

Διαστρωμάτωση επιδημιολογικών δεδομένων[#]

Π. Γαλάνης,^{##}
Λ. Σπάρος

Εργαστήριο Κλινικής Επιδημιολογίας,
Τμήμα Νοσηλευτικής, Πανεπιστήμιο
Αθηνών, Αθήνα

Stratification
of epidemiological data

Abstract at the end of the article

1. Εισαγωγή
2. Η έννοια του συχυτή
3. Διαστρωματική ανάλυση δεδομένων
4. Εξουδετέρωση των συχυτών με τη μέθοδο της διαστρωμάτωσης
 - 4.1. Μελέτες με κλειστούς πληθυσμούς: Επίπτωση-ποσοστό, επιπολασμός
 - 4.2. Μελέτες με κλειστούς πληθυσμούς: Επίπτωση-πυκνότητα
 - 4.3. Μελέτες «ασθενών-μαρτύρων»
5. Διαστρωμάτωση στην περίπτωση δύο ή περισσότερων συχυτών
6. Σημασία της διαστρωμάτωσης

Λέξεις ευρετηρίου

Διαστρωμάτωση
Προσδιοριστής
Σύχχυση
Συχυτή
Τροποποίηση του μέτρου αποτελέσματος

Υποβλήθηκε 14.9.2005
Εγκρίθηκε 28.11.2005

1. ΕΙΣΑΓΩΓΗ

Η *διαστρωματική ανάλυση*¹⁻³ (stratified analysis) ή *διαστρωμάτωση* (stratification) είναι ένας τρόπος μελέτης και ελέγχου της επίδρασης άλλων προσδιοριστών (εξωγενών), εκτός του μελετώμενου προσδιοριστή, στη συχνότητα εμφάνισης του νοσήματος, στη φάση της ανάλυσης* των δεδομένων. Διαστρωμάτωση σημαίνει ότι τα δεδομένα χωρίζονται σε υποομάδες ή στρώματα (strata). Π.χ., διαστρωμάτωση κατά φύλο ή ηλικία σημαίνει ότι τα δεδομένα χωρίζονται σε κατηγορίες ανδρών και γυναικών ή κατά ηλικία.⁶

Παράγοντας κινδύνου (risk factor) ή *έκθεση* (exposure) ή *προσδιοριστής* (determinant), όπως τελικά επι-

κράτησε να λέγεται σήμερα, είναι το χαρακτηριστικό (συγγενές, περιβαλλοντικό ή συμπεριφοράς) των ατόμων από το οποίο εξαρτάται (σχετίζεται ή συναρτάται) η συχνότητα της μελετώμενης έκβασης.^{5,7,8} Η ακαταλληλότητα του όρου παράγοντας κινδύνου (risk factor) προέκυψε από τη λέξη factor (παράγοντας), που σημαίνει αιτία, ενώ ο παράγοντας κινδύνου δεν σχετίζεται κατ' ανάγκη αιτιακά με τη συχνότητα του νοσήματος. Οι προσδιοριστές (της συχνότητας) διακρίνονται σε περιγραφικούς (descriptive) και σε αιτιακούς (causal). Παλαιότερα, οι περιγραφικοί προσδιοριστές φέρονταν με τον όρο δείκτες κινδύνου (risk indicators), ενώ οι αιτιακοί προσδιοριστές με τον όρο παράγοντες κινδύνου. Ο προσδιοριστής της συχνότητας μιας μελετώμενης έκβασης περιλαμβάνει δύο, τουλάχιστον, 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 εκλογής του ερευνητή, άρα συνέπεια του σχεδιασμού της μελέτης. Από την επιλογή της κατηγορίας αναφοράς εξαρτάται η ερμηνεία των αποτελεσμά-

* Ο όρος «ανάλυση δεδομένων» περιλαμβάνει⁴ (α) τα περιγραφικά στατιστικά μέτρα (descriptive statistic), που συνοψίζουν τα δεδομένα, και (β) τα διαλογισμικά στατιστικά μέτρα (inferential statistic), που χρησιμοποιούνται στα στατιστικά υποδείγματα ή μοντέλα, για την εξαγωγή συμπερασμάτων για τα αντικείμενα μιας μελέτης, που είναι, ουσιαστικά, η παράμετρος που ενδιαφέρει τους ερευνητές. Πρέπει να σημειωθεί ότι ο όρος «ανάλυση δεδομένων» είναι εσφαλμένος εννοιολογικά, αφού οι παρατηρήσεις δεν είναι δεδομένα και η επεξεργασία των παρατηρήσεων είναι σύνθεση και όχι ανάλυση. Στην ουσία, η ανάλυση των δεδομένων είναι το σύνολο της μαρτυρίας (evidence) που η εμπειρική έρευνα χρησιμοποιεί για τον έλεγχο της υπόθεσης.⁵

[#] Μερική χρηματοδότηση: Ειδικός Λογαριασμός Κονδυλίων Έρευνας Πανεπιστημίου Αθηνών

^{##} Ο Π. Γαλάνης είναι υπότροφος του Ιδρύματος Α.Σ. ΩΝΑΣΗ

των μιας μελέτης. Για παράδειγμα, προσδιοριστής της συχνότητας της νεφρικής νόσου 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 της νεφρικής νόσου, σε σχέση πάντοτε με τη συχνότητα της νεφρικής νόσου στην κατηγορία αναφοράς, στην προκειμένη περίπτωση στην κατηγορία των «νορμοπαισικών». Ο όρος έκθεση χρησιμοποιείται⁹ (α) ως μια ειδική (ενδεικτική) κατηγορία ενός (δυσνητικού) περιβαλλοντικού (σε αντίθεση με το συγγενή ή συμπεριφορικό) προσδιοριστή, που λειτουργεί ως αιτιακός ή προληπτικός παράγοντας της μελετώμενης έκβασης σε σχέση με μια άλλη συγκεκριμένη κατηγορία (κατηγορία αναφοράς), (β) ως η ενδεικτική κατηγορία (έκθεση) ενός περιβαλλοντικού προσδιοριστή (δυσνητικού) της μελετώμενης έκβασης, (γ) ως περιβαλλοντικός προσδιοριστής, (δ) εσφαλμένα ως προσδιοριστής (συγγενής, περιβαλλοντικός ή συμπεριφορικός) ή ως ενδεικτική κατηγορία ενός προσδιοριστή.

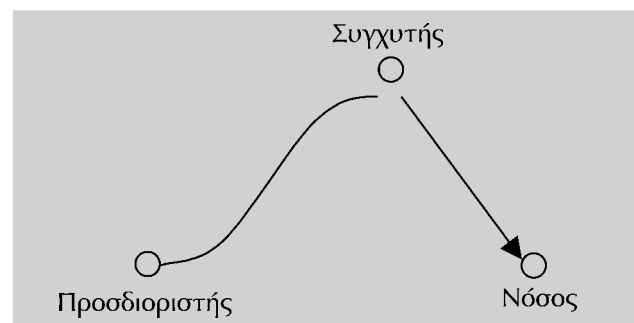
Ο πρωταρχικός λόγος για τη χρησιμοποίηση της διαστρωματικής ανάλυσης είναι να εκτιμηθεί και, αν είναι αναγκαίο, να ελεγχθεί η σύγχυση που προκύπτει όταν κάποιος προσδιοριστής (εξωγενής), άλλος από το μελετώμενο προσδιοριστή, είναι περισσότερο ή λιγότερο επικρατής στην ενδεικτική κατηγορία απ' ό,τι στην κατηγορία αναφοράς. Όταν τα δεδομένα στρωματοποιούνται σε επίπεδα του συγχυτή, π.χ. σε άνδρες και γυναίκες, κάθε στρώμα είναι ελεύθερο σύγχυσης από τη μεταβλητή διαστρωμάτωσης. Έτσι, αν η συσχέτιση μεταξύ έκθεσης και νόσου αναλυθεί ξεχωριστά, π.χ. σε άνδρες και γυναίκες, καθεμιά από τις δύο ομάδες θα δώσει μια εκτίμηση του αποτελέσματος της έκθεσης που είναι ανεξάρτητη του φύλου. Συχνά, τα επιμέρους αυτά αποτελέσματα κατά στρώματα δεν αναφέρονται ξεχωριστά, αλλά ενοποιημένα (pooled) σ' ένα μόνο αποτέλεσμα.^{4,10}

Ο δεύτερος λόγος εφαρμογής της διαστρωματικής ανάλυσης είναι η μελέτη της *τροποποίησης του μέτρου αποτελέσματος* (effect measure modification).^{1-3,11-14} Τροποποίηση του μέτρου αποτελέσματος σημαίνει ότι το αποτέλεσμα του μελετώμενου προσδιοριστή είναι ισχυρότερο σε μερικά στρώματα σε σχέση με άλλα. Παραδείγματος χάριν, αν η σχετική επίπτωση-πυκνότητα μεταξύ προσδιοριστή και νόσου είναι 2 στις γυναίκες και 3 στους άνδρες, το φύλο θα θεωρηθεί *τροποποιητής του αποτελέσματος* (effect modifier).^{6,12-14} Η παρούσα ανασκόπηση επικεντρώνεται στο πώς χρησιμοποιείται η διαστρωμάτωση των δεδομένων για την εξουδετέρωση των συγχυτών μιας μελέτης.

