηλευτικής, Πανεπιστήμιο
Αθηνών, Αθήνα

Systematic review
and meta-analysis

Abstract at the end of the article

Λέξεις ευρητηρίου

Μετα-ανάλυση
Μοντέλο σταθερών αποτελεσμάτων
Μοντέλο τυχαίων αποτελεσμάτων
Συστηματική ανασκόπηση
Σφάλμα δημοσίευσης

Υποβλήθηκε 19.12.2008
Εγκρίθηκε 12.1.2009

απαιτούμενων δεδομένων για την πραγματοποίηση της ανασκόπησης και (δ) η εσφαλμένη ερμηνεία της έκβασης. Έτσι, η συστηματική ανασκόπηση αποτελεί τη μοναδική αξιόπιστη επιλογή επίλυσης των προβλημάτων αυτών. Οι συστηματικές ανασκοπήσεις βεβαίως δεν εξουδετερώνουν πάντοτε τα μειονεκτήματα των αφηγηματικών ανασκοπήσεων, αλλά οπωσδήποτε μειώνουν την πιθανότητα σφάλματος και περιγράφουν με σαφήνεια τους περιορισμούς και τα σφάλματα που τυχόν υπάρχουν.^{7,8} Μια συστηματική ανασκόπηση, για παράδειγμα, που περιλαμβάνει μελέτες με μειωμένη στατιστική ισχύ, εξαιτίας της έλλειψης μελετών με αυξημένη στατιστική ισχύ, θα χαρακτηρίζεται και εκείνη από μειωμένη στατιστική ισχύ.

Σημειώνεται ότι σε ορισμένες περιπτώσεις η συστηματική ανασκόπηση και η μετα-ανάλυση χρησιμοποιούνται λανθασμένα ως ταυτόσημες έννοιες. Η συστηματική ανασκόπηση αποτελεί μια ανασκόπηση της βιβλιογραφίας σχετικά με μια συγκεκριμένη επιστημονική υπόθεση –π.χ. σχέση μεταξύ καπνισματικής συνήθειας και καρκίνου του πνεύμονα– και αποβλέπει στην αναγνώριση, την εκτίμηση και την επιλογή των καλύτερα μεθοδολογικά σχεδιασμένων μελετών. Θεωρητικά, οι συστηματικές ανασκοπήσεις προσφέρουν την πλέον αξιόπιστη πληροφορία σχετικά με μια επιστημονική υπόθεση και για το λόγο αυτόν είναι απαραίτητο οι επιστήμονες υγείας να γνωρίζουν τον τρόπο διεξαγωγής

τους, καθώς και τον τρόπο ερμηνείας των αποτελεσμάτων τους. Η μετα-ανάλυση αποτελεί μια μαθηματική διαδικασία που συνδυάζει στατιστικά τα αποτελέσματα των μελετών που επιλέχθηκαν έπειτα από τη συστηματική ανασκόπηση. Είναι προφανές ότι η εξαγωγή ασφαλών συμπερασμάτων με την εφαρμογή της μετα-ανάλυσης προϋποθέτει μια καλά σχεδιασμένη συστηματική ανασκόπηση της βιβλιογραφίας, έτσι ώστε να συμπεριληφθούν στη μετα-ανάλυση οι πλέον κατάλληλες μελέτες. Έτσι, τα συμπεράσματα της μετα-ανάλυσης είναι ασφαλή μόνον εφόσον έχει προηγηθεί ενδελεχής συστηματική ανασκόπηση όλων των μελετών, οι οποίες ενδεχομένως θα μπορούσαν να συμπεριληφθούν στη μετα-ανάλυση. Επιπλέον, η διεξαγωγή μόνο της συστηματικής ανασκόπησης χωρίς την εφαρμογή της μετα-ανάλυσης για τον υπολογισμό ενός συνδυαστικού αποτελέσματος με βάση τα ξεχωριστά αποτελέσματα των επιμέρους μελετών αποτελεί μια μη ολοκληρωμένη διαδικασία, καθώς δεν εξάγεται ένα συγκεντρωτικό αποτέλεσμα.

Στην παρούσα εργασία, προτείνεται η σαφής διάκριση ανάμεσα στη συστηματική ανασκόπηση και τη μετα-ανάλυση, με την πρώτη να αποτελεί τη θεωρητική διαδικασία καθορισμού με ορισμένα κριτήρια των καλύτερα μεθοδολογικά σχεδιασμένων μελετών σχετικά με μια συγκεκριμένη επιστημονική υπόθεση και τη δεύτερη να αποτελεί τη μαθηματική διαδικασία υπολογισμού ενός συγκεντρωτικού αποτελέσματος με βάση τα αποτελέσματα των μελετών, οι οποίες επιλέχθηκαν έπειτα από τη συστηματική ανασκόπηση της βιβλιογραφίας. Έτσι, η συστηματική ανασκόπηση και η μετα-ανάλυση αποτελούν το πρώτο και το δεύτερο βήμα, αντίστοιχα, μιας διαδικασίας συνδυασμού των αποτελεσμάτων ενός αριθμού μελετών αναφορικά με μια συγκεκριμένη επιστημονική υπόθεση, με σκοπό τον υπολογισμό με μεγαλύτερη ακρίβεια και εγκυρότητα –σε σχέση με κάθε επιμέρους μελέτη ξεχωριστά– ενός συγκεντρωτικού αποτελέσματος, το οποίο εκτιμά ουσιαστικά τη σχέση μεταξύ προσδιοριστή* και έκβασης.** Ουσιαστικά, η συστηματική ανασκόπηση και η μετα-ανάλυση αποτελούν δύο αλληλένδετες διαδικασίες και μόνον ο συνδυασμός τους μπορεί να οδηγήσει σε ασφαλή συμπεράσματα.

Σε ορισμένες περιπτώσεις, η μετα-ανάλυση αναφέρεται καταχρηστικά ως η διαδικασία που περιλαμβάνει τόσο τα βήματα μιας συστηματικής ανασκόπησης όσο και το στατιστικό συνδυασμό των αποτελεσμάτων των επιμέρους μελετών, ενώ σε ορισμένες άλλες περιπτώσεις αναφέρεται μόνο η εφαρμογή της συστηματικής ανασκόπησης, υπονοώντας όμως και την εφαρμογή της μετα-ανάλυσης. Είναι προτιμότερο πάντως να γίνεται διάκριση ανάμεσα στη συστηματική ανασκόπηση, ως της διαδικασίας αναζήτησης και επιλογής με ορισμένα κριτήρια των κατάλληλων μελετών, και στη μετα-ανάλυση, ως της στατιστικής διαδι-

κασίας υπολογισμού ενός συγκεντρωτικού αποτελέσματος με βάση τα αποτελέσματα των επιμέρους μελετών. Στην πράξη, η μετα-ανάλυση είναι μια μαθηματική διαδικασία που χρησιμοποιεί τα αποτελέσματα της συστηματικής ανασκόπησης, ενώ η εξαγωγή ασφαλών συμπερασμάτων μέσω της μετα-ανάλυσης είναι αδύνατη χωρίς προηγουμένως να έχει πραγματοποιηθεί η κατάλληλη συστηματική ανασκόπηση.

Η συστηματική ανασκόπηση και η μετα-ανάλυση είναι απαραίτητες πριν από τη διεξαγωγή μιας νέας μελέτης, έτσι ώστε να είναι δυνατόν να διαπιστωθεί εάν τα αποτελέσματα της μελέτης αυτής προσθέτουν πληροφορία σχετικά με μια συγκεκριμένη επιστημονική υπόθεση.^{4,12-14} Η μετα-ανάλυση είναι περισσότερο χρήσιμη όταν τα αποτελέσματα των διαφόρων μελετών είναι αντιφατικά μεταξύ τους και όταν ο αριθμός των συμμετεχόντων στις μελέτες είναι μικρός, καθώς στις περιπτώσεις αυτές ο συνδυασμός των μελετών αυξάνει τη στατιστική ισχύ.¹⁵⁻¹⁷ Η μετα-ανάλυση αποτελεί ουσιαστικά μια στατιστική ανάλυση στην οποία ως «δεδομένα» χρησιμοποιούνται τα αποτελέσματα των επιμέρους μελετών και όχι τα πρωτογενή δεδομένα των μελετών αυτών. Στην περίπτωση πάντως που είναι διαθέσιμα και τα πρωτογενή δεδομένα των μελετών, συνιστάται η εφαρμογή της μεθόδου της *συνένωσης* (pooling) για τον υπολογισμό ενός συγκεντρωτικού αποτελέσματος.^{18,19} Η συνένωση είναι δυνατή σε ελάχιστες μόνο περιπτώσεις εξαιτίας της αδυναμίας συγκέντρωσης των πρωτογενών δεδομένων, αλλά παρουσιάζει αρκετά πλεονεκτήματα, με σημαντικότερο τη δυνατότητα εξουδετέρωσης των συγχυτών.²⁰

* *Παράγοντας κινδύνου* (risk factor) ή *έκθεση* (exposure) ή *προσδιοριστής* (determinant), όπως τελικά επικράτησε να λέγεται σήμερα, είναι το χαρακτηριστικό –συγγενές, περιβαλλοντικό ή συμπεριφορικό– των ατόμων από το οποίο εξαρτάται (σχετίζεται ή συναρτάται) η συχνότητα εμφάνισης της μελετώμενης έκβασης (η οποία στην αιτιογνωστική επιδημιολογία είναι η εμφάνιση της πάθησης).⁹⁻¹¹ Ο προσδιοριστής της συχνότητας εμφάνισης μιας έκβασης περιλαμβάνει δύο κατηγορίες, την *ενδεικτική κατηγορία* (index category) και την *κατηγορία αναφοράς* (reference category). Η επιλογή της κατηγορίας αναφοράς του προσδιοριστή είναι μείζονος σημασίας στο σχεδιασμό της μελέτης. Είναι κρίσιμης σημασίας να γίνει αντιληπτό ότι η κατηγορία αναφοράς δεν είναι συμφυές χαρακτηριστικό του προσδιοριστή, αλλά αποτέλεσμα επιλογής του ερευνητή, άρα συνέπεια του σχεδιασμού της μελέτης. Από την επιλογή της κατηγορίας αναφοράς εξαρτάται η ερμηνεία των αποτελεσμάτων μιας μελέτης. Προσδιοριστής, π.χ., της συχνότητας εμφάνισης της νεφρικής πάθησης δεν είναι η αρτηριακή υπέρταση, αλλά η αρτηριακή πίεση. Η αρτηριακή υπέρταση είναι μια κατηγορία και συνήθως η ενδεικτική κατηγορία του προσδιοριστή, στην οποία μελετάται η συχνότητα εμφάνισης της νεφρικής πάθησης σε σχέση πάντοτε με τη συχνότητα εμφάνισής της στην κατηγορία αναφοράς, στην προκειμένη περίπτωση στην κατηγορία των ατόμων που δεν έχουν αρτηριακή υπέρταση.

** Ο όρος *έκβαση* χρησιμοποιείται (α) στην αιτιογνωστική επιδημιολογία για να δηλώσει την εμφάνιση ενός ανεπιθύμητου συμβάντος (πάθησης ή θανάτου), ενώ (β) στην προγνωστική επιδημιολογία για να δηλώσει την ίαση, το θάνατο ή την εμφάνιση καταλοίπων έπειτα από την ύπαρξη μιας πάθησης.

Η μετα-ανάλυση* στις επιστήμες υγείας, αρχικά, εφαρμόστηκε για την εκτίμηση της αποτελεσματικότητας θεραπευτικών παρεμβάσεων σε κλινικές δοκιμές.^{4,21–24} Η μετα-ανάλυση, για παράδειγμα, 33 κλινικών δοκιμών που εκτιμούσαν την αποτελεσματικότητα της ενδοφλέβιας χορήγησης στρεπτοκινάσης σε σχέση με τη χορήγηση ανενεργούς ουσίας (placebo**) σε πάσχοντες με έμφραγμα του μυοκαρδίου κατέδειξε την υπεροχή της στρεπτοκινάσης, ενώ μόνο οι 6 από τις 33 κλινικές δοκιμές είχαν καταλήξει σε στατιστικά σημαντική υπεροχή της στρεπτοκινάσης.²⁵ Σήμερα, η δημοσίευση κλινικών δοκιμών απαιτεί την εφαρμογή αυστηρών κριτηρίων, διευκολύνοντας έτσι σημαντικά την εφαρμογή της μετα-ανάλυσης.^{26,27}

Εξάλλου, στην περίπτωση της αιτιογνωστικής επιδημιολογίας, όπου δεν είναι δυνατή η πραγματοποίηση τυχαιοποιημένων μελετών,*** παρά μόνο η πραγματοποίηση μελετών «ασθενών-μαρτύρων» –και σπανιότερα μελετών παρακολούθησης– εφαρμόζονται όλο και πιο συχνά μετα-αναλύσεις για την ασφαλέστερη εκτίμηση της σχέσης μεταξύ προσδιοριστή και συχνότητας εμφάνισης μιας πάθησης.²⁹ Σημειώνεται, πάντως, ότι το πλεονέκτημα της τυχαιοποίησης που διαθέτουν οι κλινικές δοκιμές έναντι των μελετών «ασθενών-μαρτύρων» και των μελετών παρακολούθησης αντανακλάται και στην εφαρμογή της μετα-

ανάλυσης στα διάφορα αυτά είδη μελετών.^{30–32} Μολονότι η μετα-ανάλυση μελετών που αφορούν στην αιτιογνωστική επιδημιολογία δεν είναι πάντοτε δυνατή^{33,34} (και ιδιαίτερα στις περιπτώσεις όπου σκοπός είναι ο υπολογισμός ενός συγκεντρωτικού μέτρου σχέσης****), ο αριθμός των αντίστοιχων δημοσιεύσεων συνεχώς αυξάνεται, καθώς πριν από το 1992 είχαν δημοσιευτεί 678 μετα-αναλύσεις, ενώ μεταξύ των ετών 1992–1995 δημοσιεύτηκαν 525 και το 1996 δημοσιεύτηκαν 400.

2. ΣΥΣΤΗΜΑΤΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

2.1. Διατύπωση επιστημονικής υπόθεσης

Το πρώτο βήμα μιας συστηματικής ανασκόπησης αποτελεί η διατύπωση μιας σαφούς επιστημονικής υπόθεσης που αφορά στη σχέση μεταξύ του μελετώμενου προσδιοριστή και της συχνότητας εμφάνισης μιας έκβασης, λαμβάνοντας υπόψη και την ύπαρξη τυχόν συγχυτών ή ακόμη και τροποποιητών.^{4,6,17} Στην περίπτωση των κλινικών δοκιμών, η τυχαιοποίηση περιορίζει σημαντικά την ύπαρξη των συγχυτών –τόσο των γνωστών όσο και των άγνωστων– κάτι όμως που δεν συμβαίνει στην περίπτωση των μελετών που εφαρμόζονται στην αιτιογνωστική επιδημιολογία, οπότε και απαιτείται πιο προσεκτικός σχεδιασμός.

Το πλέον σημαντικό σε μια συστηματική ανασκόπηση είναι ο σωστός σχεδιασμός του πρωτοκόλλου της ανασκόπησης και έπειτα η αυστηρή εφαρμογή του. Είναι εξαιρετικής σημασίας ο καθορισμός με ακρίβεια του τρόπου μέτρησης τόσο του μελετώμενου προσδιοριστή όσο και της έκβασης.⁶

2.2. Κριτήρια ένταξης και αποκλεισμού των μελετών

Σε όλες τις μελέτες, ο μελετώμενος πληθυσμός ή, αλλιώς, το «δείγμα» των παρατηρήσεων καθορίζει σε σημαντικό βαθμό τη χρησιμότητα και την επιστημονική εγκυρότητα των αποτελεσμάτων. Σε μια συστηματική ανασκόπηση, το «δείγμα» των παρατηρήσεων αποτελείται από τις μελέτες –δημοσιευμένες και μη– ή, ορθότερα, τα αποτελέσματα των μελετών που, τελικά, περιλαμβάνονται στη μετα-ανάλυση. Στο πρωτόκολλο μιας συστηματικής ανασκόπησης κα-

* Η μετα-ανάλυση, ως η μαθηματική διαδικασία συνδυασμού των αποτελεσμάτων διαφόρων μελετών ως τον υπολογισμό ενός συγκεντρωτικού αποτελέσματος, εφαρμόστηκε για πρώτη φορά το 1904 από τον Karl Pearson σε μια προσπάθειά του να αντιμετωπίσει το πρόβλημα της μειωμένης στατιστικής ισχύος που παρουσίαζαν μελέτες με μικρό αριθμό συμμετεχόντων. Στις επιστήμες υγείας, η πρώτη μετα-ανάλυση δημοσιεύτηκε το 1955. Ο όρος μετα-ανάλυση (meta-analysis) εισήχθη το 1976 από τον Gene Glass, το ενδιαφέρον του οποίου εστιαζόταν στην έρευνα στο χώρο της εκπαίδευσης.²⁸

** Πλασέμπο (placebo): (α) Φάρμακο, το οποίο χορηγείται σε πάσχοντα με σκοπό περισσότερο την ψυχολογική του στήριξη και την ανακούφιση παρά τη θεραπεία της ίδιας του της πάθησης και (β) αδρανής ουσία, σε σύγκριση προς την οποία δοκιμάζεται η χρήση μιας άλλης (π.χ. φαρμάκου) σε ελεγχόμενες συνθήκες εργαστηρίου.

*** Οι επιδημιολογικές μελέτες διακρίνονται στις πειραματικές μελέτες ή, αλλιώς, τυχαιοποιημένες ελεγχόμενες δοκιμές, και στις μη πειραματικές μελέτες.^{9–11} Οι πειραματικές μελέτες διακρίνονται στις τυχαιοποιημένες μελέτες θεραπευτικής παρέμβασης ή, αλλιώς, κλινικές δοκιμές (στις οποίες τα μελετώμενα άτομα είναι πάσχοντα) και στις τυχαιοποιημένες μελέτες προληπτικής παρέμβασης, που περιλαμβάνουν τις δοκιμές πεδίου (στις οποίες η παρέμβαση καθορίζεται σε υγιή μελετώμενα άτομα) και τις δοκιμές παρέμβασης στην κοινότητα (στις οποίες η παρέμβαση καθορίζεται σε υγιείς ομάδες ατόμων και όχι ατομικά). Όταν οι επιδημιολογικές πειραματικές μελέτες δεν είναι εφικτές, τότε πραγματοποιούνται μη πειραματικές μελέτες, που περιλαμβάνουν τις οιοσδήποτε πειραματικές μελέτες, τις μελέτες παρακολούθησης (στις οποίες τα μελετώμενα άτομα ταξινομούνται σε ομάδες ανάλογα με το αν ανήκουν στην ενδεικτική κατηγορία ή κατηγορία αναφοράς του μελετώμενου προσδιοριστή), τις μελέτες «ασθενών-μαρτύρων» (στις οποίες τα μελετώμενα άτομα διακρίνονται σε ασθενείς ή, καλύτερα, πάσχοντες, και «μάρτυρες» ανάλογα με το αν πάσχουν ή όχι από τη μελετώμενη πάθηση), τις οικολογικές μελέτες, τις συγχρονικές μελέτες και τις μελέτες αναλογικής θνησιμότητας.

**** Η ποσοτικοποίηση της νοσηρότητας γίνεται με τα μέτρα συχνότητας, ενώ η μελέτη των σχέσεων που συνδέουν τα χαρακτηριστικά των ατόμων (προσδιοριστές) με τη συχνότητα εμφάνισης των παθήσεων πραγματοποιείται με τα μέτρα σχέσης.^{9–11} Τα κυριότερα μέτρα σχέσης είναι (α) η διαφορά των μέτρων συχνότητας, (β) ο λόγος των μέτρων συχνότητας, (γ) η σχετική περίσσεια των μέτρων συχνότητας, (δ) το αποδοτέο κλάσμα στους εκτεθειμένους, (ε) το αποδοτέο κλάσμα στο σύνολο του πληθυσμού, (στ) ο λόγος των συμπληρωματικών πιθανοτήτων (λόγος των odds) και (ζ) ο συντελεστής της λογιστικής παλινδρόμησης.

θορίζονται με σαφήνεια και στη συνέχεια εφαρμόζονται επακριβώς τα κριτήρια ένταξης και αποκλεισμού των διαφόρων μελετών, έτσι ώστε να προκύψουν, τελικά, οι πλέον κατάλληλες μελέτες, τα αποτελέσματα των οποίων θα χρησιμοποιηθούν στη μετα-ανάλυση.^{4,6,17} Τα κριτήρια που εφαρμόστηκαν πρέπει να είναι αντικειμενικά και επιστημονικά και όχι να αποβλέπουν στη διευκόλυνση των ερευνητών. Παραδείγματος χάρη, η απόφαση να μη συμπεριληφθούν στη συστηματική ανασκόπηση μελέτες που πραγματοποιήθηκαν πριν από μια συγκεκριμένη ημερομηνία πρέπει να στηρίζεται στην πληροφορία ότι τη συγκεκριμένη αυτή ημερομηνία μεταβλήθηκε η έως τότε καθιερωμένη θεραπευτική παρέμβαση, με συνέπεια να μην είναι εφικτή η σύγκριση των αποτελεσμάτων των παλαιότερων μελετών με εκείνα των πλέον πρόσφατων. Προφανώς, η απόφαση αποκλεισμού των μελετών που διεξήχθησαν πριν από μια συγκεκριμένη ημερομηνία δεν πρέπει να στηρίζεται στο γεγονός ότι οι μελέτες αυτές, εξαιτίας του ότι πραγματοποιήθηκαν παλαιότερα, δεν συμπεριλαμβάνονται στις ηλεκτρονικές βάσεις αναζήτησης μελετών στο διαδίκτυο.

Ένα ακόμη εξαιρετικά σημαντικό πρόβλημα στην αναζήτηση των μελετών κατά τη συστηματική ανασκόπηση αποτελούν οι μελέτες που για διάφορους λόγους δεν έχουν δημοσιευτεί σε περιοδικά. Στην κατηγορία αυτή ανήκουν κυρίως οι μελέτες που δεν δημοσιεύτηκαν επειδή κατέληξαν σε αρνητικές σχέσεις (μη δημοσίευση, π.χ., μελετών που κατέληξαν στο ότι η εφαρμογή μιας νέας θεραπευτικής παρέμβασης για την αντιμετώπιση του καρκίνου του πνεύμονα είχε ως αποτέλεσμα μεγαλύτερη θνητότητα σε σχέση με την καθιερωμένη θεραπευτική παρέμβαση), καθώς επίσης και μελέτες διδακτορικών διατριβών, μελέτες που παρουσιάζονται σε συνέδρια και μελέτες κρατικών ή ιδιωτικών οργανισμών που δεν δημοσιεύονται για διάφορους λόγους (συνήθως οικονομικούς ή λόγους διασφάλισης του απορρήτου).

Σε αρκετές περιπτώσεις, οι ερευνητές αποκλείουν τις μελέτες που δεν είναι δημοσιευμένες στην αγγλική γλώσσα εξαιτίας της ανικανότητας μετάφρασής τους, γεγονός που ενδεχομένως να εισάγει σημαντικό σφάλμα, ιδιαίτερα στην περίπτωση μελετών που δημοσιεύονται σε έγκυρα περιοδικά τα οποία δεν εκδίδονται στην αγγλική γλώσσα.³⁵⁻³⁷

Σημειώνεται ότι με την εφαρμογή αυστηρότερων κριτηρίων αποκλεισμού των μελετών επιτυγχάνεται ευκολότερα η ομοιογένεια των μελετών που, τελικά, συμπεριλαμβάνονται στη μετα-ανάλυση. Με την εφαρμογή, εξάλλου, ευρύτερων κριτηρίων ένταξης των μελετών αυξάνεται η πληροφορία, καθώς περισσότερες μελέτες περιλαμβάνονται, τελικά, στη μετα-ανάλυση αρκεί όμως η ετερογένεια των μελετών να

μην είναι μεγάλη. Πάντως, σε κάθε περίπτωση, τα κριτήρια ένταξης και αποκλεισμού των μελετών αποτελούν υποκειμενική απόφαση των ερευνητών, με την προϋπόθεση βέβαια ότι η απόφαση αυτή είναι ορθολογική και εξυπηρετεί τις επιδιώξεις της επιστήμης και όχι προσωπικά ή οικονομικά συμφέροντα. Ορισμένοι, εξάλλου, προτείνουν να εκτιμάται η ποιότητα των μελετών που τελικά περιλαμβάνονται στη μετα-ανάλυση, με σκοπό να προσδίδεται μεγαλύτερη βαρύτητα κατά τον υπολογισμό του συγκεντρωτικού αποτελέσματος στις μελέτες εκείνες που θεωρούνται καλύτερες. Επιπλέον, με τον τρόπο αυτόν μπορεί να εκτιμηθεί και ο βαθμός στον οποίο το μέτρο σχέσης που υπολογίζεται σε μια μελέτη επηρεάζεται από την ποιότητα της μελέτης.^{38,39} Η εφαρμογή, για παράδειγμα, του σχήματος της τυχαιοποίησης σε μια κλινική δοκιμή, που αποτελεί σημαντικό χαρακτηριστικό της ποιότητας της δοκιμής, σχετίζεται άμεσα με το υπολογιζόμενο μέτρο σχέσης.⁴⁰⁻⁴² Εντούτοις, η εκτίμηση της ποιότητας των μελετών με την εφαρμογή μιας ad hoc διαδικασίας μπορεί να οδηγήσει σε μειωμένη εγκυρότητα, ενώ συχνές είναι και οι περιπτώσεις μη ύπαρξης σχέσης ανάμεσα στο υπολογιζόμενο μέτρο σχέσης και την ποιότητα των μελετών.^{33,43,44}

2.3. Καθορισμός των μελετών

Ο σκοπός της διαδικασίας αναζήτησης των μελετών είναι, αρχικά, η αναγνώριση όλων των σχετικών μελετών –δημοσιευμένων και μη– που θα μπορούσαν να συμπεριληφθούν στη μετα-ανάλυση και κατόπιν ο καθορισμός των μελετών που θα συμπεριληφθούν τελικά στη μετα-ανάλυση. Είναι απαραίτητο να καθορίζεται αναλυτικά και με σαφήνεια η στρατηγική αναζήτησης των κατάλληλων μελετών, έτσι ώστε να είναι δυνατή τόσο η επανάληψη της μετα-ανάλυσης από άλλους ερευνητές όσο και η εκτίμηση της εξωτερικής εγκυρότητας (external validity). Διερευνώντας σε βάθος τη στρατηγική αναζήτησης των μελετών είναι δυνατόν να ερμηνευτούν τα διαφορετικά συμπεράσματα που προέκυψαν από διαφορετικές μετα-αναλύσεις σχετικά με μια συγκεκριμένη επιστημονική υπόθεση.⁴⁵ Το 1996, σε μια ανασκόπηση⁴⁶ που περιελάμβανε 103 μετα-αναλύσεις, βρέθηκε ότι η στρατηγική αναζήτησης στις περισσότερες μετα-αναλύσεις δεν περιγραφόταν επαρκώς, λιγότερες από τις μισές μετα-αναλύσεις περιέγραφαν με λεπτομέρεια τα πρωτογενή δεδομένα των μελετών και μόλις το 22% των μετα-αναλύσεων εκτιμούσε την ποιότητα των πρωτογενών δεδομένων.