2. Η ΕΝΝΟΙΑ ΤΟΥ ΣΥΓΧΥΤΗ

Η *σύγχυση* (confounding)^{1-3,11-14} στις επιδημιολογικές μελέτες ορίζεται ως η απόκλιση ενός μέτρου σχέσης που προκαλείται από έναν εξωγενή προσδιοριστή και συμβαίνει μόνο μέσα στα πλαίσια μιας συγκεκριμένης μελέτης. Τα κυριότερα μέτρα αποτελέσματος σε μια αιτιολογική μελέτη είναι δύο: (α) η διαφορά των επιπτώσεων-πυκνοτήτων (IDe-IDo) και (β) ο λόγος των επιπτώσεων-πυκνοτήτων (ή σχετική επίπτωση-πυκνότητα, IDe/IDo). Με IDe συμβολίζεται η επίπτωση-πυκνότητα στην ενδεικτική κατηγορία (ή κατηγορία των εκτεθειμένων), ενώ με IDo συμβολίζεται η επίπτωση-πυκνότητα στην κατηγορία αναφοράς (ή κατηγορία των μη εκτεθειμένων). Η σύγχυση δεν είναι ένα χαρακτηριστικό του κάθε εξωγενή προσδιοριστή. Δηλαδή, ένας εξωγενής προσδιοριστής που είναι συγχυτής μιας σχέσης σε μια μελέτη, μπορεί να μην είναι συγχυτής της ίδιας σχέσης σε μια άλλη μελέτη.^{7,12,14} Ο *συγχυτής* (confounder) πρέπει να έχει τα παρακάτω χαρακτηριστικά (εικ. 1):^{6,9,12}

- Να σχετίζεται με τη νόσο (να αποτελεί αιτία ή απλά προσδιοριστή της νόσου, αλλά όχι αποτέλεσμα της νόσου).
- Να σχετίζεται με το μελετώμενο προσδιοριστή, να ανισοκατανέμεται δηλαδή στις δύο κατηγορίες του προσδιοριστή.
- Να μην αποτελεί αποτέλεσμα του μελετώμενου προσδιοριστή. Για παράδειγμα, σε μια δίαιτα πλούσια σε κορεσμένα λίπη, η οποία προκαλεί αύξηση των επιπέδων της λιποπρωτεΐνης χαμηλής περιεκτικότητας (LDL) στο αίμα, η αυξημένη LDL αυξάνει την πιθανότητα εμφάνισης της αθηρωματικής νόσου. Έτσι, η αυξημένη LDL σχετίζεται τόσο με τη νόσο (αθηρωματική νόσο) όσο και με το μελετώμενο προσδιοριστή (δίαιτα πλούσια σε κορεσμένα λίπη). Η αυξημένη LDL, ωστόσο, δεν είναι συγχυτής της σχέσης μεταξύ



Εικόνα 1. Σχηματική απεικόνιση των χαρακτηριστικών ενός συγχυτή.⁶

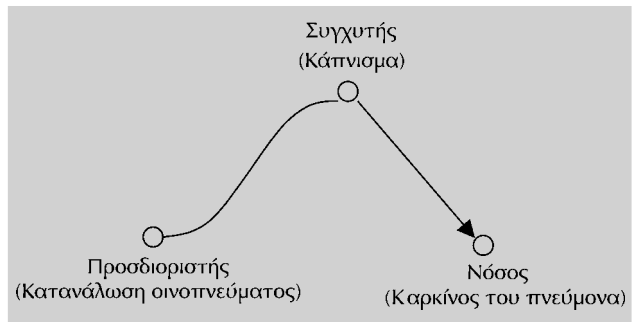
δίαιτας και αθηρωματικής νόσου, γιατί αποτελεί αποτέλεσμα του μελετώμενου προσδιοριστή, καθώς η πλούσια σε κορεσμένα λίπη διαίτα προκαλεί αύξηση της LDL.

Παράδειγμα συγχυτικής σχέσης αποτελεί η διαπίστωση ότι άτομα που καταναλώνουν μεγάλη ποσότητα οινοπνεύματος έχουν μεγαλύτερο κίνδυνο εμφάνισης καρκίνου του πνεύμονα από εκείνα που δεν καταναλώνουν οινοπνεύμα. Η συγχυτική αυτή σχέση οφείλεται στο ότι τα άτομα της πρώτης κατηγορίας είναι σε μεγαλύτερη αναλογία βαρείς καπνιστές σε σχέση με τα άτομα της δεύτερης κατηγορίας. Η συγχυτική επίδραση του καπνίσματος εξηγεί την παρατηρηθείσα σχέση μεταξύ κατανάλωσης οινοπνεύματος και καρκίνου του πνεύμονα (εικ. 2).⁶ Το κάπνισμα είναι συγχυτής της σχέσης οινοπνεύματος και καρκίνου του πνεύμονα, γιατί έχει τα παρακάτω χαρακτηριστικά:

- Σχετίζεται με τη νόσο, καθώς αποτελεί προσδιοριστή του καρκίνου του πνεύμονα.
- Σχετίζεται με το μελετώμενο προσδιοριστή, καθώς τα άτομα που καταναλώνουν μεγαλύτερη ποσότητα οινοπνεύματος είναι, συνήθως, και άτομα που καπνίζουν περισσότερο.
- Δεν αποτελεί συνέπεια του μελετώμενου προσδιοριστή, δηλαδή της κατανάλωσης οινοπνεύματος.

Μια καλά σχεδιασμένη μελέτη επιδιώκει την εξουδετέρωση των συγχυτών,* η οποία μπορεί να γίνει με διάφορους τρόπους, όπως η διαστρωμάτωση, η εξομοίωση, η πολυμεταβλητή ανάλυση, η προτύπωση, η περιοριστική μέθοδος και το πείραμα.^{6,13,14} Η εξουδετέρωση όλων των δυνατών, γνωστών και άγνωστων, συγχυτών είναι απαραίτητη προϋπόθεση για την ανάδειξη μιας σχέσης σε αιτιακή. Η μη εξουδετέρωση όλων των συγχυτών

* *Εξομοίωση (matching)*: Επιλογή μιας ομάδας ατόμων, έτσι ώστε να είναι παρόμοια με μια άλλη ομάδα, ως προς την κατανομή μίας ή περισσότερων μεταβλητών.^{6,13,15} *Διαστρωμάτωση (stratification)*: Υποδιαίρεση σε κατηγορίες (στρώματα) ως προς μία ή περισσότερες μεταβλητές.^{6,13,15} *Προτύπωση (standardization)*: Χρησιμοποιείται, συνήθως, για τη σύγκριση της θνησιμότητας μεταξύ πληθυσμών που έχουν διαφορετική κατανομή κατά ηλικία ή φύλο, ιδιαίτερα στην επαγγελματική επιδημιολογία. Διακρίνεται σε άμεση (direct) και έμμεση (indirect).^{6,13,15} *Περιοριστική μέθοδος (restriction)*:^{6,13} Η μελέτη αφορά μόνο στο ένα επίπεδο του συγχυτή. Για παράδειγμα, αν μελετηθεί το αποτέλεσμα της δράσης ενός φαρμάκου μόνο σε μια ηλικιακή κατηγορία, τότε εξουδετερώνεται η σύγχυση που προκαλείται από τη διαφορετική κατανομή κατά ηλικία. Με τον τρόπο αυτόν, ωστόσο, τα αποτελέσματα της μελέτης περιορίζονται σε μία συγκεκριμένη ηλικιακή κατηγορία. *Πείραμα (experiment)*:^{6,13} Μελέτη, όπου η έκθεση ελέγχεται από τον ερευνητή με σκοπό τη βελτίωση της εγκυρότητας της μελέτης, π.χ. με τυχαία χορήγηση της έκθεσης στα άτομα του πληθυσμού της μελέτης (τυχαιοποιημένο πείραμα).



Εικόνα 2. Παράδειγμα συγχυτικής επίδρασης: Η σχέση μεταξύ κατανάλωσης οινοπνεύματος και καρκίνου του πνεύμονα.⁶

εισάγει συστηματικό σφάλμα και καθιστά τα αποτελέσματα της μελέτης μη έγκυρα.

Για την καλύτερη κατανόηση της έννοιας του συγχυτή θα αναλυθούν τα δεδομένα του πίνακα 1, στον οποίο φαίνονται οι θνησιμότητες στους άνδρες και τις γυναίκες που πάσχουν από νευραλγία του τριδύμου.¹⁶ Ο λόγος των θνησιμοτήτων (θνησιμότητα** στους άνδρες/θνησιμότητα στις γυναίκες) είναι 1,10, γεγονός που υποδηλώνει ότι, με βάση τα αδρά αυτά δεδομένα*** (crude data), η θνησιμότητα στους άνδρες είναι λίγο μεγαλύτερη από τη θνησιμότητα στις γυναίκες. Στο συγκεκριμένο παράδειγμα, προσδιοριστής της έκβασης (δη-

Πίνακας 1. Θνησιμότητες ασθενών με νευραλγία του τριδύμου, ανάλογα με το φύλο.¹⁹

	Άνδρες	Γυναίκες
Θάνατοι	90	131
Ανθρωπο-έτη	2465	3946
Θνησιμότητα (έτος) ⁻¹	0,0365	0,0332
Λόγος θνησιμοτήτων	1,10	
90% διάστημα εμπιστοσύνης	0,88-1,38	

** Όταν η μελετώμενη έκβαση είναι ο θάνατος, τότε η επίπτωση του θανάτου καλείται *θνησιμότητα (mortality)*.^{5,17-19} Στην περίπτωση των κλειστών πληθυσμών η θνησιμότητα υπολογίζεται είτε ως επίπτωση-ποσοστό είτε ως επίπτωση-πυκνότητα, ενώ στην περίπτωση των ανοικτών πληθυσμών η θνησιμότητα μπορεί να υπολογιστεί μόνο ως επίπτωση-πυκνότητα, οπότε:

$$\text{Θνησιμότητα} = \frac{\text{Αριθμός θανάτων από ένα νόσημα X}}{\text{Άθροισμα των χρονικών περιόδων παρακολούθησης που τα μελετώμενα άτομα βρίσκονται σε κίνδυνο να πεθάνουν από το νόσημα X}}$$

*** Δεδομένα (data) ή εμπειρικά δεδομένα (empirical data), με την έννοια ότι υπάρχουν ανεξάρτητα από τη νόσηση, δεν υφίστανται. Αποτελούν, ουσιαστικά, παρατηρήσεις (observations) νοητικά φορτισμένες. Αυτό, εξάλλου, διακρίνει τις παρατηρήσεις από τις εντυπώσεις (impressions).²⁰

λαδή του θανάτου) είναι το φύλο. Ενδεικτική κατηγορία (ή κατηγορία των εκτεθειμένων) θεωρούνται οι άνδρες, ενώ κατηγορία αναφοράς (ή κατηγορία των μη εκτεθειμένων) θεωρούνται οι γυναίκες. Η σχέση, ωστόσο, μεταξύ προσδιοριστή (φύλου) και έκβασης (θανάτου) δεν μπορεί να θεωρηθεί αιτιακή, λόγω της ύπαρξης συγχυτών, γνωστών και άγνωστων, που δεν έχουν εξουδετερωθεί. Η καλύτερη μέθοδος για την εξουδετέρωση των γνωστών συγχυτών, όταν ο αριθμός τους είναι μικρός, θεωρείται η διαστρωματική ανάλυση των δεδομένων. Στο παραπάνω παράδειγμα, η ηλικία θεωρείται ότι είναι ένας πιθανός συγχυτής της σχέσης μεταξύ φύλου και θανάτου. Για το λόγο αυτόν, τα δεδομένα στρωματοποιούνται σύμφωνα με τον πιθανό συγχυτή (δυναμικό συγχυτή, potential confounder), σύμφωνα δηλαδή με την ηλικία, οπότε προκύπτουν τα δεδομένα του πίνακα 2. Αναφορικά με την ηλικία, οι ασθενείς κατανομούνται σε δύο στρώματα. Στο ένα στρώμα κατανομούνται οι ασθενείς που είναι <65 ετών και στο άλλο στρώμα κατανομούνται οι ασθενείς που είναι >65 ετών. Τα αποτελέσματα που προκύπτουν, μετά τη διαστρωμάτωση, έχουν ως εξής:

- Οι ασθενείς, τόσο οι άνδρες όσο και οι γυναίκες, που ανήκουν στην κατηγορία «>65 ετών» έχουν πολύ μεγαλύτερη θνησιμότητα από τους ασθενείς που ανήκουν στην κατηγορία «<65 ετών». Πιο συγκεκριμένα, η θνησιμότητα στους άνδρες ασθενείς >65 ετών είναι $0,0801$ (έτος)⁻¹, ενώ η θνησιμότητα στους άνδρες ασθενείς <65 ετών είναι $0,0092$ (έτος)⁻¹. Επιπλέον, η θνησιμότητα στις γυναίκες ασθενείς >65 ετών είναι $0,0539$ (έτος)⁻¹, ενώ η θνησιμότητα στις γυναίκες ασθενείς <65 ετών είναι $0,0059$ (έτος)⁻¹. Η εντυπωσιακά μεγαλύτερη θνησιμότητα στις μεγαλύτερες ηλικίες δεν είναι χαρακτηριστικό μόνο του συγκεκριμένου πληθυσμού των ασθενών, αλλά οποιουδήποτε πληθυσμού και βέβαια και του γενικού.
- Η διαστρωμάτωση αποκαλύπτει τη διαφορά των ηλικιακών κατανομών στους άνδρες και στις γυναίκες ασθενείς. Στους άνδρες, ο πληθυσμο-χρόνος είναι

μεγαλύτερος για την κατηγορία <65 ετών (1516 ανθρωπο-έτη), ενώ, αντίθετα, στις γυναίκες ο πληθυσμο-χρόνος είναι μεγαλύτερος για την κατηγορία >65 ετών (2245 ανθρωπο-έτη). Η διαφορά αυτή στην ηλικιακή κατανομή θα μπορούσε να οδηγήσει σε μικρότερη συνολική (και στις δύο, δηλαδή, ηλικιακές κατηγορίες) θνησιμότητα στους άνδρες σε σχέση με τις γυναίκες, καθώς η σύγκριση των θνησιμοτήτων μεταξύ ανδρών και γυναικών θα μπορούσε να θεωρηθεί, ως ένα βαθμό, ως σύγκριση θνησιμοτήτων μεταξύ νεότερων (ανδρών) και γηραιότερων (γυναικών).

- Με βάση τα δεδομένα του πίνακα 1, ο λόγος των θνησιμοτήτων (άνδρες/γυναίκες) είναι 1,10, ενώ με βάση τα δεδομένα του πίνακα 2, οι θνησιμότητες στις δύο ηλικιακές ομάδες είναι 1,57 (<65 ετών) και 1,49 (>65 ετών). Η διαφορά ανάμεσα στον αδρό λόγο θνησιμοτήτων (1,10) και τους λόγους θνησιμοτήτων (1,57 και 1,49), που προέκυψαν έπειτα από τη διαστρωμάτωση, είναι αποτέλεσμα της σύγχυσης που επιφέρει η ηλικία και, πιο συγκεκριμένα, του ότι οι γυναίκες ασθενείς είναι μεγαλύτερες ηλικιακά από τους άνδρες ασθενείς. Συμπερασματικά, η ηλικία είναι συγχυτής της σχέσης μεταξύ φύλου και θανάτου στη συγκεκριμένη μελέτη. Επειδή οι άνδρες είναι μικρότερης ηλικίας από τις γυναίκες, η αδρή θνησιμότητα είναι πολύ μικρότερη από τις δύο επιμέρους θνησιμότητες, ανάλογα με την ηλικία, που προκύπτουν έπειτα από τον έλεγχο της σύγχυσης.

3. ΔΙΑΣΤΡΩΜΑΤΙΚΗ ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ

Όπως προαναφέρθηκε, ο κύριος λόγος για τη χρησιμοποίηση της διαστρωματικής ανάλυσης^{1-3,6,10,12-14,21-25} είναι η εκτίμηση και, αν είναι αναγκαίο, ο έλεγχος της σύγχυσης που προκύπτει όταν κάποιο χαρακτηριστικό (εξωγενής προσδιοριστής), άλλο από το μελετώμενο προσδιοριστή, κατανέμεται ανισότιμα στην ενδεικτική κατηγορία σε σχέση με την κατηγορία αναφοράς του μελετώμενου προσδιοριστή. Όταν τα δεδομένα στρωματοποιούνται κατά επίπεδα του συγχυτή, για παράδειγμα σε άνδρες και γυναίκες, κάθε στρώμα θα είναι ελεύθερο σύγχυσης από τη μεταβλητή διάστρωσης. Έτσι, αν η σχέση μεταξύ προσδιοριστή και νοσήματος αναλυθεί χωριστά, π.χ. σε άνδρες και γυναίκες, σε καθένα από τα δύο στρώματα η εκτίμηση του αποτελέσματος είναι ανεξάρτητη του φύλου. Συχνά, τα επιμέρους αυτά αποτελέσματα κατά στρώματα δεν αναφέρονται χωριστά, αλλά ενοποιημένα σε ένα μόνο συνοπτικό (ή συνολικό ή συγκεντρωτικό) αποτέλεσμα.

Πίνακας 2. Θνησιμότητες ασθενών με νευραλγία του τριδύμου, ανάλογα με το φύλο και την ηλικία.¹⁹

	Ηλικία (έτη)			
	<65		≥65	
	Άνδρες	Γυναίκες	Άνδρες	Γυναίκες
Θάνατοι	14	10	76	121
Ανθρωπο-έτη	1516	1701	949	2245
Θνησιμότητα (έτος) ⁻¹	0,0092	0,0059	0,0801	0,0539
Λόγος θνησιμοτήτων	1,57		1,49	

Με τη διαστρωμάτωση, τα δεδομένα παρουσιάζονται σε καθένα από τα στρώματα χωριστά. Οι ερευνητές, ωστόσο, συχνά θέλουν να παρουσιάσουν τη σχέση μεταξύ προσδιοριστή και νόσου μ' έναν απλούστερο τρόπο. Έτσι, αρχικά πραγματοποιούν τις ουσιώδεις συγκρίσεις μεταξύ των διαφόρων στρωμάτων και στη συνέχεια συνενώνουν την πληροφορία που προέρχεται από το σύνολο των συγκρίσεων. Υπάρχουν δύο μέθοδοι για τη συνένωση της πληροφορίας που προκύπτει από το σύνολο των στρωμάτων, η *συνένωση* (pooling) και η *προτύπωση* (standardization). Οι δύο αυτές μέθοδοι χρησιμοποιούν διαφορετικούς τρόπους για να συνενώσουν τα αποτελέσματα των διαφόρων στρωμάτων. Στην παρούσα ανασκόπηση θα αναφερθεί εκτενώς η μέθοδος της συνένωσης.

Απαραίτητη προϋπόθεση για να πραγματοποιηθεί η συνένωση είναι το αποτέλεσμα του μελετώμενου προσδιοριστή να είναι το ίδιο σε όλα τα στρώματα. Με άλλη διατύπωση, η συνένωση δικαιολογείται μόνο εάν φαίνεται απίθανο, είτε από τα δεδομένα μιας συγκεκριμένης μελέτης είτε από τα δεδομένα άλλων μελετών, ο συγχυτής να είναι ταυτόχρονα και τροποποιητής του μέτρου αποτελέσματος.^{13,14} Απαραίτητη προϋπόθεση, δηλαδή, για να πραγματοποιηθεί η συνένωση είναι ο συγχυτής να μην είναι ταυτόχρονα και τροποποιητής σε μια συγκεκριμένη μελέτη. Σύμφωνα με την υπόθεση αυτή, κάθε στρώμα παρέχει μια ξεχωριστή εκτίμηση του αποτελέσματος, που αναφέρεται ως ειδική για κάθε στρώμα εκτίμηση (stratum-specific estimate). Εφόσον κάθε στρώμα παρέχει μια ειδική εκτίμηση του αποτελέσματος, η αρχή στην οποία στηρίζεται η συνένωση είναι ο υπολογισμός του μέσου όρου αυτών των ειδικών –για κάθε στρώμα– εκτιμήσεων του αποτελέσματος. Ο μέσος όρος υπολογίζεται ως «σταθμισμένος» μέσος όρος (weighed average). Με τον τρόπο αυτό διαπιστώνεται ότι μερικά στρώματα δίνουν περισσότερες πληροφορίες από άλλα. Όσο μεγαλύτερος είναι ο αριθμός των δεδομένων ενός στρώματος τόσο μεγαλύτερη είναι η πληροφορία που το στρώμα αυτό παρέχει για την εκτίμηση του τελικού αποτελέσματος. Έτσι, ουσιαστικά, κατά τη διαδικασία της συνένωσης δίνεται μεγαλύτερο βάρος στα στρώματα που έχουν περισσότερα δεδομένα. Η διαδικασία της συνένωσης είναι δυνατόν να πραγματοποιηθεί με διάφορες τεχνικές.