Η βιβλιογραφική αναζήτηση απαιτείται να είναι συστηματική και αναλυτική, με τον ερευνητή να χρησιμοποιεί διάφορες πηγές πληροφορίας, όπως μηχανές αναζήτησης στο διαδίκτυο, γραπτές εργασίες, παραπομπές μελετών που

έχουν δημοσιευτεί, καθώς και μελέτες ακαδημαϊκών, κρατικών ή ιδιωτικών οργανισμών που για διάφορους λόγους δεν έχουν δημοσιευτεί.^{47,48} Η ολοκληρωμένη αναζήτηση δεν θα πρέπει να περιλαμβάνει μόνο τις μηχανές αναζήτησης στο διαδίκτυο, καθώς σε ορισμένες συστηματικές ανασκοπήσεις με την επιλογή αυτή αναγνωρίστηκε μόλις το 50% των μελετών που σχετίζονταν με μια συγκεκριμένη επιστημονική υπόθεση.² Εκτιμάται ότι το 25–50% των κλινικών δοκιμών που διεξάγονται δεν δημοσιεύονται, με αποτέλεσμα η έλλειψη της πληροφορίας που προέρχεται από τις μελέτες που δεν δημοσιεύονται να εισάγει σφάλμα στη μετα-ανάλυση και να περιορίζει την ακρίβεια της εκτίμησης του συγκεντρωτικού αποτελέσματος.⁴⁹

Η αναζήτηση και η εύρεση όλων των μελετών που θα έπρεπε να συμπεριληφθούν σε μια συστηματική ανασκόπηση είναι μια εξαιρετικά χρονοβόρα και επίπονη διαδικασία, ιδιαίτερα στην περίπτωση κατά την οποία ο αριθμός των συναφών μελετών που δεν έχουν δημοσιευτεί είναι μεγάλος. Η ύπαρξη των μηχανών αναζήτησης στο διαδίκτυο έχει διευκολύνει σημαντικά τη διαδικασία αυτή, χωρίς όμως να αποτελεί πανάκεια. Οι σημαντικότερες μηχανές αναζήτησης στο διαδίκτυο που αφορούν στις επιστήμες υγείας είναι οι εξής:

- **MEDLINE** (Medical Literature Analysis and Retrieval System Online)
Η βιβλιογραφική αυτή βάση αποτελεί το μεγαλύτερο κομμάτι της PubMed, της βιβλιογραφικής βάσης της National Library of Medicine (NLM) των ΗΠΑ, και καλύπτει τα επιστημονικά πεδία της Ιατρικής, της Οδοντιατρικής, της Νοσηλευτικής, της Κτηνιατρικής, των βασικών επιστημών κ.ά. Περιέχει βιβλιογραφικές αναφορές και περιλήψεις άρθρων από >4.800 περιοδικά που εκδίδονται στις ΗΠΑ και σε 70 ακόμη χώρες. Η βάση περιέχει περισσότερες από 18 εκατομμύρια βιβλιογραφικές αναφορές, που χρονολογούνται από το 1966 έως σήμερα. Μολονότι η κάλυψη του υλικού είναι διεθνής, η πλειοψηφία των αναφορών είναι στην αγγλική γλώσσα ή διαθέτουν περίληψη στα αγγλικά.
- **EMBASE** (The Excerpta Medica Database)
Η βάση αυτή περιλαμβάνει περισσότερες από 12 εκατομμύρια βιβλιογραφικές αναφορές από το 1974 έως σήμερα, ενώ ετησίως προστίθενται περίπου 600.000 αναφορές και περιλήψεις. Η EMBASE είναι ιδιαίτερα χρήσιμη στην περίπτωση μελετών που αφορούν σε φαρμακευτικά προϊόντα, καθώς περιλαμβάνει σχεδόν όλα τα περιοδικά που ειδικεύονται στο επιστημονικό αυτό πεδίο.
- **CANCERLIT** (Cancer Literature)
Η βάση αυτή αποτελεί δημιουργία του National Cancer

Institute των ΗΠΑ και περιλαμβάνει περίπου 2 εκατομμύρια αναφορές από το 1963 έως σήμερα. Η CANCERLIT περιλαμβάνει >3.500 περιοδικά που σχετίζονται κυρίως με τις νεοπλασματικές παθήσεις. Σημαντικό πλεονέκτημα της βάσης αυτής αποτελεί η ύπαρξη αναφορών που αφορούν σε βιβλία, πρακτικά συνεδρίων και σεμιναρίων, εργασίες κρατικών οργανισμών και διδακτορικές διατριβές.

- **TOXLIN** (Toxicology Literature Online)
Η βάση αυτή ανήκει στη βιβλιογραφική βάση της National Library of Medicine των ΗΠΑ και περιλαμβάνει περισσότερες από 3 εκατομμύρια βιβλιογραφικές αναφορές περιοδικών που σχετίζονται κατά κύριο λόγο με την τοξικολογία, τα φαρμακευτικά προϊόντα και διάφορες χημικές ουσίες.
- **DAI** (Dissertation Abstracts International)
Αποτελεί μια εξαιρετικά σημαντική βάση, καθώς περιλαμβάνει δεδομένα που δύσκολα μπορούν να βρεθούν και, πιο συγκεκριμένα, τις περιλήψεις του 95–98% των διδακτορικών διατριβών που έχουν πραγματοποιηθεί στις ΗΠΑ και τον Καναδά από το 1861 έως σήμερα. Από το 1976 γίνεται συστηματική προσπάθεια, έτσι ώστε στη βάση αυτή να συμπεριληφθούν και διδακτορικές διατριβές που διεξάγονται σε χώρες εκτός από τις ΗΠΑ και τον Καναδά, χωρίς όμως η προσπάθεια αυτή να έχει αποδώσει έως σήμερα τα αναμενόμενα αποτελέσματα.
- **Εθνικό Κέντρο Τεκμηρίωσης** (National Documentation Centre)
Πρόκειται για ελληνική βάση που είναι διαθέσιμη στην ηλεκτρονική διεύθυνση www.ekt.gr και η οποία περιλαμβάνει περίπου το 80% των ελληνικών διδακτορικών διατριβών που πραγματοποιήθηκαν από το 1985 έως σήμερα, καθώς και το 5% των αντίστοιχων διατριβών από το 1932–1985.

Εξαιρετικά σημαντική είναι η συμβολή της *Cochrane Collaboration** (διαθέσιμη στην ηλεκτρονική διεύθυνση www.cochrane.org), η οποία από το 1993 αποτελεί ένα διεθνή μη κερδοσκοπικό οργανισμό στην πραγματοποίηση συστηματικών ανασκοπήσεων και μετα-αναλύσεων ιδιαίτερα στην περίπτωση των κλινικών δοκιμών.^{50,51} Περισσότεροι από 6.000 ερευνητές παγκοσμίως συνεργάζονται σήμερα αφιλοκερδώς με την Cochrane Collaboration, στη βάση δεδομένων της οποίας υπάρχουν 2.893 συστηματικές ανασκοπήσεις και 1.646 ερευνητικά πρωτόκολλα που

* Η Cochrane Collaboration δημιουργήθηκε στην Οξφόρδη το 1993 και το όνομά της οφείλεται στο Βρετανό επιδημιολόγο Archibald Lemman Cochrane (1909–1988).

σχετίζονται κυρίως με κλινικές δοκιμές και την αποτελεσματικότητα θεραπευτικών παρεμβάσεων.

Σε αρκετές περιπτώσεις, κατά τη διεξαγωγή μιας συστηματικής ανασκόπησης οι ερευνητές επικοινωνούν με πιο εξειδικευμένους επιστήμονες υγείας σχετικά με μια συγκεκριμένη επιστημονική υπόθεση, έτσι ώστε να αυξήσουν την πληροφορία τους και να εξάγουν πιο αξιόπιστα συμπεράσματα. Στην περίπτωση, εξάλλου, κατά την οποία από μια μελέτη προέκυψαν δύο ή περισσότερες δημοσιεύσεις με παρόμοια συμπεράσματα, θα πρέπει στη συστηματική ανασκόπηση να περιλαμβάνεται μόνο η πιο πρόσφατη δημοσίευση.⁵²

Όταν η αναζήτηση των μελετών γίνεται ηλεκτρονικά μέσω του διαδικτύου, τότε θα πρέπει να αναφέρεται λεπτομερώς η στρατηγική, έτσι ώστε να είναι δυνατή η επανάληψή της από άλλους ερευνητές μελλοντικά. Η αναφορά της στρατηγικής αναζήτησης θα πρέπει να περιλαμβάνει τις λέξεις-κλειδιά που χρησιμοποιήθηκαν, το πεδίο αναζήτησης –π.χ., εάν οι λέξεις-κλειδιά βρίσκονται στον τίτλο, στην περίληψη ή σε ολόκληρο το άρθρο– και το λογισμικό (software) που ενδεχομένως χρησιμοποιήθηκε για την αναζήτηση, όπως για παράδειγμα το OVID.*

Σημειώνεται ότι ορισμένοι ερευνητές προτείνουν την ποιοτική εκτίμηση των μελετών που συμπεριλαμβάνονται, τελικά, σε μια μετα-ανάλυση, έτσι ώστε κατά τον υπολογισμό του συνδυαστικού αποτελέσματος της μετα-ανάλυσης να δίνεται μεγαλύτερη βαρύτητα στις καλύτερα σχεδιασμένες μελέτες.^{43,44} Επιπλέον, με την κατηγοριοποίηση αυτή των μελετών είναι δυνατόν να εκτιμηθεί ο βαθμός στον οποίο το μέγεθος της σχέσης μεταξύ προσδιοριστή και έκβασης εξαρτάται από την ποιότητα της εκάστοτε μελέτης, χωρίς όμως αυτό να ισχύει πάντοτε.^{33,38,39}

3. ΜΕΤΑ-ΑΝΑΛΥΣΗ

Το τελευταίο βήμα της συστηματικής ανασκόπησης αποτελεί ο καθορισμός των μελετών που πρόκειται να συμπεριληφθούν στη μετα-ανάλυση για τον υπολογισμό, με μαθηματικό τρόπο, ενός συνδυαστικού αποτελέσματος. Στη συνέχεια, εφαρμόζεται η μετα-ανάλυση για το συνδυασμό, με μαθηματικό τρόπο, των αποτελεσμάτων των μελετών που επιλέχθηκαν κατά τη συστηματική βιβλιογραφική ανασκόπηση, έτσι ώστε να εξαχθεί ένα συγκεντρωτικό αποτέλεσμα. Για παράδειγμα, εάν σε μια συστηματική

ανασκόπηση επιλέχθηκαν τρεις μελέτες με 40, 60 και 100 συμμετέχοντες σε καθεμιά, τότε η μετα-ανάλυση των τριών αυτών μελετών θα εκφράζει ουσιαστικά το αποτέλεσμα μιας μελέτης με 200 συμμετέχοντες. Με τον τρόπο αυτόν, η μετα-ανάλυση** οδηγεί στον υπολογισμό ενός συγκεντρωτικού αποτελέσματος με μεγαλύτερη ακρίβεια και εγκυρότητα απ' ό,τι κάθε μελέτη χωριστά. Η μετα-ανάλυση παρέχει την καλύτερη δυνατή ένδειξη (evidence) όταν περιλαμβάνει τυχαίοποιημένες ελεγχόμενες δοκιμές, ενώ παρουσιάζει μειωμένη εγκυρότητα στην περίπτωση μελετών παρακολούθησης και μελετών «ασθενών-μαρτύρων».