Η συχνότερα χρησιμοποιούμενη, και η οποία θα αναφερθεί εκτενώς στη συνέχεια, είναι η *τεχνική κατά Mantel-Haenszel*. Μια δεύτερη τεχνική είναι αυτή της *μεγίστης πιθανοφάνειας* (maximum likelihood technique) που χρησιμοποιείται στις πολυμεταβλητές μεθόδους, όπως η *πολυμεταβλητή λογιστική παλινδρόμηση* (multiple logistic

regression) και το μοντέλο των αναλογικών κινδύνων. Στη συνέχεια, θα παρατεθούν οι τύποι με τους οποίους υπολογίζεται ο «σταθμισμένος» μέσος όρος, στηριζόμενοι στην τεχνική κατά Mantel-Haenszel. Όταν το αποτέλεσμα του μελετώμενου προσδιοριστή είναι μεγαλύτερο σε ορισμένα στρώματα, τότε δεν είναι δυνατόν να εφαρμοστεί η μέθοδος της συνένωσης για την εξαγωγή ενός και μόνου αποτελέσματος. Στην περίπτωση αυτή, για την εκτίμηση του αποτελέσματος του μελετώμενου προσδιοριστή στα διάφορα στρώματα χρησιμοποιείται η μέθοδος της προτύπωσης.

4. ΕΞΟΥΔΕΤΕΡΩΣΗ ΤΩΝ ΣΥΓΧΥΤΩΝ ΜΕ ΤΗ ΜΕΘΟΔΟ ΤΗΣ ΔΙΑΣΤΡΩΜΑΤΩΣΗΣ

4.1. Μελέτες με κλειστούς πληθυσμούς: Επίπτωση-ποσοστό,* επιπολασμός

Στη δημογραφία, η κοόρτη⁷ είναι συνώνυμη του κλειστού πληθυσμού (closed population).^{5,6,12,19,21} Είναι ένα κλειστό σύνολο ατόμων, όπου η ιδιότητα του μέλους καθορίζεται από ένα συμβάν (event) σε μια συγκεκριμένη τοποχρονική περιοχή. Ο χαρακτηρισμός «κλειστός» σημαίνει ότι «απαγορεύεται» η έξοδος των μελών από τον πληθυσμό αυτόν. Η ιδιότητα του μέλους ενός κλειστού πληθυσμού δεν χάνεται ούτε με το θάνατο του μέλους («επιδημιολογική αθανασία»). Παράδειγμα κλειστού πληθυσμού είναι τα άτομα που γεννήθηκαν στην Ελλάδα το 2000.

Ανοικτός (open) ή δυναμικός (dynamic) πληθυσμός είναι ο πληθυσμός μιας πόλης ή μιας χώρας, οι νοσηλευόμενοι σ' ένα νοσοκομείο κ.ά.^{5,6,12,19} Τα μέλη του ανοι-

* Η *επίπτωση-ποσοστό* είναι το ποσοστό των προσωπο-στιγμών στην αρχή της παρακολούθησης (T_0 =επιστημονικός χρόνος) που εμφανίσε τις περιπτώσεις της νόσου κατά τη διάρκεια μιας ορισμένης χρονικής περιόδου.^{5,8,9,19,20} Ο όρος προσωπο-στιγμή (person-moment ή instance) εισήχθη για να δηλώσει ότι μια περίπτωση νοσήματος συμβαίνει σ' ένα πρόσωπο σε μια συγκεκριμένη χρονική στιγμή.^{5,9,19,20} *Σημειακός επιπολασμός*: Όταν η συχνότητα του νοσήματος που μελετάται δεν αφορά σε ενάρξεις του νοσήματος (νέες περιπτώσεις), που αποτελούν σημεία στο χρόνο (συμβάντα), αλλά καταστάσεις που έχουν διάρκεια, τότε το μέτρο συχνότητας ονομάζεται *επιπολασμός* και είναι ποσοστό που προκύπτει από ένα κοινό κλάσμα, το οποίο αριθμητή έχει τις υπάρχουσες (νέες και παλαιές) περιπτώσεις νοσήματος και παρονομαστή (τάξη αναφοράς) μια σειρά προσωπο-στιγμών.^{5,8,13,14,17-20,26,27} Η βάση μελέτης, στην περίπτωση αυτή, είναι η τομή ενός πληθυσμού και ο υπολογισμός γίνεται σε μια ορισμένη στιγμή του ημερολογιακού χρόνου. *Επιπολασμός περιόδου*: Στην περίπτωση αυτή, οι περιπτώσεις επιπολασμού αφορούν σε μια χρονική περίοδο, συνήθως μελλοντική.^{6,8,19,20,26} Ο επιπολασμός μιας νόσου αντιστοιχεί στο ποσοστό του πληθυσμού που έχει τη νόσο σε ένα ορισμένο σημείο στο χρόνο. Ο επιπολασμός είναι ποσοστό, καθώς ο αριθμητής περιέχεται στον παρονομαστή, οπότε λαμβάνει τιμές από 0-1. Επιπλέον, δεν έχει χρονική διάσταση, καθώς αναφέρεται σε μια συγκεκριμένη χρονική στιγμή.

κού πληθυσμού εναλλάσσονται στο χρόνο. Ο ανοικτός πληθυσμός, σε αντίθεση με τον κλειστό, δεν γηράσκει. Η ιδιότητα του μέλους του ανοικτού πληθυσμού προσδιορίζεται από μια κατάσταση (state) και διαρκεί όσο διαρκεί η κατάσταση αυτή. Ένα άτομο είναι μέλος του ανοικτού πληθυσμού της Αθήνας, όσο ζει στην Αθήνα και για το χρονικό διάστημα που ζει σ' αυτή. Χάνει, όμως, την ιδιότητα του μέλους εφόσον απομακρυνθεί από αυτή, κάτι που δεν ισχύει για τον κλειστό πληθυσμό.

Σε μελέτες όπου οι μελετώμενοι πληθυσμοί είναι κλειστοί (κοόρτες), όπως στις προγνωστικές μελέτες παρέμβασης, υπολογίζονται, όπως είναι γνωστό, επιπτώσεις-ποσοστά (ή ο επιπολασμός). Για το συμβολισμό κάθε στρώματος χρησιμοποιείται ο δείκτης i , που μπορεί να λάβει τιμές από ένα έως το συνολικό αριθμό των στρωμάτων. Ο συμβολισμός για το στρώμα i σ' ένα σύνολο στρωμάτων, σε μια μελέτη με κλειστούς πληθυσμούς όπου υπολογίζονται επιπτώσεις-ποσοστά, φαίνεται στον πίνακα 3. Η επίπτωση-ποσοστό για τους εκτεθειμένους είναι a_i/N_{1i} , ενώ για τους μη εκτεθειμένους είναι b_i/N_{0i} . Όπως προαναφέρθηκε, σκοπός είναι τα επιμέρους αποτελέσματα κατά στρώματα να συνενωθούν σ' ένα μόνο αποτέλεσμα. Τα μέτρα σχέσης που μπορούν να υπολογιστούν είναι η διαφορά των επιπτώσεων-ποσοστών (incidence-proportion difference, IPD) και ο λόγος των επιπτώσεων-ποσοστών (incidence-proportion ratio, IPR). Έτσι, είναι δυνατόν να υπολογιστεί η «συγκεντρωτική» διαφορά των επιπτώσεων-ποσοστών και ο «συγκεντρωτικός» ή η συνοπτική διαφορά μπορεί να εκτιμηθεί, από τα στρωματοποιημένα δεδομένα, με τον παρακάτω τύπο:

$$IPD_{MH} = \frac{\sum_i \frac{a_i N_{0i} - b_i N_{1i}}{T_i}}{\sum_i \frac{N_{1i} N_{0i}}{T_i}} \quad (1)$$

Με Σ συμβολίζεται το άθροισμα όλων των εμπειρικών τιμών των μέτρων σχέσης όλων των στρωμάτων. Ο

Πίνακας 3. Μελέτη με κλειστούς πληθυσμούς, όπου είναι δυνατός ο υπολογισμός των επιπτώσεων-ποσοστών.

	Εκτεθειμένοι	Μη εκτεθειμένοι	Σύνολο
Σειρά περιπτώσεων νοσήματος	a_i	b_i	M_{1i}
Μη περιπτώσεις (βασικές περιπτώσεις)	c_i	d_i	M_{0i}
Σύνολο ατόμων σε κίνδυνο	N_{1i}	N_{0i}	T

δείκτης MH αποτελεί τα αρχικά των ερευνητών Mantel και Haenszel, που ήταν οι πρώτοι οι οποίοι κατέληξαν στην ισότητα 1.

Ο «συγκεντρωτικός» (ή συνοπτικός) λόγος των επιπτώσεων-ποσοστών μπορεί να εκτιμηθεί, από τα στρωματοποιημένα δεδομένα, με τον παρακάτω τύπο:

$$IPR_{MH} = \frac{\sum_i \frac{a_i N_{0i}}{T_i}}{\sum_i \frac{b_i N_{1i}}{T_i}} \quad (2)$$

Στον πίνακα 4 φαίνονται τα δεδομένα μιας μελέτης²⁸ με κλειστούς πληθυσμούς με διαβητικούς ασθενείς, οι οποίοι ελάμβαναν τολβουταμίδη ή εικονικό φάρμακο (placebo).* Οι ερευνητές πίστευαν ότι η λήψη τολβουταμίδης μειώνει την πιθανότητα εμφάνισης επιπλοκών από το σακχαρώδη διαβήτη τύπου II, που ενδεχομένως οδηγούν στο θάνατο. Τα αποτελέσματα της μελέτης, ωστόσο, ήταν ακριβώς αντίθετα. Με βάση τα αδρά δεδομένα του πίνακα 4, η διαφορά των επιπτώσεων-ποσοστών θανάτου στις δύο ομάδες ήταν 0,045, γεγονός που σημαίνει ότι η επίπτωση-ποσοστό του θανάτου στους ασθενείς που ελάμβαναν τολβουταμίδη ήταν μεγαλύτερη από τους ασθενείς που ελάμβαναν placebo. Οι ερευνητές θεώρησαν ότι η διαφορά αυτή των επιπτώσεων-ποσοστών οφείλεται στην ύπαρξη συγχυτικών παραγόντων. Η ηλικία θεωρήθηκε ως ένας πιθανός συγχυτής. Από τύχη οι ασθενείς που ελάμβαναν τολβουταμίδη ήταν μεγαλύτερης ηλικίας από τους ασθενείς που ελάμβαναν εικονικό φάρμακο. Η ηλικιακή αυτή διαφορά φαίνεται στον πίνακα 4: Το 48% (98/204) των ασθενών που ελάμβαναν τολβουταμίδη ήταν >55 ετών, ενώ μόνο το 41% (85/205) των ασθενών που ελάμβαναν placebo ήταν >55 ετών. Προφανώς, όσο μεγαλύτερη είναι η ηλικία ενός ατόμου τόσο μεγαλύτερη είναι και η πιθανότητα θανάτου, κάτι που φαίνεται και στον πίνακα 4. Πιο συγκεκριμένα, για τους ασθενείς που ελάμβαναν placebo, η επίπτωση-ποσοστό θανάτου κατά τη χρονική περίοδο διεξαγωγής της μελέτης ήταν 0,1888 (16/85) για εκείνους >55 ετών και μόνο 0,042 (5/120) για εκείνους <55 ετών. Κάτι αντίστοιχο ισχύει και για τους ασθενείς που ελάμβαναν τολβουταμίδη. Επομένως, η μεγαλύτερη επίπτωση-ποσοστό του θανάτου για τους ασθενείς που ελάμβαναν τολβουταμίδη οφειλόταν, ως ένα βαθμό, στη συγχυτική επίδραση της ηλικίας. Για τη

* Placebo: (α) Φάρμακο, το οποίο χορηγείται σε ασθενή με σκοπό περισσότερο την ψυχολογική του στήριξη και ανακούφιση παρά τη θεραπεία της ίδιας του της αρρώστιας και (β) αδρανής ουσία, σε σύγκριση προς την οποία δοκιμάζεται η χρήση μιας άλλης (π.χ. φαρμάκου) σε ελεγχόμενες συνθήκες εργαστηρίου.²⁹

Πίνακας 4. Επιπτώσεις-ποσοστά θανάτου για ασθενείς που ελάμβαναν τολβουταμίδη ή placebo, ανάλογα με την ηλικία τους και συνολικά.²⁸

	Ηλικία (έτη)				Σύνολο	
	<55		≥55		Τολβουταμίδη	Placebo
	Τολβουταμίδη	Placebo	Τολβουταμίδη	Placebo		
Θάνατοι	8	5	22	16	30	21
Σύνολο ατόμων σε κίνδυνο	106	120	98	85	204	205
Επίπτωση-ποσοστό θανάτου	0,076	0,042	0,224	0,188	0,147	0,102
Διαφορά επιπτώσεων-ποσοστών	0,034		0,036		0,045	
Λόγος επιπτώσεων-ποσοστών	1,81		1,19		1,44	

διερεύνηση της υπόθεσης αυτής απαιτείται η διαστρωματική ανάλυση των δεδομένων, με βάση την ηλικία, και στη συνέχεια ο υπολογισμός της «συγκεντρωτικής» διαφοράς των επιπτώσεων-ποσοστών. Με βάση την ισότητα 1, η «συγκεντρωτική» διαφορά των επιπτώσεων-ποσοστών ισούται με:

$$IPD_{MH} = \frac{8 \cdot 120 - 5 \cdot 106}{\frac{226}{106 \cdot 120} + \frac{183}{226}} + \frac{22 \cdot 85 - 16 \cdot 98}{\frac{183}{98 \cdot 85} + \frac{183}{183}} = \frac{1903 + 1650}{56283 + 45519} = 0,035$$

Έπειτα από την εφαρμογή της διαστρωμάτωσης για την εξουδετέρωση της συγχυτικής επίδρασης της ηλικίας προκύπτει ότι η «συγκεντρωτική» διαφορά των επιπτώσεων-ποσοστών θανάτου είναι 0,035, μικρότερη δηλαδή από τη διαφορά των επιπτώσεων-ποσοστών θανάτου για αδρά δεδομένα, που είναι 0,045. Οι δύο επιμέρους διαφορές των επιπτώσεων-ποσοστών θανάτου στα δύο ηλικιακά στρώματα ήταν 0,034 για ασθενείς <55 ετών και 0,036 για ασθενείς >55 ετών. Εύκολα διαπιστώνεται ότι η «συγκεντρωτική» διαφορά των επιπτώσεων-ποσοστών θανάτου βρίσκεται μέσα στο στενό εύρος των δύο ηλικιακών στρωμάτων. Μαθηματικά, η «συγκεντρωτική» διαφορά των επιπτώσεων-ποσοστών θανάτου είναι ένας «σταθμισμένος» μέσος όρος των τιμών που λαμβάνονται στα επιμέρους στρώματα, οπότε η τιμή που λαμβάνει θα βρίσκεται πάντοτε μέσα στο εύρος των τιμών του υπολογιζόμενου μέτρου σχέσης στα επιμέρους στρώματα. Η αδρή εκτίμηση, ωστόσο, του αποτελέσματος δεν βρίσκεται μέσα στο εύρος αυτό, όπως φάνηκε και στο συγκεκριμένο παράδειγμα. Η αδρή διαφορά των επιπτώσεων-ποσοστών θανάτου (0,045) διαφέρει από τη «συγκεντρωτική» διαφορά των επιπτώσεων-ποσοστών θανάτου (0,035), επειδή η αδρή διαφορά αντανακλά όχι μόνο το αποτέλεσμα της τολβουταμίδης (που στη συγκεκριμένη μελέτη εκτιμήθηκε ως 0,035), αλλά και

το αποτέλεσμα της συγχυτικής δράσης της ηλικίας. Η αδρή διαφορά ήταν μεγαλύτερη από τη «συγκεντρωτική» διαφορά, που προκύπτει έπειτα από την εξουδετέρωση της συγχυτικής επίδρασης της ηλικίας, επειδή οι ασθενείς που ελάμβαναν τολβουταμίδη ήταν μεγαλύτεροι ηλικιακά από τους ασθενείς που ελάμβαναν placebo. Εάν οι ασθενείς που ελάμβαναν τολβουταμίδη ήταν νεότεροι από τους ασθενείς που ελάμβαναν placebo, τότε η συγχυτική δράση της ηλικίας θα οδηγούσε σε αντίθετα αποτελέσματα, η αδρή δηλαδή διαφορά θα ήταν μικρότερη από τη «συγκεντρωτική» διαφορά που θα προέκυπε έπειτα από τη διαστρωμάτωση.

Από τα δεδομένα του πίνακα 4 μπορεί, επίσης, να υπολογιστεί ο «συγκεντρωτικός» λόγος των επιπτώσεων-ποσοστών θανάτου χρησιμοποιώντας την ισότητα 2:

$$IPR_{MH} = \frac{\frac{8 \cdot 120}{226} + \frac{22 \cdot 85}{183}}{\frac{5 \cdot 106}{226} + \frac{16 \cdot 98}{183}} = \frac{4248 + 10219}{2345 + 8568} = 1,33$$

Ο αδρός λόγος των επιπτώσεων-ποσοστών θανάτου ήταν 1,44, ενώ μετά από την εξουδετέρωση της συγχυτικής επίδρασης της ηλικίας ο «συγκεντρωτικός» λόγος ήταν 1,33. Από τον πίνακα 4 προκύπτει πως οι λόγοι των επιπτώσεων-ποσοστών θανάτου στα δύο επιμέρους ηλικιακά στρώματα διαφέρουν σημαντικά μεταξύ τους. Πιο συγκεκριμένα, για τους ασθενείς >55 ετών ο λόγος των επιπτώσεων-ποσοστών θανάτου ήταν 1,19, ενώ για τους ασθενείς <55 ετών ο λόγος των επιπτώσεων-ποσοστών θανάτου ήταν 1,81. Το μεγάλο εύρος των δύο αυτών λόγων έχει ως αποτέλεσμα τόσο ο αδρός όσο και ο «συγκεντρωτικός» λόγος να βρίσκονται μέσα στο εύρος αυτό. Όταν οι τιμές του μέτρου σχέσης στα επιμέρους στρώματα είναι σχεδόν ίδιες, όπως συνέβη στην περίπτωση όπου το υπολογιζόμενο μέτρο σχέσης ήταν η διαφορά των επιπτώσεων-ποσοστών θανάτου, τότε είναι εύκολο να γίνει μια σχετικά ακριβής εκτίμηση του «συγκεντρωτικού» μέτρου σχέσης στηριζόμενοι στις τιμές του

μέτρου σχέσης στα επιμέρους στρώματα. Όταν, όμως, οι τιμές του μέτρου αποτελέσματος στα επιμέρους στρώματα διαφέρουν σημαντικά μεταξύ τους, όπως συνέβη στην περίπτωση όπου το υπολογιζόμενο μέτρο σχέσης ήταν ο λόγος των επιπτώσεων-ποσοστών, τότε η εκτίμηση του «συγκεντρωτικού» μέτρου σχέσης είναι πολύ δύσκολη.