3.1. Σφάλμα δημοσίευσης

Πριν από την πραγματοποίηση της μετα-ανάλυσης απαιτείται η εκτίμηση του σφάλματος δημοσίευσης (publication bias), το οποίο μπορεί να οφείλεται στο γεγονός ότι:^{4,49,53–55}

- Τα περισσότερα περιοδικά δημοσιεύουν συχνότερα μελέτες οι οποίες καταλήγουν σε θετικά ευρήματα έναντι μελετών που καταλήγουν σε αρνητικά ευρήματα, δημοσιεύουν δηλαδή συχνότερα μελέτες που καταλήγουν στο ότι η ενδεικτική κατηγορία του μελετώμενου προσδιοριστή αυξάνει τη συχνότητα εμφάνισης της μελετώμενης έκβασης (π.χ. μελέτες που καταλήγουν στο συμπέρασμα ότι οι καπνιστές εμφανίζουν συχνότερα καρκίνο του πνεύμονα)
- Οι ερευνητές αποστέλλουν συχνότερα προς κρίση στα περιοδικά μελέτες που καταλήγουν σε θετικά ευρήματα έναντι μελετών που καταλήγουν σε αρνητικά ευρήματα
- Οι μελέτες που καταλήγουν σε θετικά ευρήματα δημοσιεύονται συχνότερα σε περιοδικά τα οποία εκδίδονται στην αγγλική γλώσσα, με αποτέλεσμα να έχουν μεγαλύτερη πιθανότητα να εντοπιστούν από τους ερευνητές που διεξάγουν μια συστηματική ανασκόπηση
- Οι μελέτες που καταλήγουν σε θετικά ευρήματα έχουν μεγαλύτερη πιθανότητα να εντοπιστούν κατά τη διαδικασία της συστηματικής ανασκόπησης
- Οι μελέτες που καταλήγουν σε θετικά ευρήματα έχουν

** Σημειώνεται ότι ο όρος μετα-ανάλυση μάλλον χρησιμοποιείται καταχρηστικά και ότι προτιμότερος θα ήταν ο όρος μετα-σύνθεση, καθώς ουσιαστικά πραγματοποιείται σύνθεση και όχι ανάλυση των αποτελεσμάτων ενός αριθμού μελετών για τον υπολογισμό ενός συγκεντρωτικού αποτελέσματος. Επιπλέον, ο όρος ανάλυση δεδομένων είναι εσφαλμένος εννοιολογικά, καθώς οι παρατηρήσεις δεν είναι δεδομένα, ενώ η επεξεργασία των παρατηρήσεων είναι σύνθεση και όχι ανάλυση. Στην ουσία, η ανάλυση των δεδομένων είναι το σύνολο της μαρτυρίας ή της ένδειξης που η εμπειρική έρευνα χρησιμοποιεί για τον έλεγχο της υπόθεσης.

* Το OVID είναι το χρησιμότερο ίσως λογισμικό αναζήτησης δεδομένων στις επιστήμες υγείας, περιλαμβάνοντας >1.700 περιοδικά και βιβλία, καθώς και 200 ηλεκτρονικές βάσεις δεδομένων. Το OVID δημιουργήθηκε το 1984 στις ΗΠΑ, έχοντας σήμερα περίπου 13 εκατομμύρια χρήστες.

μεγαλύτερη πιθανότητα να δημοσιευτούν σε περισσότερα από ένα περιοδικά, με αποτέλεσμα να έχουν και μεγαλύτερη πιθανότητα να συμπεριληφθούν σε μια συστηματική ανασκόπηση

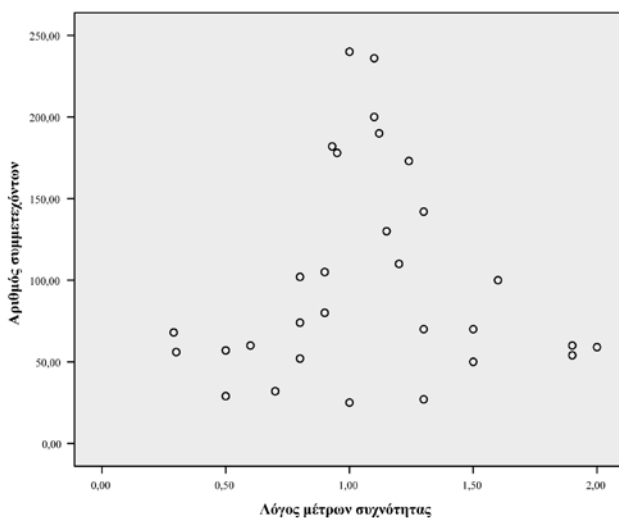
- Ορισμένες μελέτες δεν αποστέλλονται ποτέ για δημοσίευση και ιδιαίτερα εκείνες που δεν καταλήγουν σε θετικά ευρήματα, εκείνες που χρηματοδοτούνται από φαρμακευτικές εταιρείες και δεν καταλήγουν στα επιθυμητά αποτελέσματα και εκείνες που αφορούν σε μεταπτυχιακές ή διδακτορικές εργασίες.

Η ύπαρξη σφάλματος δημοσίευσης αποτελεί συστηματικό σφάλμα που μειώνει σημαντικά την εγκυρότητα μιας μετα-ανάλυσης και για το λόγο αυτόν απαιτείται ο εντοπισμός του. Ο πλέον συνηθισμένος τρόπος εκτίμησης του σφάλματος δημοσίευσης είναι ο σχεδιασμός ενός διαγράμματος που είναι γνωστό ως *διάγραμμα «χωνί»*⁵⁶ (funnel plot) (εικ. 1). Στο διάγραμμα αυτό, στον κάθετο άξονα αντιστοιχεί ο αριθμός των συμμετεχόντων στις διάφορες μελέτες που συμπεριλαμβάνονται στη μετα-ανάλυση, ενώ στον οριζόντιο άξονα αντιστοιχεί το μέτρο σχέσης που υπολογίζεται στις μελέτες (διαφορά ή λόγος μέτρων συχνότητας, λόγος συμπληρωματικών πιθανοτήτων κ.ά.). Όταν δεν υπάρχει σφάλμα δημοσίευσης, τότε τα σημεία σε ένα διάγραμμα «χωνί» που αντιστοιχούν στις διάφορες μελέτες δημιουργούν την οπτική απεικόνιση ενός συμμετρικού ανεστραμμένου κώνου ή χωνιού, όπως φαίνεται στην εικόνα 1. Στην περίπτωση αυτή, της απουσίας δηλαδή σφάλματος δημοσίευσης, οι μελέτες με μεγαλύτερο αριθμό συμμετεχόντων βρίσκονται στην κορυφή του ανεστραμμένου χωνιού και χαρακτηρίζονται από μεγαλύτερη ακρίβεια, με τις τιμές του εκτιμώμενου μέτρου σχέσης να

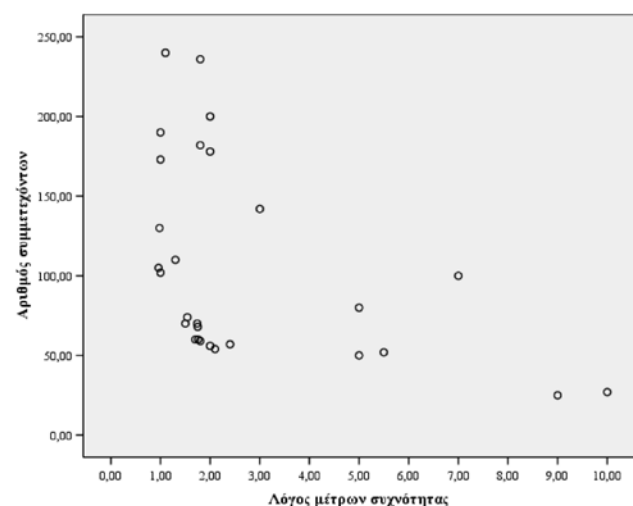
μην παρουσιάζουν μεγάλη διασπορά, ενώ οι μελέτες με μικρότερο αριθμό συμμετεχόντων βρίσκονται στο κάτω μέρος του ανεστραμμένου χωνιού και χαρακτηρίζονται από μικρότερη ακρίβεια, με τις τιμές του μέτρου σχέσης να παρουσιάζουν μεγαλύτερη διασπορά.

Η ύπαρξη ασυμμετρίας, εξάλλου, στο διάγραμμα «χωνί» υποδηλώνει την ύπαρξη σφάλματος δημοσίευσης.³⁷ Η περίπτωση αυτή αφορά συνήθως στο κάτω μέρος του χωνιού, όπου μελέτες με μικρό αριθμό συμμετεχόντων δεν παρουσιάζουν μεγάλη διασπορά στο εκτιμώμενο μέτρο σχέσης. Αυτό οφείλεται στο γεγονός ότι δημοσιεύονται συχνότερα μελέτες με μικρό αριθμό συμμετεχόντων που καταλήγουν σε θετικά ευρήματα σε σχέση με εκείνες που καταλήγουν σε αρνητικά ευρήματα. Στην εικόνα 2 φαίνεται ένα παράδειγμα ύπαρξης σφάλματος δημοσίευσης. Πιο συγκεκριμένα, στο διάγραμμα «χωνί» της εικόνας 2 υπάρχει ασυμμετρία, καθώς οι μελέτες με μικρό αριθμό συμμετεχόντων κατέληξαν όλες σε θετική σχέση μεταξύ προσδιοριστή και έκβασης, όπως υποδηλώνεται από την εκτίμηση του μέτρου σχέσης, που ήταν ο λόγος μέτρων συχνότητας. Έτσι, δεν υπάρχουν μελέτες με μικρό αριθμό συμμετεχόντων στις οποίες η ενδεικτική κατηγορία του προσδιοριστή να σχετίζεται με μικρότερη συχνότητα εμφάνισης της μελετώμενης έκβασης σε σχέση με την κατηγορία αναφοράς.

Ακόμη και στην περίπτωση όπου η συστηματική ανασκόπηση περιλαμβάνει ενδελεχή βιβλιογραφική ανασκόπηση, είναι απαραίτητο να κατασκευαστεί το διάγραμμα «χωνί», έτσι ώστε να εντοπιστεί η τυχόν ύπαρξη σφάλματος δημοσίευσης.⁵⁷ Μάλιστα, η εκτίμηση του σφάλματος δημοσίευσης μπορεί να γίνει και πιο αξιόπιστα, σε σχέση



Εικόνα 1. Διάγραμμα «χωνί», με την ύπαρξη συμμετρίας να υποδηλώνει την απουσία σφάλματος δημοσίευσης.



Εικόνα 2. Διάγραμμα «χωνί», με την ύπαρξη ασυμμετρίας να υποδηλώνει την παρουσία σφάλματος δημοσίευσης.

με το διάγραμμα «χωνί» και με ορισμένους στατιστικούς ελέγχους, με συνηθέστερους το συντελεστή συσχέτισης διατάξεων του Kendall (Kendall's tau rank correlation coefficient),⁵⁸ ο οποίος εφαρμόζεται χωρίς περιορισμούς αλλά εμφανίζει μειωμένη στατιστική ισχύ, και το συντελεστή συσχέτισης διατάξεων του Spearman (Spearman's rank correlation coefficient),⁵⁹ ο οποίος επίσης εφαρμόζεται χωρίς περιορισμούς και επιπλέον είναι απλούστερος στον υπολογισμό. Τέλος, η ανάλυση ευαισθησίας (sensitivity analysis) αποτελεί την καλύτερη ίσως στατιστική μέθοδο για την εκτίμηση του σφάλματος δημοσίευσης.