Οι ισότητες 1 και 2 χρησιμοποιούνται για τον υπολογισμό ενός «συγκεντρωτικού» μέτρου σχέσης όταν ισχύει η υπόθεση ότι το αποτέλεσμα είναι ομοιόμορφο στα επιμέρους στρώματα. Αυτό σημαίνει ότι η ανάδειξη ενός εξωγενούς προσδιοριστή σε συγχυτή μιας σχέσης προϋποθέτει ότι ο προσδιοριστής αυτός δεν είναι ταυτόχρονα και τροποποιητής της μελετώμενης σχέσης. Στο παράδειγμα του πίνακα 4, η υπόθεση αυτή ισχύει όταν το υπολογιζόμενο μέτρο σχέσης ήταν η διαφορά των επιπτώσεων-ποσοστών. Και αυτό γιατί η διαφορά των επιπτώσεων-ποσοστών θανάτου στους ασθενείς >55 ετών ήταν 0,036, ενώ στους ασθενείς <55 ετών ήταν 0,034, οπότε το αποτέλεσμα στα δύο επιμέρους στρώματα ήταν ομοιόμορφο. Το ερώτημα που προκύπτει είναι πώς μπορεί να χρησιμοποιηθεί η υπόθεση αυτή για να εκτιμηθεί ο «συγκεντρωτικός» λόγος των επιπτώσεων-ποσοστών θανάτου, όταν οι λόγοι των επιπτώσεων-ποσοστών στα δύο επιμέρους ηλικιακά στρώματα είναι τόσο διαφορετικοί; Η υπόθεση αυτή δεν υποδηλώνει ότι το αποτέλεσμα θα πρέπει να είναι ακριβώς το ίδιο στα διάφορα στρώματα, αλλά λαμβάνει υπόψη της τη στατιστική μεταβλητότητα στα διάφορα στρώματα. Είναι δυνατόν να διεξαχθεί ένας στατιστικός έλεγχος, που καλείται έλεγχος της ετερογένειας (test of heterogeneity) ή έλεγχος της ομοιογένειας (test of homogeneity), για να καθοριστεί εάν η μεταβλητότητα του αποτελέσματος στα διάφορα στρώματα είναι συμβατή με την υπόθεση ότι το αποτέλεσμα είναι ομοιόμορφο. Πρέπει να σημειωθεί, πάντως, ότι η υπόθεση όπου το αποτέλεσμα είναι ομοιόμορφο είναι λαθεμένη στις περισσότερες περιπτώσεις. Είναι περισσότερο ρεαλιστικό να θεωρήσει κάποιος την υπόθεση αυτή ως μια «φανταστική άνεση», που διευκολύνει τον υπολογισμό ενός «συγκεντρωτικού» μέτρου σχέσης. Για τα δεδομένα του πίνακα 4, η μεταβλητότητα του λόγου των επιπτώσεων-ποσοστών στα δύο ηλικιακά στρώματα δεν είναι τόσο μεγάλη, έτσι ώστε να θεωρηθεί ότι παραβιάζεται η υπόθεση της ομοιομορφίας του αποτελέσματος.

4.2. Μελέτες με κλειστούς πληθυσμούς: Επίπτωση-πυκνότητα*

Τα δεδομένα μιας μελέτης με κλειστούς** πληθυσμούς, στην οποία υπολογίζονται επιπτώσεις-πυκνότητες, φαίνονται στον πίνακα 5.¹³ Όπως και στην περίπτωση των

επιπτώσεων-ποσοστών, είναι δυνατόν να υπολογιστεί ο «συγκεντρωτικός» (ή συνοπτικός) λόγος των επιπτώσεων-πυκνοτήτων (incidence-density ratio, IDR) ή η «συγκεντρωτική» διαφορά των επιπτώσεων-πυκνοτήτων (incidence-density difference, IDD). Πιο συγκεκριμένα, η «συγκεντρωτική» διαφορά των επιπτώσεων-πυκνοτήτων στα διαστρωματωμένα δεδομένα υπολογίζεται με την παρακάτω ισότητα:

$$IDD_{MH} = \frac{\sum_i \frac{a_i PT_{0i} - b_i PT_{1i}}{T_i}}{\sum_i \frac{PT_{1i} PT_{0i}}{T_i}} \quad (3)$$

Ο «συγκεντρωτικός» λόγος των επιπτώσεων-πυκνοτήτων υπολογίζεται με την παρακάτω ισότητα:

$$IDR_{MH} = \frac{\sum_i \frac{a_i PT_{0i}}{T_i}}{\sum_i \frac{b_i PT_{1i}}{T_i}} \quad (4)$$

Στον πίνακα 6 φαίνονται τα δεδομένα μιας μελέτης³¹ με κλειστούς πληθυσμούς, όπου υπολογίστηκαν οι επιπτώσεις-πυκνότητες θανάτου (ή θνησιμότητες) για παροντικούς και παρελθοντικούς λήπτες κλοζαπίνης, ενός φαρμάκου που χρησιμοποιείται για τη θεραπεία ανθεκτικών μορφών σχιζοφρένειας. Οι ερευνητές υπέθεσαν ότι η λήψη κλοζαπίνης μειώνει τη θνησιμότητα στους παροντικούς λήπτες. Για να διερευνηθεί η υπόθεση αυτή, χρησι-

Πίνακας 5. Δεδομένα μελέτης με κλειστούς πληθυσμούς, στην οποία υπολογίζονται επιπτώσεις-πυκνότητες.

	Εκτεθειμένοι	Μη εκτεθειμένοι	Σύνολο
Σειρά περιπτώσεων νοσήματος	a_i	b_i	M_i
Πληθυσμο-χρόνος	PT_{1i}	PT_{0i}	T_i

* Η *επίπτωση-πυκνότητα* (incidence-density) ποσοτικοποιεί την εμφάνιση ενός συμβάντος με τάξη αναφοράς μια πεπερασμένη ποσότητα πληθυσμο-χρόνου (population-time),^{5,9,13,14,19,20,30} Ο ιδιόμορφος αυτός χρόνος, δηλαδή ο πληθυσμο-χρόνος (που δεν είναι ημερολογιακός ή ηλικιακός), προκύπτει κατά την κίνηση ενός πληθυσμού (ανοικτού ή κλειστού) στον ημερολογιακό χρόνο και συνίσταται από άπειρο αριθμό προσωπο-στιγμών. Αποτελεί το άθροισμα των χρονικών περιόδων παρακολούθησης των μελών του πληθυσμού. Με δεδομένο ότι τάξη αναφοράς στο εμπειρικό αυτό μέτρο συχνότητας είναι ο πληθυσμο-χρόνος, η συχνότητα αφορά μόνο στην εμφάνιση συμβάντων και όχι καταστάσεων. Η επίπτωση-πυκνότητα δεν είναι καθαρός αριθμός, αλλά έχει αφενός αριθμητική τιμή και αφετέρου μονάδα μέτρησης, που είναι το αντίστροφο του χρόνου [π.χ. (έτος)⁻¹].

** Υπενθυμίζεται ότι κλειστοί πληθυσμοί χρησιμοποιούνται σχεδόν αποκλειστικά στις κλινικές δοκιμές. Στην περίπτωση αυτή, οι μελετώμενοι πληθυσμοί είναι κλειστοί και σχηματίζονται με επιλογή από τον πληθυσμό-πηγή, ο οποίος μπορεί να είναι είτε ανοικτός είτε κλειστός.⁷

Πίνακας 6. Θνησιμότητες για παροντικούς και παρελθοντικούς λήπτες κλοζαπίνης, ανάλογα με την ηλικία τους και συνολικά.³¹

	Ηλικία (έτη)				Σύνολο	
	10-54		55-94		Τωρινοί	Παρελθόντες
	Τωρινοί	Παρελθόντες	Τωρινοί	Παρελθόντες		
Θάνατοι	196	111	167	157	363	268
Πληθυσμο-χρόνος (ανθρωπο-έτη)	62.119	15.763	6085	2780	68.204	18.543
Θνησιμότητα ($\times 10^{-5}$ έτη ⁻¹)	315,5	704,2	2744	5647	532,2	1445
Διαφορά θνησιμοτήτων ($\times 10^{-5}$ έτη ⁻¹)		-388,7		-2903		-912,8
Λόγος θνησιμοτήτων		0,45		0,49		0,37

μποιήθηκε ως κατηγορία αναφοράς η εμπειρία των ασθενών που ελάμβαναν κλοζαπίνη κατά το παρελθόν. Για τον έλεγχο της συγχυτικής επίδρασης της ηλικίας, τα δεδομένα διαστρωματώθηκαν, με βάση την ηλικία, σε δύο ηλικιακές ομάδες (ασθενείς ηλικίας 10-54 ετών και ασθενείς ηλικίας 55-94 ετών).

Όπως συμβαίνει στο γενικό πληθυσμό, έτσι ακριβώς και στους ασθενείς της συγκεκριμένης μελέτης η θνησιμότητα ήταν πολύ μεγαλύτερη για τους ασθενείς μεγαλύτερης ηλικίας σε σχέση με τους νεαρότερους ασθενείς. Επιπλέον, διαπιστώθηκε σχέση μεταξύ της ηλικίας και του χρόνου λήψης (παρόν ή παρελθόν) της κλοζαπίνης. Στους παροντικούς λήπτες κλοζαπίνης, το 9% (6085/68204) του πληθυσμο-χρόνου ανήκε στη μεγαλύτερη ηλικιακή κατηγορία, ενώ στους παρελθοντικούς λήπτες κλοζαπίνης, το 15% (2780/18543) του πληθυσμο-χρόνου ανήκε στη μεγαλύτερη ηλικιακή κατηγορία. Η διαφορά αυτή ήταν αρκετή για να προκαλέσει σύγχυση στη μελετώμενη σχέση. Επειδή ο πληθυσμο-χρόνος για τους παρελθοντικούς λήπτες ήταν μεγαλύτερος, η διαφορά αυτή στην ηλικία είχε ως συνέπεια μικρότερη θνησιμότητα στους παροντικούς λήπτες. Σύμφωνα με τα αδρά δεδομένα του πίνακα 6, η θνησιμότητα ήταν μικρότερη στους παροντικούς λήπτες. Η αδρή διαφορά των θνησιμοτήτων ήταν $-0,009128$ (έτος)⁻¹. Ένα μέρος της διαφοράς αυτής οφείλεται στη συγχυτική δράση της ηλικίας. Η «συγκεντρωτική» διαφορά των θνησιμοτήτων, μετά από την εξουδετέρωση της συγχυτικής δράσης της ηλικίας, προκύπτει έπειτα από την εφαρμογή της ισότητας 3:

$$IDD_{MH} = \frac{196 \cdot 15763 - 111 \cdot 62119}{77882} + \frac{167 \cdot 2780 - 157 \cdot 6085}{8865} = \frac{62119 \cdot 15763 - 6085 \cdot 2780}{77882 + 8865} = -720 \times 10^{-5} \text{ (έτος)}^{-1}$$

Η τιμή που προκύπτει [$-0,0072$ (έτος)⁻¹], μετά από την εξουδετέρωση της ηλικίας, είναι αρκετά μικρότερη

από την τιμή που προκύπτει με τη χρήση των αδρών δεδομένων [$-0,009128$ (έτος)⁻¹], γεγονός αναμενόμενο έπειτα από τη διαφορά που διαπιστώθηκε στις ηλικιακές κατανομές των δύο μελετώμενων σειρών (παροντικοί και παρελθοντικοί λήπτες κλοζαπίνης). Ο βαθμός της σύγχυσης ήταν μέτριος, παρά το ότι η ηλικία είναι ισχυρός προσδιοριστής, ακριβώς επειδή η διαφορά στην ηλικιακή κατανομή μεταξύ των τωρινών και των παρελθόντων ληπτών κλοζαπίνης ήταν μέτρια. Πρέπει, πάντως, να σημειωθεί ότι η εναπομείνουσα «συγκεντρωτική» διαφορά [$-0,0072$ (έτος)⁻¹] δεν εξουδετερώνει πλήρως τη συγχυτική δράση της ηλικίας κι αυτό γιατί οι ασθενείς χωρίστηκαν σε δύο πολύ μεγάλες ηλικιακές ομάδες. Ενδεχομένως, μέσα σ' αυτές τις δύο ομάδες να υπάρχει ακόμη η συγχυτική δράση της ηλικίας. Η πλήρης εξουδετέρωση της συγχυτικής δράσης της ηλικίας ίσως να επιτυχανόταν με τη διαίρεση των ασθενών σε περισσότερα από δύο στρώματα, με βάση την ηλικία.

Ο «συγκεντρωτικός» λόγος των θνησιμοτήτων για τα δεδομένα του πίνακα 6 υπολογίζεται με βάση την ισότητα 4:

$$IDR_{MH} = \frac{196 \cdot 15763}{77882} + \frac{167 \cdot 2780}{8865} = \frac{111 \cdot 62119}{77882} + \frac{157 \cdot 6085}{8865} = \frac{39,67 + 52,37}{88,53 + 107,77} = 0,47$$

Η τιμή 0,47 υποδηλώνει ότι, μετά από την εξουδετέρωση της συγχυτικής δράσης της ηλικίας στα δύο επιμέρους ηλικιακά στρώματα, η θνησιμότητα στους παρελ-

Πίνακας 7. Μελέτη ασθενών-μαρτύρων.

	Εκτεθειμένοι	Μη εκτεθειμένοι	Σύνολο
Περιπτώσεις νοσήματος	a _i	b _i	M _{1i}
«Μάρτυρες»	c _i	d _i	M _{0i}
Σύνολο	N _{1i}	N _{0i}	T _i

θοντικούς λήπτες ήταν διπλάσια σε σχέση με τη θνησιμότητα στους παροντικούς λήπτες.

4.3. Μελέτες «ασθενών-μαρτύρων»

Στον πίνακα 7 φαίνονται τα δεδομένα μιας μελέτης «ασθενών-μαρτύρων»,* όπου χρησιμοποιείται ο συμβολισμός i για ένα στρώμα της διαστρωματικής ανάλυσης. Ο «συγκεντρωτικός» λόγος των οιοει επιπτώσεων-πυκνοτήτων υπολογίζεται με την παρακάτω ισότητα:

$$OR_{MH} = \frac{\sum_i \frac{a_i d_i}{T_i}}{\sum_i \frac{b_i c_i}{T_i}} \quad (5)$$

Τα δεδομένα στον πίνακα 8 προέρχονται από μια μελέτη «ασθενών-μαρτύρων»³² με παιδιά με συγγενή καρδιακή ανεπάρκεια. Σκοπός της μελέτης ήταν να διερευνηθεί η σχέση μεταξύ της λήψης σπερματοκτόνου από τη μητέρα πριν από τη σύλληψη και της συχνότητας εμφάνισης του συνδρόμου Down στο υποσύνολο των παιδιών που είχαν και συγγενή καρδιακή ανεπάρκεια και σύνδρομο Down. Τα παιδιά με συγγενή καρδιακή ανεπάρκεια ήταν >300, ενώ εκείνα που έπασχαν και από σύνδρομο Down ήταν μόλις 16. Χρησιμοποιώντας τα αδρά δεδομένα της μελέτης, βρέθηκε ότι ο αδρός λόγος των οιοει επιπτώσεων-πυκνοτήτων^{12,20} ήταν 3,50. Οι ερευνητές διερεύνησαν την τυχόν συγχυτική επίδραση της ηλικίας στη μελετώμενη σχέση της λήψης σπερματοκτόνου και του συνδρόμου Down. Εφαρμόζοντας την ισότητα 5 για τα δεδομένα του πίνακα 8, προκύπτει ότι ο «συγκεντρωτικός» λόγος των οιοει επιπτώσεων πυκνοτήτων ισούται με:

$$OR_{MH} = \frac{\frac{3 \cdot 1059}{1175} + \frac{1 \cdot 86}{95}}{\frac{104 \cdot 9}{1175} + \frac{5 \cdot 3}{95}} = \frac{2704 + 0,905}{0,797 + 0,158} = 3,78$$

Έτσι, ο «συγκεντρωτικός» λόγος των οιοει επιπτώσεων-πυκνοτήτων ήταν λίγο μεγαλύτερος από τον αδρό λόγο των οιοει επιπτώσεων-πυκνοτήτων, γεγονός που υποδηλώνει τη μέτρια συγχυτική δράση της ηλικίας.

* Ο όρος μελέτες «ασθενών-μαρτύρων» είναι εσφαλμένος όχι μόνο ως διατύπωση, αλλά, κυρίως, από άποψη περιεχομένου. Ο όρος case-referent studies εισήχθη από τον Miettinen το 1976 και είναι ετυμολογικά ορθότερος από τον όρο μελέτες «ασθενών-μαρτύρων», λόγω της λαθεμένης χρήσης του όρου «μάρτυρες». Ο όρος μελέτες «ασθενών-μαρτύρων» είναι εσφαλμένος, καθώς μια ομάδα ασθενών από μια συγκεκριμένη νόσο (που μπορεί να είναι εκτεθειμένοι ή όχι στον προσδιοριστή που μελετάται) συγκρίνεται με μια ομάδα υγιών ως προς τη μελετώμενη νόσο (που μπορεί και αυτοί να είναι εκτεθειμένοι ή όχι).⁹

5. ΔΙΑΣΤΡΩΜΑΤΩΣΗ ΣΤΗΝ ΠΕΡΙΠΤΩΣΗ ΔΥΟ Ή ΠΕΡΙΣΣΟΤΕΡΩΝ ΣΥΓΧΥΤΩΝ

Στην παρούσα ανασκόπηση, για λόγους διευκόλυνσης, χρησιμοποιήθηκαν παραδείγματα με ένα μόνο δυνητικό συγχυτή, οπότε τα στρώματα που προέκυπταν κάθε φορά ήταν πολύ λίγα. Διαστρωμάτωση, βεβαίως, μπορεί να πραγματοποιηθεί και με δύο ή περισσότερους δυνητικούς συγχυτές. Εάν υποθεθεί ότι σκοπός είναι ο έλεγχος της σύγχυσης που προκαλείται από το φύλο και την ηλικία συγχρόνως, και η ηλικία των μελετώμενων ατόμων χωρίζεται σε πέντε ομάδες, τότε από το συνδυασμό του φύλου και της ηλικίας προκύπτουν δέκα στρώματα. Προφανώς, και στην περίπτωση αυτή ισχύουν οι ισότητες που περιγράφηκαν αναλυτικά παραπάνω. Όσο μεγαλύτερος, όμως, είναι ο αριθμός των δυνητικών συγχυτών τόσο μεγαλύτερος είναι και ο αριθμός των στρωμάτων που προκύπτουν. Παραδείγματος χάριν, εάν οι δυνητικοί συγχυτές είναι πέντε και καθένας έχει τρεις κατηγορίες, τότε στη διαστρωματική ανάλυση προκύπτουν $3 \times 3 \times 3 \times 3 \times 3 = 243$ στρώματα. Όταν δημιουργηθεί μεγάλος αριθμός στρωμάτων, τότε σε κάθε στρώμα θα υπάρχει μικρός αριθμός παρατηρήσεων, με αποτέλεσμα η πληροφορία που θα παρέχει το κάθε στρώμα να είναι πολύ μικρή. Όταν, μάλιστα, ο αριθμός των δεδομένων σε κάποια στρώματα γίνει πολύ μικρός, τότε, ενδεχομένως, τα στρώματα αυτά δεν θα προσφέρουν πληροφορία στο συνολικό αποτέλεσμα, οπότε με τον τρόπο αυτό χάνονται πολύτιμες πληροφορίες της μελέτης. Έτσι, η διαστρωματική ανάλυση δεν είναι ιδιαίτερα πρακτική και εύχρηστη όταν πρόκειται να ελεγχθεί η σύγχυση που οφείλεται σε πολλούς δυνητικούς συγχυτές.

6. ΣΗΜΑΣΙΑ ΤΗΣ ΔΙΑΣΤΡΩΜΑΤΩΣΗΣ

Όπως αναφέρθηκε, η διαστρωμάτωση δεν είναι ιδιαίτερα πρακτική και εύχρηστη όταν οι δυνητικοί συγχυτές μιας μελετώμενης σχέσης είναι αρκετοί. Στην περίπτωση αυτή προτιμάται η *πολυμεταβλητή ανάλυση*** (multivariable analysis). Η εφαρμογή, ωστόσο, της πολυμεταβλητής ανάλυσης απαιτεί τη χρήση ηλεκτρονικού υπολογιστή και των κατάλληλων στατιστικών προγραμμάτων. Αντίθετα, η διαστρωματική ανάλυση δεν απαιτεί εξειδικευμένα στατιστικά προγράμματα, καθώς οι ισότητες που προαναφέρθηκαν απαιτούν υπολογισμούς που πραγματοποιούνται με απλή αριθμομηχανή ή ακόμη και

** Ανάλυση δεδομένων: (α) Ανάλυση παλινδρόμησης, όπου μέτρο αποτελέσματος είναι ο συντελεστής β και (β) συνοπτική ανάλυση (tabular analysis).