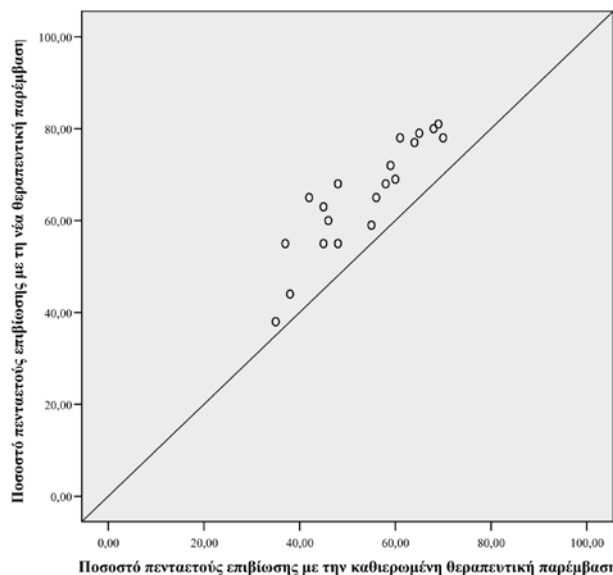
3.2. Ετερογένεια των μελετών

Η μετα-ανάλυση χρησιμοποιείται για τον υπολογισμό ενός συγκεντρωτικού αποτελέσματος με βάση τα αποτελέσματα ενός αριθμού παρόμοιων μελετών που προέκυψαν έπειτα από τη συστηματική ανασκόπηση. Η ακρίβεια και η εγκυρότητα μιας μετα-ανάλυσης εξαρτώνται σημαντικά από το βαθμό στον οποίο οι επιμέρους μελέτες είναι αρκετά ομοιογενείς μεταξύ τους, έτσι ώστε τα αποτελέσματά τους να μπορούν να συνδυαστούν για τον υπολογισμό ενός συγκεντρωτικού αποτελέσματος. Έτσι, πρέπει να υπάρχει ομοιογένεια στο μεθοδολογικό σχεδιασμό, στους μελετώμενους πληθυσμούς, στον τρόπο μέτρησης του προσδιοριστή και της έκβασης, στις μεθόδους εξουδετέρωσης των συγχυτών κ.λπ. Σε κάθε περίπτωση, βέβαια, τα αποτελέσματα των επιμέρους μελετών είναι λογικό να παρουσιάζουν μια ορισμένη μεταβλητότητα που οφείλεται στην τύχη. Όταν όμως τα αποτελέσματα των επιμέρους μελετών που πρόκειται να συμπεριληφθούν στη μετα-ανάλυση παρουσιάζουν μεγαλύτερη ετερογένεια από εκείνη που αναμένεται εκ τύχης, τότε ο υπολογισμός ενός μόνο συγκεντρωτικού αποτελέσματος μπορεί να οδηγήσει σε εσφαλμένα συμπεράσματα. Στην περίπτωση αυτή, συνιστάται η εφαρμογή της διαστρωματικής ανάλυσης ή της ανάλυσης παλινδρόμησης με τη χρήση των διαφόρων χαρακτηριστικών του μεθοδολογικού σχεδιασμού των μελετών ως προβλεπτικών παραγόντων για την εκτίμηση του βαθμού στον οποίο τα χαρακτηριστικά αυτά επηρεάζουν τη σχέση μεταξύ προσδιοριστή και έκβασης.⁶⁰ Οι μελέτες των οποίων τα αποτελέσματα διαφέρουν σημαντικά από τα αποτελέσματα της πλειονότητας των μελετών δεν πρέπει να απορρίπτονται απλά και μόνο εξαιτίας της ασυμφωνίας αυτής, αλλά πρέπει να εξετάζονται διεξοδικά τα διάφορα χαρακτηριστικά του μεθοδολογικού σχεδιασμού ή της ανάλυσης των δεδομένων που μπορεί να οδήγησαν στην ασυμφωνία αυτή.⁶¹

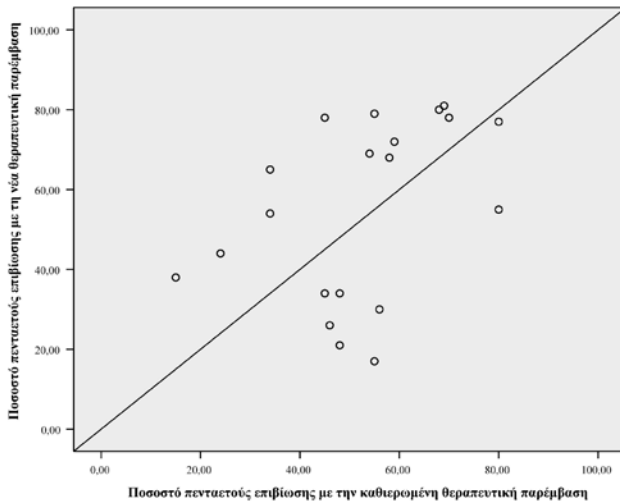
Η εκτίμηση της ομοιογένειας των αποτελεσμάτων των μελετών που πρόκειται να συμπεριληφθούν στη μετα-

ανάλυση μπορεί να πραγματοποιηθεί είτε με το *διάγραμμα L'Abbé* (L'Abbé plot) είτε με την εφαρμογή των κατάλληλων στατιστικών ελέγχων. Στο διάγραμμα L'Abbé, στον κάθετο άξονα αντιστοιχεί το μέτρο συχνότητας για εκείνους που ανήκουν στην ενδεικτική κατηγορία του μελετώμενου προσδιοριστή, ενώ στον οριζόντιο άξονα αντιστοιχεί το μέτρο συχνότητας για εκείνους που ανήκουν στην κατηγορία αναφοράς του προσδιοριστή. Διεξάγεται μια μετα-ανάλυση, π.χ., για την εκτίμηση της αποτελεσματικότητας μιας νέας θεραπευτικής παρέμβασης για την αντιμετώπιση του καρκίνου του μαστού. Το μέτρο συχνότητας που υπολογίζεται είναι το ποσοστό πενταετούς επιβίωσης. Στην εικόνα 3 φαίνεται το διάγραμμα L'Abbé της μετα-ανάλυσης αυτής. Η διαγώνιος γραμμή που φέρεται από την αρχή των αξόνων σε γωνία 45° αποτελεί το όριο της αποτελεσματικότητας της νέας θεραπευτικής παρέμβασης, με τις τιμές που βρίσκονται πάνω από τη διαγώνιο αυτή να εκφράζουν την υπεροχή της νέας θεραπευτικής παρέμβασης έναντι της καθιερωμένης και τις τιμές που βρίσκονται κάτω από τη διαγώνιο να εκφράζουν την υπεροχή της καθιερωμένης θεραπευτικής παρέμβασης. Όσο πιο συμπαγές είναι το διάγραμμα L'Abbé, όσο πιο μεγάλη συγκέντρωση δηλαδή παρουσιάζουν οι διάφορες τιμές, τόσο πιο ομοιογενείς είναι οι μελέτες από τις οποίες προέκυψαν τα δεδομένα του διαγράμματος. Στην εικόνα 3 οι τιμές του διαγράμματος παρουσιάζουν μεγάλη συγκέντρωση, η οποία υποδηλώνει ομοιογένεια των μελετών, ενώ το αντίστροφο συμβαίνει στην εικόνα 4.

Η ερμηνεία του διαγράμματος L'Abbé εναπόκειται στην



Εικόνα 3. Διάγραμμα L'Abbé, με τις τιμές που αντιστοιχούν στις 20 μελέτες να παρουσιάζουν μεγάλη συγκέντρωση, γεγονός που υποδηλώνει την ύπαρξη ομοιογένειας.



Εικόνα 4. Διάγραμμα L'Abbé, με τις τιμές που αντιστοιχούν στις 20 μελέτες να παρουσιάζουν μικρή συκέντρωση, γεγονός που υποδηλώνει την ύπαρξη ετερογένειας.

υποκειμενική κρίση των ερευνητών και απαιτεί ιδιαίτερη εμπειρία για την αποφυγή λανθασμένων συμπερασμάτων. Για το λόγο αυτόν, το διάγραμμα L'Abbé χρησιμοποιείται επικουρικά για την εκτίμηση της ομοιογένειας των μελετών και τα συμπεράσματα στηρίζονται κυρίως στην εφαρμογή των κατάλληλων στατιστικών ελέγχων. Στην περίπτωση αυτή, συνήθως προτιμάται η χρήση του στατιστικού ελέγχου Q των DerSimonian και Laird,^{4,62,63} που παρουσιάζει τη μεγαλύτερη εγκυρότητα και στατιστική ισχύ και επιπλέον υπολογίζεται απλούστερα. Η μηδενική υπόθεση που ελέγχεται είναι ότι οι μελέτες είναι ομοιογενείς μεταξύ τους, οπότε εάν το παρατηρούμενο επίπεδο της στατιστικής σημαντικότητας (τιμή p) είναι μεγαλύτερο από το προκαθορισμένο επίπεδο της στατιστικής σημαντικότητας (τιμή α), τότε δεν απορρίπτεται η μηδενική υπόθεση και, επομένως, υπάρχει ομοιογένεια μεταξύ των μελετών. Αντίθετα, εάν η τιμή p που προκύπτει από την εφαρμογή του στατιστικού ελέγχου είναι μικρότερη από την τιμή α , τότε απορρίπτεται η μηδενική υπόθεση και υπάρχει ετερογένεια μεταξύ των μελετών. Η εκτίμηση της ύπαρξης ετερογένειας είναι εξαιρετικής σημασίας, καθώς καθορίζει τη μαθηματική μέθοδο που πρόκειται να εφαρμοστεί για τον υπολογισμό του συγκεντρωτικού αποτελέσματος.

3.3. Μαθηματική διαδικασία

Η εκτίμηση του σφάλματος δημοσίευσης και της ομοιογένειας μεταξύ των μελετών αποτελούν απαραίτητες προϋποθέσεις για την εφαρμογή της κατάλληλης μαθηματικής μεθόδου, με σκοπό τον υπολογισμό του συγκεντρωτικού αποτελέσματος της μετα-ανάλυσης. Εάν διαπιστωθεί ομοι-

ογένεια μεταξύ των μελετών, τότε εφαρμόζεται το μοντέλο των σταθερών αποτελεσμάτων (fixed-effects model),^{4,64} ενώ εάν διαπιστωθεί ετερογένεια μεταξύ των μελετών, τότε εφαρμόζεται το μοντέλο των τυχαίων αποτελεσμάτων (random-effects model).^{4,65-67} Στην περίπτωση της ύπαρξης ομοιογένειας, η επιλογή του μοντέλου δεν επηρεάζει τα αποτελέσματα.

Όταν εφαρμόζεται το μοντέλο των σταθερών αποτελεσμάτων, ισχύει η υπόθεση ότι το αποτέλεσμα του προσδιοριστή είναι σταθερό στις διάφορες μελέτες και ότι η μεταβλητότητα του αποτελέσματος οφείλεται στην τυχαία μεταβλητότητα που παρουσιάζει η κάθε μελέτη εξαιτίας του γεγονότος ότι λαμβάνεται ένα «δείγμα» από τον πληθυσμό-πηγή. Επιπλέον, όταν εφαρμόζεται το μοντέλο των σταθερών αποτελεσμάτων, τότε τα συμπεράσματα της μετα-ανάλυσης είναι έγκυρα μόνο για τις μελέτες που συμπεριελήφθησαν στην ανάλυση αυτή και δεν μπορούν να γενικευτούν για όλες τις παρόμοιες μελέτες.

Στην περίπτωση του μοντέλου των τυχαίων αποτελεσμάτων, η μεταβλητότητα του αποτελέσματος του προσδιοριστή οφείλεται τόσο στη μεταβλητότητα που παρουσιάζει η κάθε μελέτη εξαιτίας του ότι χρησιμοποιούνται «δείγματα» και όχι πληθυσμοί-πηγή, όσο και στη μεταβλητότητα μεταξύ των διαφόρων μελετών.⁶⁵ Στο μοντέλο αυτό, ισχύει η υπόθεση ότι οι μελέτες που επιλέχθηκαν να συμπεριληφθούν στη μετα-ανάλυση αποτελούν ένα τυχαίο «δείγμα» όλων των παρόμοιων μελετών, οπότε είναι δυνατή η γενίκευση των αποτελεσμάτων.⁶⁶

Για τον υπολογισμό του συγκεντρωτικού αποτελέσματος στην περίπτωση του μοντέλου των σταθερών αποτελεσμάτων εφαρμόζεται η μέθοδος των Mantel και Haenszel, η μέθοδος του Peto ή η μέθοδος της αντίστροφης διασποράς (inverse variance method), ενώ στην περίπτωση του μοντέλου των τυχαίων αποτελεσμάτων εφαρμόζεται η μέθοδος των DerSimonian και Laird.⁴ Όπως προαναφέρθηκε, όταν υπάρχει ομοιογένεια μεταξύ των μελετών, τότε η επιλογή του μοντέλου δεν επηρεάζει το συγκεντρωτικό αποτέλεσμα. Σε όλες τις μεθόδους σταθμίζεται η βαρύτητα των μελετών στον υπολογισμό του συγκεντρωτικού αποτελέσματος, με εκείνες που έχουν μεγαλύτερο αριθμό συμμετεχόντων και μικρότερη τυχαία μεταβλητότητα να έχουν μεγαλύτερη βαρύτητα. Η μέθοδος των Mantel και Haenszel και η μέθοδος της αντίστροφης διασποράς χρησιμοποιούνται συχνότερα, ενώ η μέθοδος του Peto είναι απλούστερη στους υπολογισμούς, ιδιαίτερα των διαστημάτων εμπιστοσύνης, αλλά η εφαρμογή της απαιτεί την ύπαρξη ορισμένων προϋποθέσεων. Πιο συγκεκριμένα, η μέθοδος του Peto μπορεί να εφαρμοστεί μόνον όταν (α) το μέτρο σχέσης που υπολογίζεται είναι ο λόγος των συμπληρωματικών πιθανοτήτων, (β) η μετα-ανάλυση αφορά σε κλινικές δοκιμές, (γ) ο αριθμός

των περιπτώσεων έκβασης είναι σχετικά μεγάλος και (δ) ο αριθμός των συμμετεχόντων που ανήκουν στην ενδεικτική κατηγορία του μελετώμενου προσδιοριστή είναι περίπου ίσος με τον αριθμό των συμμετεχόντων στην κατηγορία αναφοράς του προσδιοριστή.

Ως παράδειγμα αναφέρεται μια υποθετική μετα-ανάλυση 7 κλινικών δοκιμών για τη σύγκριση της αποτελεσματικότητας του ενεργοποιητή του ιστικού προδρόμου της πλασμίνης έναντι της στρεπτοκινάσης* για την πρόληψη του θανάτου έπειτα από οξύ έμφραγμα του μυοκαρδίου. Στη μετα-ανάλυση αυτή, η μελετώμενη έκβαση ήταν ο θάνατος εξαιτίας οξέος εμφράγματος του μυοκαρδίου, ο μελετώμενος προσδιοριστής ήταν η θεραπευτική αγωγή –στρεπτοκινάση έναντι του ενεργοποιητή του ιστικού προδρόμου της πλασμίνης– και το υπολογιζόμενο μέτρο συχνότητας ήταν η θνητότητα** σε ένα έτος. Το μέτρο σχέσης που χρησιμοποιήθηκε για την εφαρμογή της μετα-

ανάλυσης ήταν ο λόγος των θνητοτήτων στους εμφραγματίες που ελάμβαναν στρεπτοκινάση σε σχέση με τους εμφραγματίες που ελάμβαναν ενεργοποιητή του ιστικού προδρόμου της πλασμίνης. Έτσι, η ενδεικτική κατηγορία του προσδιοριστή ήταν η ομάδα των πασχόντων που ελάμβαναν στρεπτοκινάση, ενώ η κατηγορία αναφοράς ήταν η ομάδα των πασχόντων που ελάμβαναν ενεργοποιητή του ιστικού προδρόμου της πλασμίνης.

Στον πίνακα 1 φαίνονται τα υποθετικά δεδομένα για τις 7 κλινικές δοκιμές που συμπεριελήφθησαν στη μετα-ανάλυση. Πραγματοποιώντας το στατιστικό έλεγχο Q των DerSimonian και Laird για τον έλεγχο της ομοιογένειας μεταξύ των μελετών προέκυψε ότι $p=0,87$, γεγονός που υποδηλώνει την ύπαρξη ομοιογένειας, οπότε είναι δυνατή η εφαρμογή του μοντέλου των σταθερών αποτελεσμάτων για τον υπολογισμό του συγκεντρωτικού αποτελέσματος. Στον πίνακα 2 φαίνονται τα «βάρη» με τα οποία σταθμίζονται οι

Πίνακας 1. Υποθετικά δεδομένα 7 κλινικών δοκιμών για τη σύγκριση της αποτελεσματικότητας της στρεπτοκινάσης έναντι του ενεργοποιητή του ιστικού προδρόμου της πλασμίνης για την πρόληψη του θανάτου έπειτα από οξύ έμφραγμα του μυοκαρδίου.

Μελέτη	Ενδεικτική κατηγορία (στρεπτοκινάση)		Κατηγορία αναφοράς (ενεργοποιητής)		Λόγος θνητοτήτων
	Πάσχοντες (σύνολο)	Θάνατοι (σε 1 έτος)	Πάσχοντες (σύνολο)	Θάνατοι (σε 1 έτος)	
A	100	10	100	5	2
B	80	5	81	3	1,6875
Γ	45	3	44	2	1,4667
Δ	250	30	248	10	2,976
E	52	4	50	3	1,2821
ΣΤ	40	5	40	2	2,5
Z	62	5	61	4	1,2298

Πίνακας 2. «Βάρη» με τα οποία σταθμίστηκαν οι 7 κλινικές δοκιμές του πίνακα 1 τόσο με την εφαρμογή της μεθόδου των Mantel και Haenszel όσο και με την εφαρμογή της μεθόδου της αντίστροφης διασποράς.

Μελέτη	Λόγος θνητοτήτων	95% διάστημα εμπιστοσύνης	Τιμή p	«Βάρη» (%) (μέθοδος Mantel-Haenszel)	«Βάρη» (%) (μέθοδος αντίστροφης διασποράς)
A	2	0,71–5,64	0,191	17	17
B	1,6875	0,42–6,83	0,463	10	10
Γ	1,4667	0,26–8,36	0,666	7	6
Δ	2,976	1,49–5,96	0,002	34	39
E	1,2821	0,31–5,44	0,736	11	9
ΣΤ	2,5	0,52–12,14	0,256	7	7
Z	1,2298	0,35–4,36	0,748	14	12

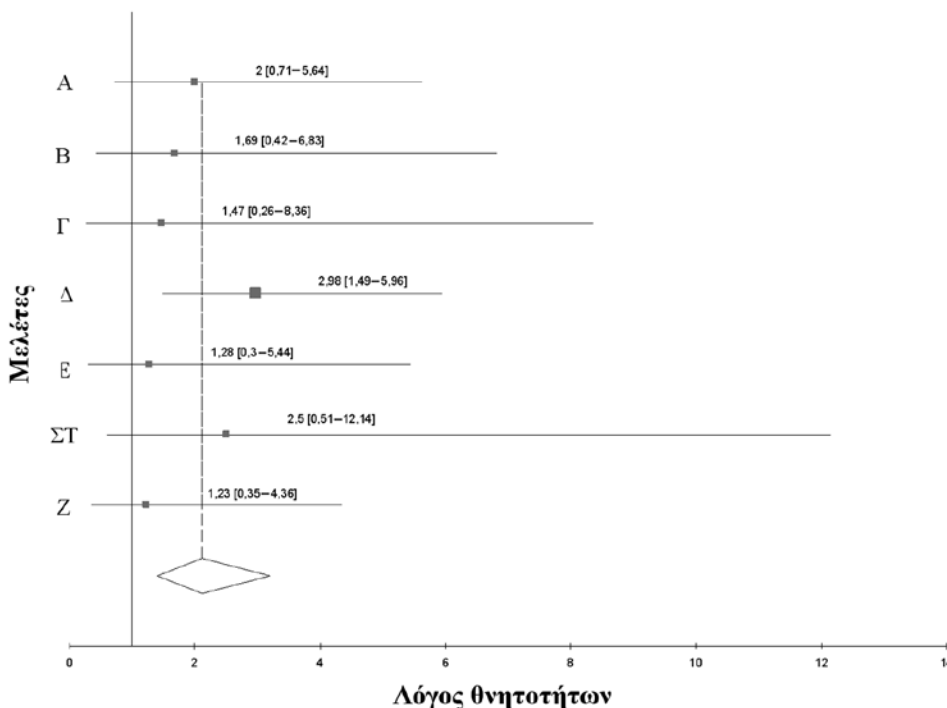
* Ο ενεργοποιητής του ιστικού προδρόμου της πλασμίνης κοστίζει αρκετά περισσότερο από τη στρεπτοκινάση, αλλά παρουσιάζει δύο σημαντικά πλεονεκτήματα έναντι αυτής. Πιο συγκεκριμένα, η στρεπτοκινάση είναι μια πρωτεΐνη που δεν προέρχεται από τον ανθρώπινο οργανισμό και για το λόγο αυτόν προκαλεί συχνότερα αλλεργικές αντιδράσεις και επιπλέον οδηγεί συχνότερα σε αιμορραγία σε σχέση με τον ενεργοποιητή του ιστικού προδρόμου της πλασμίνης.

** Η θνητότητα (fatality) αποτελεί μέτρο συχνότητας των παθήσεων και υπολογίζεται μόνο σε κλειστούς πληθυσμούς πασχόντων.^{9,68} Η θνητότητα ή, αλλιώς, επίπτωση-ποσοστό του θανάτου είναι το ποσοστό των πασχόντων που απεβίωσαν μέσα σ' ένα συγκεκριμένο χρονικό διάστημα εξαιτίας της πάθησης από την οποία έπασχαν. Λαμβάνει τιμές 0–1 και δεν έχει μονάδες μέτρησης.

κλινικές δοκιμές τόσο με την εφαρμογή της μεθόδου των Mantel και Haenszel όσο και με την εφαρμογή της μεθόδου της αντίστροφης διασποράς. Η στάθμιση είναι σχεδόν η ίδια και στις δύο μεθόδους, γεγονός που αντανακλάται και στα αποτελέσματα της μετα-ανάλυσης. Έτσι, με την εφαρμογή της μεθόδου των Mantel και Haenszel προέκυψε ότι ο λόγος των θνητοτήτων ήταν ίσος με 2,12 (95% διάστημα εμπιστοσύνης: 1,4–3,21, $p=0,0004$), ενώ με την εφαρμογή της μεθόδου της αντίστροφης διασποράς προέκυψε ότι ο λόγος των θνητοτήτων ήταν ίσος με 2,08 (95% διάστημα εμπιστοσύνης: 1,35–3,21, $p=0,0009$). Και με τις δύο μεθόδους, δηλαδή, το αποτέλεσμα της μετα-ανάλυσης ήταν ότι στους εμφραγματίες που ελάμβαναν στρεπτοκινάση η θνητότητα σε ένα έτος εξαιτίας του εμφράγματος ήταν περίπου διπλάσια σε σχέση με τη θνητότητα των εμφραγματιών που ελάμβαναν ενεργοποιητή του ιστικού προδρόμου της πλασμίνης. Όπως προαναφέρθηκε, όταν υπάρχει ομοιογένεια μεταξύ των μελετών που συμπεριλαμβάνονται στη μετα-ανάλυση, τότε το μοντέλο των σταθερών αποτελεσμάτων καταλήγει στα ίδια αποτελέσματα με το μοντέλο των τυχαίων αποτελεσμάτων. Έτσι, και στο συγκεκριμένο παράδειγμα, όπου υπήρχε ομοιογένεια μεταξύ των μελετών, η εφαρμογή της μεθόδου των DerSimonian και Laird –μοντέλο των τυχαίων αποτελεσμάτων– κατέληξε στα ίδια αποτελέσματα με την εφαρμογή της μεθόδου της αντίστροφης διασποράς (μοντέλο των σταθερών αποτελεσμάτων).

Στην πλειονότητα των περιπτώσεων, η μετα-ανάλυση συνοδεύεται από ένα διάγραμμα που είναι γνωστό ως διά-

γραμμα «δάσος» (forest plot) και στο οποίο απεικονίζονται τα αποτελέσματα τόσο των επιμέρους μελετών όσο και της μετα-ανάλυσης. Το διάγραμμα «δάσος» είναι εξαιρετικά σημαντικό, καθώς προσφέρει μια σαφή και άμεση εικόνα της μετα-ανάλυσης και η ερμηνεία του δεν απαιτεί ιδιαίτερες γνώσεις. Στην εικόνα 5 φαίνεται το διάγραμμα «δάσος» της μετα-ανάλυσης των 7 κλινικών δοκιμών του πίνακα 1. Στον κάθετο άξονα αντιστοιχούν οι επιμέρους μελέτες που συμπεριελήφθησαν στη μετα-ανάλυση, ενώ στον οριζόντιο άξονα αντιστοιχεί το εκτιμώμενο μέτρο σχέσης, που στο συγκεκριμένο παράδειγμα ήταν ο λόγος των θνητοτήτων. Σε κάθε μελέτη αντιστοιχεί ένα τετράγωνο, που αποτελεί το εκτιμώμενο μέτρο σχέσης στη μελέτη αυτή, καθώς και μια οριζόντια γραμμή διαμέσου του τετραγώνου, τα άκρα της οποίας αποτελούν το διάστημα εμπιστοσύνης του μέτρου σχέσης. Έτσι, στην εικόνα 5, το τετράγωνο που αντιστοιχεί στη μελέτη Α αποτελεί το μέτρο σχέσης που υπολογίστηκε στη μελέτη αυτή και είναι ίσο με 2, ενώ τα άκρα της οριζόντιας γραμμής διαμέσου του τετραγώνου αποτελούν το διάστημα εμπιστοσύνης του μέτρου σχέσης, που είναι ίσο με 0,71–5,64. Παρομοίως, το τετράγωνο που αντιστοιχεί στη μελέτη Β αποτελεί το μέτρο σχέσης που υπολογίστηκε στη μελέτη αυτή και είναι ίσο με 1,69, ενώ τα άκρα της οριζόντιας γραμμής διαμέσου του τετραγώνου αποτελούν το διάστημα εμπιστοσύνης του μέτρου σχέσης, που είναι ίσο με 0,42–6,83. Το μέγεθος των τετραγώνων εξαρτάται από το «βάρος» που προσδίδεται σε κάθε μελέτη για τον υπολογισμό του συγκεντρωτικού αποτελέσματος της μετα-ανάλυσης, με μεγαλύτερα τετράγωνα να αντι-



Εικόνα 5. Διάγραμμα «δάσος» που αντιστοιχεί στα αποτελέσματα της μετα-ανάλυσης των 7 κλινικών δοκιμών του πίνακα 1.

στοιχούν σε μελέτες που έχουν μεγαλύτερο «βάρος». Έτσι, στην εικόνα 5, το τετράγωνο που αντιστοιχεί στη μελέτη Δ είναι το μεγαλύτερο, καθώς η μελέτη αυτή είχε το μεγαλύτερο «βάρος» κατά τον υπολογισμό του συγκεντρωτικού αποτελέσματος της μετα-ανάλυσης.