Πίνακας 8. Παιδιά με συγγενή καρδιακή ανεπάρκεια και σύνδρομο Down και υγιείς «μάρτυρες» (βασικές περιπτώσεις), ανάλογα με τη λήψη σπερματοκτόνου από τη μητέρα πριν από τη σύλληψη και την ηλικία της μητέρας κατά τον τοκετό.³²

	Ηλικία μητέρας (έτη) κατά τον τοκετό					
	<35			≥35		
	Λήψη σπερματοκτόνου			Λήψη σπερματοκτόνου		
	Ναι	Όχι	Σύνολο	Ναι	Όχι	Σύνολο
Περιπτώσεις	3	9	12	1	3	4
«Μάρτυρες»	104	1059	1163	5	86	91
Σύνολο	107	1068	1175	6	89	95
Λόγος των οιοει επιπτώσεων-πυκνοτήτων		3,39			5,73	

με το χέρι. Η διαστρωματική ανάλυση πρέπει πάντοτε να προτιμάται έναντι της πολυμεταβλητής ανάλυσης για τον έλεγχο των συγχυτών. Η πολυμεταβλητή ανάλυση θα πρέπει να χρησιμοποιείται μόνο όταν οι πιθανοί συγχυτές μιας σχέσης είναι πολλοί, οπότε είναι πρακτικά αδύνατη η διαστρωματική ανάλυση των δεδομένων. Τα κύρια πλεονεκτήματα της διαστρωμάτωσης έναντι της πολυμεταβλητής ανάλυσης είναι τα εξής:¹³

- Με τη διαστρωμάτωση, ο ερευνητής έχει μια σαφή εικόνα της κατανομής των μελετώμενων ατόμων ανάλογα με τον προσδιοριστή, την έκβαση και τους πιθανούς συγχυτές. Αξιοσημείωτα χαρακτηριστικά των κατανομών γίνονται άμεσα αντιληπτά. Οι κατανομές αυτές δεν είναι άμεσα αντιληπτές όταν πραγματοποιείται η πολυμεταβλητή ανάλυση.
- Όχι μόνο ο ερευνητής αλλά και οι αναγνώστες μιας μελέτης είναι δυνατόν να έχουν μια ευκρινή εικόνα της κατανομής των μελετώμενων ατόμων. Μάλιστα,

από τους αναλυτικούς πίνακες της διαστρωματικής ανάλυσης, οι αναγνώστες μπορούν να ελέγξουν τους διάφορους υπολογισμούς ή και να υπολογίσουν μόνοι τους τα «συγκεντρωτικά» μέτρα σχέσης.

- Η διαστρωμάτωση απαιτεί λιγότερες υποθέσεις, οπότε μειώνεται η πιθανότητα σφάλματος.

Ο πρωταρχικός έλεγχος για την ύπαρξη τυχόν συγχυτών απαιτείται να πραγματοποιείται με τη χρήση της διαστρωματικής ανάλυσης των δεδομένων. Σπάνια, η χρήση της πολυμεταβλητής ανάλυσης μεταβάλλει την ερμηνεία των αποτελεσμάτων που προκύπτουν από τη χρήση της διαστρωμάτωσης. Ακόμη και αν είναι λογικό να πραγματοποιηθεί πολυμεταβλητή ανάλυση, όπως στην περίπτωση όπου ο αριθμός των συγχυτών είναι μεγάλος, θα πρέπει αρχικά να πραγματοποιηθεί διαστρωματική ανάλυση, τουλάχιστον για τους συγχυτές που θεωρείται ότι έχουν τη μεγαλύτερη σημασία.

ABSTRACT

Stratification of epidemiological data

P. GALANIS, L. SPAROS

Laboratory of Clinical Epidemiology, School of Nursing, University of Athens, Athens, Greece

Archives of Hellenic Medicine 2006, 23(6):626–637

Stratification constitutes the mainstay of epidemiologic analysis. Even with studies that ultimately require more complicated analyses, stratification is an important interim tool. It familiarizes the investigator with distributions of key variables and patterns in the data in ways that other approaches cannot provide. Stratification means that data are separated into categories or strata. For example stratification by sex or age means that data are separated into men and women or into categories by age. Several analytic concerns motivate stratification, the most prominent of which is evaluation and control of confounding. A simple definition of confounding would be the confusion, or mixing, of effects. This definition implies that the effect of the determinant under study is mixed together with the effect of a confounder, leading to a systematic bias. When data are stratified by confounder, for example into men and women, each stratum would be free of confounding. Thus, if the correlation between exposure and disease is analyzed separately, for example in men and women, each category

of gender would give an estimation of the result of exposure, independent of gender. Stratification is also used to evaluate and describe effect-measure modification. In addition the study of biologic interactions between two factors is most easily accomplished using stratification methods.

Key words: Confounder, Confounding, Determinant, Effect measure modification, Stratification

Βιβλιογραφία

1. ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Διαστρωματική ανάλυση δεδομένων. *Αρχ Ελλ Ιατρ* 2005, 21:378–384
2. ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Τροποποίηση του μέτρου αποτελέσματος και σύγκριση στην εφαρμοσμένη ιατρική έρευνα. *Αρχ Ελλ Ιατρ* 2005, 22:170–177
3. ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Στατιστική αλληλεπίδραση και τροποποίηση του μέτρου αποτελέσματος. *Αρχ Ελλ Ιατρ* 2005, 21:137–147
4. MIETTINEN OS. *Advanced data analysis. Course notes.* Erasmus University Medical School, Rotterdam, The Netherlands, 1995
5. ΣΠΑΡΟΣ Λ. *Μεταεπιδημιολογία.* Εκδόσεις ΒΗΤΑ, Αθήνα, 2001
6. AHLBOM A, NORELL S. *Εισαγωγή στη σύγχρονη επιδημιολογία.* Ιατρικές Εκδόσεις Λίτσας, Αθήνα, 1992
7. ΣΠΑΡΟΣ Λ, ΓΑΛΑΝΗΣ Π, ΖΑΧΟΣ Ι, ΤΣΙΛΙΔΗΣ Κ. *Επιδημιολογία Ι.* Εκδόσεις ΒΗΤΑ, Αθήνα, 2004
8. ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Η έννοια του αποδοτέου κλάσματος στην εφαρμοσμένη ιατρική έρευνα. *Αρχ Ελλ Ιατρ* 2005, 22:157–169
9. MIETTINEN OS. *Theoretical epidemiology. Principles of occurrence research in medicine.* John Wiley & Sons, New York, 1985
10. GAIL MH. Stratification. In: Armitage P, Colton T (eds) *Encyclopedia of biostatistics.* John Wiley & Sons, New York, 1998:43–46
11. ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Βιολογική και στατιστική αλληλεπίδραση. *Αρχ Ελλ Ιατρ* 2004, 21:123–136
12. ΓΑΛΑΝΗΣ Π. Αλληλεπίδραση παραγόντων κινδύνου στην επιδημιολογία. Μεταπτυχιακή διπλωματική εργασία. Αθήνα, 2003
13. ROTHMAN KJ. *Epidemiology: An introduction.* Oxford University Press, New York, 2002
14. ROTHMAN KJ, GREENLAND S. *Modern epidemiology.* 2nd ed. Lippincott Williams & Wilkins, Philadelphia, 1998
15. STROM BL. *Pharmacoepidemiology.* 2nd ed. John Wiley & Sons, New York, 1994
16. ROTHMAN KJ, MONSON RR. Survival in trigeminal neuralgia. *J Chron Dis* 1973, 26:303–309
17. SELVIN S. *Statistical analysis of epidemiologic data.* 2nd ed. Oxford University Press, New York, 1996:3–21
18. MORRISON AS. *Screening in chronic disease.* Oxford University Press, New York, 1985:25–29
19. ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Μέτρα συχνότητας των νοσημάτων. *Αρχ Ελλ Ιατρ* 2005, 22:178–191
20. ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Βασικές αρχές ανάλυσης των επιδημιολογικών δεδομένων. *Αρχ Ελλ Ιατρ* 2005, 22:146–156
21. ROTHMAN KJ. *Modern epidemiology.* 1st ed. Little Brown & Co, Boston, 1986
22. ΤΡΙΧΟΠΟΥΛΟΣ Δ. *Γενική και κλινική επιδημιολογία.* Εκδόσεις Παρισιάνος, Αθήνα, 2002:189–196, 220–227
23. FREEMAN J, GOLDMANN DA, MCGOWAN JE. Confounding and the analysis of multiple variables in hospital epidemiology. *Infect Control* 1987, 8:465–473
24. FREEMAN J, GOLDMANN DA, MCGOWAN JE. Methodologic issues in hospital epidemiology. IV. Risk ratios, confounding, effect modification, and the analysis of multiple variables. *Rev Infect Dis* 1988, 10:1118–1141
25. CONSONNI D, BERTAZZI PA, ZOCCHETTI C. Why and how to control for age in occupational epidemiology. *Occup Environ Med* 1997, 54:772–776
26. ΣΠΑΡΟΣ ΛΔ. Η έννοια της νοσηρότητας. *Αρχ Ελλ Ιατρ* 2001, 18:303–311
27. BEAGLEHOLE R, BONITA R, KJELLSTROM T. *Basic epidemiology.* World Health Organization, Geneva, 1993:1324
28. UNIVERSITY GROUP DIABETES PROGRAM. A study of the effects of the hypoglycemic agents on vascular complications in patients with adult onset diabetes. *Diabetes* 1970, 19(Suppl 2):747–830
29. ΜΠΑΜΠΙΝΙΩΤΗΣ Γ. *Λεξικό της νέας Ελληνικής γλώσσας.* Κέντρο Λεξικολογίας, Αθήνα, 1998:1433
30. FOWLER J, JARVIS P, CHEVANNES M. *Practical statistics for nursing and health care.* John Wiley & Sons, New York, 2002:47–50
31. WALKER AM, LANZA LL, ARELLANO F, ROTHMAN KJ. Mortality in current and former users of clozapine. *Epidemiology* 1997, 8:671–677
32. ROTHMAN KJ. Spermicide use and Down syndrome. *Am J Public Health* 1982, 72:399–401

Corresponding author:

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