Στο συγκεκριμένο παράδειγμα, όπου το μέτρο σχέσης ήταν ο λόγος θνητοτήτων, η κάθετη γραμμή που φέρεται στην τιμή 1 στον οριζόντιο άξονα αντιστοιχεί στη μη ύπαρξη σχέσης μεταξύ προσδιοριστή και έκβασης. Έτσι, η κάθετη αυτή γραμμή θα τέμνεται από τις οριζόντιες γραμμές των μελετών που δεν κατέληξαν σε στατιστικά σημαντική σχέση μεταξύ προσδιοριστή και έκβασης. Για το λόγο αυτό, στο συγκεκριμένο παράδειγμα, η μοναδική οριζόντια γραμμή που δεν τέμνει την κάθετη γραμμή η οποία φέρεται στην τιμή 1 στον οριζόντιο άξονα είναι εκείνη που αντιστοιχεί στη μελέτη Δ, η οποία είναι και η μοναδική που κατέληξε σε στατιστικά σημαντική σχέση μεταξύ προσδιοριστή και έκβασης (95% διάστημα εμπιστοσύνης: 1,49–5,96, $p=0,002$).

Στο διάγραμμα «δάσος», το συγκεντρωτικό αποτέλεσμα της μετα-ανάλυσης απεικονίζεται με ένα διαμάντι, τα άκρα του οποίου αποτελούν το διάστημα εμπιστοσύνης του εκτιμώμενου μέτρου σχέσης. Στην εικόνα 5, η κάθετη γραμμή που διέρχεται από το κέντρο του διαμαντιού προς τον οριζόντιο άξονα αντιστοιχεί στην τιμή 2,12, ενώ τα άκρα του διαμαντιού αποτελούν το διάστημα εμπιστοσύνης (95% διάστημα εμπιστοσύνης: 1,4–3,21, $p=0,0004$), όπως ακριβώς προέκυψε και από την εφαρμογή της μεθόδου των Mantel και Haenszel.

Η μετα-ανάλυση κατέληξε στο συμπέρασμα ότι η θνητότητα σε αυτούς που ελάμβαναν στρεπτοκινάση ήταν περίπου 2 φορές μεγαλύτερη σε σχέση με εκείνους που ελάμβαναν ενεργοποιητή του ιστικού προδρόμου της πλασμίνης και ότι η σχέση αυτή ήταν στατιστικά σημαντική. Σημειώνεται ότι οι 6 από τις 7 κλινικές δοκιμές που συμπεριελήφθησαν στη μετα-ανάλυση δεν είχαν καταλήξει σε στατιστικά σημαντικές σχέσεις, υποδηλώνοντας έτσι την απουσία σχέσης μεταξύ προσδιοριστή και έκβασης. Το παράδειγμα αυτό είναι χαρακτηριστικό της συνεισφοράς της μετα-ανάλυσης στην εξαγωγή ασφαλέστερων συμπερασμάτων και στον υπολογισμό ενός συγκεντρωτικού αποτελέσματος με μεγαλύτερη ακρίβεια και εγκυρότητα.

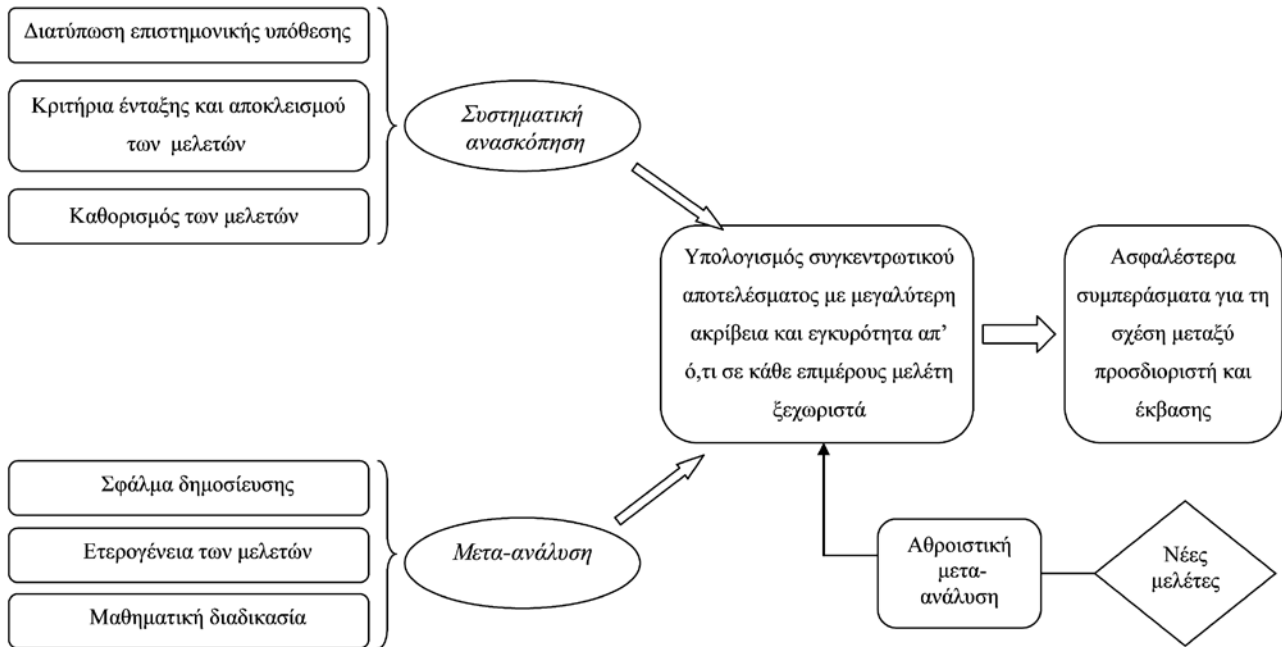
4. ΣΥΝΟΨΗ

Όπως προαναφέρθηκε, φαίνεται ότι είναι προτιμότερη η διάκριση ανάμεσα στη συστηματική ανασκόπηση και τη μετα-ανάλυση, με την πρώτη να αποτελεί τη θεωρητική διαδικασία καθορισμού με ορισμένα κριτήρια των καλύ-

τερα μεθοδολογικά σχεδιασμένων μελετών σχετικά με μια συγκεκριμένη επιστημονική υπόθεση και τη δεύτερη να αποτελεί τη μαθηματική διαδικασία υπολογισμού ενός συγκεντρωτικού αποτελέσματος με βάση τα αποτελέσματα των μελετών οι οποίες επιλέχθηκαν έπειτα από τη συστηματική ανασκόπηση της βιβλιογραφίας (εικ. 6). Έτσι, η συστηματική ανασκόπηση και η μετα-ανάλυση αποτελούν το πρώτο και το δεύτερο βήμα, αντίστοιχα, μιας διαδικασίας συνδυασμού των αποτελεσμάτων ενός αριθμού μελετών αναφορικά με μια συγκεκριμένη επιστημονική υπόθεση, με σκοπό τον υπολογισμό με μεγαλύτερη ακρίβεια και εγκυρότητα –σε σχέση με κάθε επιμέρους μελέτη χωριστά– ενός συγκεντρωτικού αποτελέσματος, το οποίο εκτιμά ουσιαστικά τη σχέση μεταξύ προσδιοριστή και έκβασης.

Τα κριτήρια ένταξης και αποκλεισμού των μελετών, η ύπαρξη σφάλματος δημοσίευσης και η ύπαρξη ετερογένειας μεταξύ των μελετών αποτελούν τα σημαντικότερα χαρακτηριστικά, τα οποία μπορούν να μειώσουν την εγκυρότητα του συγκεντρωτικού αποτελέσματος που εξάγεται μετά από τη συστηματική βιβλιογραφική ανασκόπηση και τη μετα-ανάλυση.^{4,44} Εντούτοις, η πραγματοποίηση συστηματικής ανασκόπησης και μετα-ανάλυσης σχετικά με μια συγκεκριμένη επιστημονική υπόθεση αποτελεί την καλύτερη δυνατή επιλογή και οδηγεί σε ασφαλέστερα συμπεράσματα απ' ό,τι κάθε επιμέρους μελέτη χωριστά. Για το λόγο αυτόν, η διαδικασία που φαίνεται στην εικόνα 6 για τον υπολογισμό ενός συγκεντρωτικού αποτελέσματος είναι απαραίτητη στην προσπάθεια λήψης ορθολογικών αποφάσεων σχετικά με ζητήματα δημόσιας υγείας. Χαρακτηριστικό παράδειγμα αποτελεί η αντιπαράθεση μεταξύ των καπνοβιομηχανιών και των μη κυβερνητικών οργανώσεων κατά του παθητικού καπνίσματος, με τις πρώτες να στηρίζουν τα επιχειρήματά τους στα αποτελέσματα μελετών που κατέληξαν στην ύπαρξη σχέσης μικρού βαθμού μεταξύ παθητικού καπνίσματος και καρκίνου του πνεύμονα και τις δεύτερες να στηρίζονται στα αποτελέσματα μελετών που διαπίστωσαν σημαντική αύξηση της συχνότητας εμφάνισης του καρκίνου του πνεύμονα στους παθητικούς καπνιστές.⁶⁹ Είναι προφανές ότι, στην περίπτωση αυτή, η ορθολογική λήψη αποφάσεων δεν μπορεί να στηριχθεί στα αποτελέσματα μεμονωμένων μελετών, αλλά απαιτείται ο υπολογισμός ενός συγκεντρωτικού αποτελέσματος μέσω της κατάλληλης συστηματικής ανασκόπησης της βιβλιογραφίας.^{70,71}

Το 1993, μάλιστα, εισήχθη η έννοια της αθροιστικής μετα-ανάλυσης (cumulative meta-analysis),²² κατά την οποία κάθε φορά που διεξάγεται μια νέα μελέτη σχετικά με μια επιστημονική υπόθεση, τα αποτελέσματα της μελέτης αυτής προστίθενται στα αποτελέσματα της ήδη υπάρχουσας μετα-ανάλυσης των έως εκείνη τη στιγμή μελετών που έχουν πραγματοποιηθεί, με σκοπό τον υπολογισμό



Εικόνα 6. Διαδικασία υπολογισμού του συγκεντρωτικού αποτελέσματος, που υποδηλώνει τη σχέση μεταξύ προσδιοριστή και συχνότητας εμφάνισης της έκβασης με την εφαρμογή της συστηματικής ανασκόπησης και της μετα-ανάλυσης.

ενός νέου συγκεντρωτικού αποτελέσματος με ακόμη πιο ακριβή και έγκυρο τρόπο. Σημειώνεται ότι η αθροιστική μετα-ανάλυση χρησιμοποιείται ιδιαίτερα στην περίπτωση των κλινικών δοκιμών.⁷²⁻⁷⁴ Οι ερευνητές, εξάλλου, που συνεργάζονται αφιλοκερδώς με την Cochrane Collaboration έχουν συνεισφέρει σημαντικά προς την κατεύθυνση αυτή, διεξάγοντας συνεχώς αθροιστικές μετα-αναλύσεις και

δημοσιεύοντας τα αντίστοιχα αποτελέσματα.⁷⁵⁻⁷⁷ Τέλος, η Cochrane Collaboration προσφέρει δωρεάν στην ηλεκτρονική της διεύθυνση (www.cochrane.org) τη δυνατότητα εφαρμογής ενός στατιστικού πακέτου (RevMan) για την πραγματοποίηση, με απλούστερο και γρηγορότερο τρόπο, μιας συστηματικής ανασκόπησης και μετα-ανάλυσης.

ABSTRACT

Systematic review and meta-analysis

P. GALANIS

Center for Health Services, Management and Evaluation, Department of Nursing, University of Athens, Athens, Greece

Archives of Hellenic Medicine 2009, 26(6):826-841

Systematic review and meta-analysis should be considered as two different interdependent procedures that aim to calculate a summary effect indicating the relationship between the determinant under study and the outcome. The basis of a systematic review is a comprehensive literature review to identify all the similar and relevant studies that satisfy a pre-defined set of inclusion and exclusion criteria. The basic steps in a systematic review include (a) a clear statement of the scientific hypothesis to be tested, (b) a clearly defined statement of inclusion and exclusion criteria for admission of studies into the review, and (c) a methodology for locating studies that should be included in the meta-analysis. The end result of a systematic review is a list of studies each of which provides a weighted value for the calculation of the summary result. Meta-analysis is the mathematical process of combining the separate studies that have been chosen in the systematic review to produce one overall study and to calculate a summary effect. For

example, if there are three separate studies, with 30, 50 and 100 participants, respectively, their combination would generate a total study with 180 participants. The assumption of the meta-analysis is that this combined study will provide a more reliable and precise overall result than any of the smaller individual studies. The basic steps in a meta-analysis include (a) estimation of publication bias, (b) estimation of heterogeneity among the studies, and (c) a mathematical process to combine the results of the separate studies. Estimation of publication bias is achieved with the funnel plot, the rank correlation coefficient (Kendall's or Spearman's) or the sensitivity analysis. L'Abbé plot and DerSimonian-Laird's Q test are used to estimate the heterogeneity among studies. The combined result of the meta-analysis and the results of the individual studies are displayed graphically using the forest plot. Systematic review and meta-analysis offer the best choice for researchers in making rational decisions about public health issues.

Key words: Fixed-effects model, Meta-analysis, Publication bias, Random-effects model, Systematic review

Βιβλιογραφία

- BADGETT R, O'KEEFE M, HENDERSON M. Using systematic reviews in clinical education. *Ann Intern Med* 1997, 126:886–891
- CHALMERS I, ALTMAN D. *Systematic reviews*. British Medical Journal Publishing Groups, London, 1995
- GREENLAND S. Quantitative methods in the review of epidemiologic literature. *Epidemiol Rev* 1987, 9:1–30
- STROUP D, THACKER S. Meta-analysis in epidemiology. In: Armitage P, Colton T (eds) *Encyclopedia of biostatistics*. 2nd ed. John Wiley & Sons, New York, 2005:3448–3462
- MANN C. Can meta-analysis make policy? *Science* 1994, 266:960–962
- THACKER S. Meta-analysis: A quantitative approach to research integration. *J Am Med Dir Assoc* 1988, 259:1685–1689
- BAILAR J. The promise and problems of meta-analysis. *N Engl J Med* 1997, 337:559–560
- MEINERT C. Meta-analysis: Science or religion? *Contemp Clin Trials* 1989, 10:257S–263S
- ΣΠΑΡΟΣ Λ, ΓΑΛΑΝΗΣ Π. *Δοκίμια επιδημιολογίας*. Εκδόσεις Παρισιάνου, Αθήνα, 2006
- ΣΠΑΡΟΣ Λ, ΓΑΛΑΝΗΣ Π, ΖΑΧΟΣ Ι, ΤΣΙΛΙΔΗΣ Κ. *Επιδημιολογία Ι*. Εκδόσεις ΒΗΤΑ, Αθήνα, 2004
- ROTHMAN K, GREENLAND S, LASH T. *Modern epidemiology*. 3rd ed. Lippincott Williams & Wilkins, Philadelphia, 2008
- HARWELL M. An empirical study of Hedges' homogeneity test. *Psychol Methods* 1997, 2:219–231
- KHEIFETS L, AFIFI A, BUFFLER P, ZHANG Z, MATKIN C. Occupational electric and magnetic field exposure and leukemia. A meta-analysis. *J Occup Environ Med* 1997, 39:1074–1091
- KLEIJNEN J, RIET G, KNIPSCHILD P. Vitamin B-6 in the treatment of the premenstrual syndrome – a review. *Br J Obstet Gynaecol* 1990, 97:847–852
- KRUTAN B, TAYLOR M, FREEMAN E. Vitamin B-6 in the treatment of the premenstrual syndrome. *JADA* 1990, 90:859–861
- LAW M, MORRIS J, WALD N. Environmental tobacco smoke exposure and ischaemic heart disease: An evaluation of the evidence. *Br Med J* 1997, 315:973–980
- PETITTI D. *Meta-analysis, decision analysis, and cost effectiveness analysis: Methods for quantitative synthesis in medicine*. Oxford University Press, New York, 1994
- OHLSSON A. Systematic reviews – theory and practice. *Scand J Clin Lab Invest* 1994, 219:25–32
- OLKIN I. Meta-analysis: Current issues in research synthesis. *Stat Med* 1996, 15:1253–1257
- STEINBERG K, SMITH S, LEE N, STROUP D, OLKIN I, WILLIAMSON G. A comparison of meta-analysis to pooled analysis: An application to ovarian cancer. *Am J Epidemiol* 1997, 145:1917–1925
- BOISSEL J, BLANCHARD J, PANAK E, PEYRIEUS J, SACKS H. Considerations for the meta-analysis of randomized clinical trials. *Contemp Clin Trials* 1989, 10:254–281
- CHALMERS T, LAU J. Meta-analytic stimulus for changes in clinical trials. *Stat Methods Med Res* 1993, 2:161–172
- MOSTELLER F, CHALMERS T. Some progress and problems in meta-analysis of clinical trials. *Stat Sci* 1992, 7:227–236
- SACKS H, BERRIER J, REITMAN D, ANCONA-BERK V, CHALMERS T. Meta-analyses of randomized controlled trials. *N Engl J Med* 1983, 316:450–455
- LAU J, ANTMAN E, JIMENEZ-SILVA J, KUPELNICK B, MOSTELLER F, CHALMERS T. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med* 1992, 327:248–254
- BEGG C, CHO M, EASTWOOD S, HORTON R, MOHER D, OLKIN I ET AL. Improving the quality of reporting of randomized controlled trials: The CONSORT statement. *JAMA* 1996, 276:637–639
- MOHER D, SCHULZ K, ALTMAN D. The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001, 285:1987–2007
- GLASS G. Primary, secondary, and meta-analysis of research. *Educational Researcher* 1976, 5:3–8
- STROUP D, THACKER S, OLSON C. Characteristics of meta-analyses related to acceptance for publication in a medical journal. *J Clin Epidemiol* 2001, 54:655–660
- FORD E, SMITH S, STROUP D, STEINBERG K, MUELLER P, THACKER S ET AL. Homocyst(e)ine and cardiovascular disease: A review of the evidence with special emphasis on case-control studies and nested case-control studies. *Int J Epidemiol* 2002, 31:59–70
- IYENGAR S, GREENHOUSE J. Selection models and the file drawer problem. *Statist Sci* 1988, 3:109–135
- SPITZER W. The challenge of meta-analysis. *J Clin Epidemiol* 1995, 48:1–4

33. GREENLAND S. Invited commentary: A critical look at some popular meta-analytic methods. *Am J Epidemiol* 1994, 140:290–296
34. SHAPIRO S. Meta-analysis/Shmeta-analysis. *Am J Epidemiol* 1994, 140:771–778
35. EGGER M, ZELLWEGER-ZAHNERT T, SCHNEIDER M, JUNKER C, LENGELER C, ANTES G. Language bias in randomised controlled trials published in English and German. *Lancet* 1997, 350:326–329
36. GREGOIRE G, DERDERIAN F, Le LORIER J. Selecting the language of the publications included in a meta-analysis: Is there a tower of babel bias? *J Clin Epidemiol* 1995, 48:159–163
37. MOHER D, FORTIN P, JADAD A, JUNI P, KLASSENT, Le LORIER J ET AL. Completeness of reporting of trials published in languages other than English: Implication for conduct and reporting of systematic reviews. *Lancet* 1996, 347:363–366
38. FRIEDENREICH C. Influence of methodological factors in a pooled analysis of 13 case-control studies of colorectal cancer and dietary fiber. *Epidemiology* 1994, 5:66–67
39. HASSELBLAD V, EDDY D, KOTCHMAR D. Synthesis of environmental evidence: Nitrogen dioxide epidemiology studies. *J Air Waste Manag Assoc* 1992, 42:662–671
40. BERLIN J, COLDITZ G. The role of meta-analysis in the regulatory process for foods, drugs, and devices. *J Am Med Dir Assoc* 1999, 281:830–834
41. MOHER D, PHAM B, JONES A, COOK D, JADAD A, MOHER M ET AL. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analysis? *Lancet* 1998, 352:609–613
42. SCHULZ K, CHALMERS I, HAYES R, ALTMAN D. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *J Am Med Dir Assoc* 1995, 273:408–412
43. EGGER M, SCHEIDER M, DAVEY-SMITH G. Meta-analysis: Spurious precision? Meta-analysis of observational studies. *Br Med J* 1998, 316:140–144
44. EMERSON J, BURDICK E, HOAGLIN D, MOSTELLER F, CHALMERS T. An empirical study of the possible relation of treatment differences to quality scores in controlled randomized clinical trials. *Contemp Clin Trials* 1990, 11:339–352
45. THACKER S, STROUP D. Methods and interpretation in systematic reviews: Commentary on two parallel reviews of epidural analgesia during labor. *Am J Obstet Gynecol* 2002, 186:785–805
46. SIPE T, CURLETTE W. A meta-synthesis of factors related to educational achievement: A methodological approach to summarizing and synthesizing meta-analyses. *Int J Educ Res* 1997, 25:583–698
47. BYARS DP. *The use of databases and historical controls in treatment comparison, on combining information: Historical controls, overviews, and comprehensive cohort studies. Recent results in cancer research.* Volume 111. Springer-Verlag, New York, 1988
48. COOK D, GUYATT G, RYAN G, CLIFTON J, BUCKINGHAM L, WILLAN A ET AL. Should unpublished data be included in meta-analyses? Current confusions and controversies. *J Am Med Dir Assoc* 1993, 21:2749–2753
49. DICKERSIN K, MIN YI. NIH clinical trials and publication bias. *Online J Curr Clin Trials* 1993, Document number 50
50. HIGGINS J, GREEN S. *Cochrane handbook for systematic reviews of interventions.* The Cochrane Collaboration and John Wiley & Sons Ltd, Chichester, 2008
51. EASTERBROOK P. Directory of registries of clinical trials. *Stat Med* 1992, 11:345–423
52. HUSTON P, MOHER D. Redundancy, disaggregation, and the integrity of medical research. *Lancet* 1996, 347:1024–1026
53. OLSON C, RENNIE D, COOK D, DICKERSIN K, FLANAGIN A, HOGAN J ET AL. Publication bias in editorial decision making. *JAMA* 2002, 287:2825–2828
54. EASTERBROOK P, BERLIN J, GOPALAN R, MATTHEWS D. Publication bias in clinical research. *Lancet* 1991, 337:867–872
55. ROSENTHAL R. The “file drawer problem” and tolerance for null results. *Psychol Bull* 1979, 86:638–641
56. LIGHT R, PILLEMER D. *Summing up: The science of reviewing research.* Harvard University Press, Boston, 1984
57. SONG F, KHAN K, DINNES J, SUTTON A. Asymmetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. *Int J Epidemiol* 2002, 31:88–95
58. ARMITAGE P, BERRY G. *Statistical methods in medical research.* 2nd ed. Blackwell Publishing, Oxford, 1987
59. COLTON T. *Statistics in medicine.* Little Brown & Co, Boston, 1974
60. THACKER S, STROUP D, BRANCHE C, GILCHRIST J, GOODMAN R, WEITMAN E. The prevention of ankle sprains in sports: A systematic review of the literature. *Am J Sports Med* 1999, 27:753–760
61. COLDITZ G, BURDICK E, MOSTELLER F. Heterogeneity in meta-analysis of data from epidemiologic studies: A commentary. *Am J Epidemiol* 1995, 142:371–382
62. COCHRAN W. The combination of estimates from different experiments. *Biometrics* 1954, 10:101–129
63. DerSIMONIAN R, LAIRD N. Meta-analysis in clinical trials. *Contemp Clin Trials* 1986, 7:177–188
64. HALL S, BRANNICK M. Comparison of two random-effects method of meta-analysis. *J Appl Psychol* 2002, 87:377–389
65. HEDGES L, OLKIN I. *Statistical methods for meta-analysis.* Academic Press, Boston, 1985
66. COOPER H, HEDGES L. *The handbook of research synthesis.* Russell Sage Foundation, New York, 1994
67. DuMOUCHEL W. *Hierarchical bays linear models for meta-analysis.* National institute of statistical sciences. Research Triangle Park, North Carolina, 1994
68. ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Μέτρα συχνότητας των νοσημάτων. *Αρχ Έλλ Ιατρ* 2005, 22:178–191
69. HUNT M. *How science takes stock: The story of meta-analysis.* Russell Sage Foundation, New York, 1997
70. L'ABBÉ K, DETSKY A, O'ROURKE K. Meta-analysis in clinical research. *Ann Intern Med* 1987, 107:224–233
71. WACHTER K, STRAF M. *The future of meta-analysis.* Russell Sage Foundation, New York, 1990
72. EGGER M, SMITH G, STERNE J. Meta-analysis: Is moving the goal post the answer? *Lancet* 1998, 351:15–17

73. POGUE J, YUSUF S. Cumulating evidence from randomized trials: Utilizing sequential monitoring boundaries for cumulative meta-analysis. *Contemp Clin Trials* 1997, 18:580–593
74. WARE J, MULLER J, BRAUNWALD E. The futility index: An approach to the cost-effective termination of randomized clinical trials. *Am J Med* 1985, 78:635–643
75. BLETTNER M, SAUERBREI W, SCHLEHOFER B, SCHEUCHENPFLUBT, FRIEDENREICH C. Traditional reviews, meta-analyses and pooled analyses in epidemiology. *Int J Epidemiol* 1999, 28:1–9
76. COLDITZ G, BREWERT, BERKEY C, WILSON M, BURDICK E, FINEBERG H. Efficacy of BCG vaccine in the prevention of tuberculosis: Meta-analysis of the published literature. *J Am Med Dir Assoc* 1994, 271:696–702
77. WILSON M, FINEBERG H, COLDITZ G. Geographic latitude and the efficacy of bacillus Calmette-Guerin vaccine. *Clin Infect Dis* 1995, 20:982–991

Corresponding author:

P. Galanis, 14 Dikis street, GR-157 73 Athens, Greece
e-mail: pegalan@nurs.uoa.gr