

# ΕΦΑΡΜΟΣΜΕΝΗ ΙΑΤΡΙΚΗ ΕΡΕΥΝΑ APPLIED MEDICAL RESEARCH

## Συστηματικά σφάλματα στις επιδημιολογικές μελέτες

1. Εισαγωγή
2. Συστηματικά σφάλματα επιλογής
  - 2.1. Συστηματικό σφάλμα εθελοντικής συμμετοχής ή αυτοεπιλογής
  - 2.2. «Το φαινόμενο του υγιή εργάτη»
3. Συστηματικά σφάλματα πληροφορίας
  - 3.1. Διαφορική δυσταξινόμηση
  - 3.2. Μη διαφορική δυσταξινόμηση
4. Σύγχυση
5. Εξουδετέρωση της σύγχυσης
6. Σύνοψη

ΑΡΧΕΙΑ ΕΛΛΗΝΙΚΗΣ ΙΑΤΡΙΚΗΣ 2007, 24(4):373-388  
ARCHIVES OF HELLENIC MEDICINE 2007, 24(4):373-388

Π. Γαθάνης,  
Λ. Σπάρος

Εργαστήριο Κλινικής Επιδημιολογίας,  
Τμήμα Νοσηλευτικής, Πανεπιστήμιο  
Αθηνών, Αθήνα

Systematic bias in epidemiological  
studies

Abstract at the end of the article

### Λέξεις ευρετηρίου

Δυσταξινόμηση  
Σύγχυση  
Συγχυτής  
Συστηματικό σφάλμα επιλογής  
Συστηματικό σφάλμα πληροφορίας

Υποβλήθηκε 10.2.2006

Εγκρίθηκε 1.6.2006

### 1. ΕΙΣΑΓΩΓΗ

Οι επιδημιολογικές μελέτες διακρίνονται σε αιτιολογικές (ή αιτιογνωστικές), διαγνωστικές και προγνωστικές. Οι προγνωστικές διαιρούνται σε περιγραφικές και σε προγνωστικές μελέτες παρέμβασης. Οι διαγνωστικές και οι περιγραφικές μελέτες παρέμβασης είναι περιγραφικές, ενώ οι αιτιολογικές και οι προγνωστικές μελέτες παρέμβασης είναι αιτιακές. Η ανασκόπηση αυτή θα περιοριστεί μόνο στα σφάλματα που παρατηρούνται στις αιτιολογικές μελέτες.

Σε κάθε επιδημιολογική μελέτη πραγματοποιούνται μετρήσεις.\* Για να είναι αξιόπιστη και έγκυρη η μέτρηση,\*\* θα πρέπει στο εξαγόμενο τελικό αποτέλεσμα να μην υπεισέρχονται σφάλματα.<sup>1</sup> Τα σφάλματα σε μια επιδημιολογική και γενικότερα σε μια επιστημονική μελέ-

τη διακρίνονται σε συστηματικά και τυχαία και οφείλονται στη μετρική και βιολογική μεταβλητότητα. Η φύση των δύο σφαλμάτων (τυχαίων και συστηματικών) που οφείλονται στη βιολογική μεταβλητότητα είναι τελείως διαφορετική από τη φύση των σφαλμάτων που οφείλονται στη μετρική μεταβλητότητα. Για να υπολογιστεί η βιολογική μεταβλητότητα θα πρέπει η μετρική μεταβλητότητα να αφαιρεθεί από τη «μετρούμενη μεταβλητότητα».

Στην πλειοψηφία των αιτιολογικών ή αιτιογνωστικών επιδημιολογικών μελετών, σκοπός είναι ο υπολογισμός ενός μέτρου σχέσης,\*\*<sup>3</sup> όπως π.χ. είναι ο λόγος των επιπτώσεων-πυκνοτήτων (*incidence-density ratio*).<sup>1</sup> Σε μια

\* Μερικοί συγγραφείς θεωρούν ότι η Επιδημιολογία είναι μια άσκηση στη μέτρηση και επιδημιολόγοι θεωρούνται οι επιστήμονες υγείας που γνωρίζουν να πραγματοποιούν μετρήσεις.

\*\* *Μέτρηση* είναι ο εμπειρικός καθορισμός των τιμών μιας αριθμητικής συνάρτησης, δηλαδή μιας ποσοτικής έννοιας. Ο όρος χρησιμοποιείται και με ευρύτερη έννοια, για να περιλάβει την κατανομή αντικειμένων μιας κατηγορίας σε τάξεις (ποιοτικές έννοιες). Η διαφορά μεταξύ ποιοτικών και ποσοτικών εννοιών δεν είναι διαφορά στη φύση, αλλά εννοιολογική διαφορά (γλωσσική διαφορά). Αρίθμηση είναι η μέθοδος, με την οποία καθορίζεται (προσδιορίζεται) ο απόλυτος αριθμός της τάξης αναφοράς.

<sup>3</sup> Εάν συγκριθούν επιπτώσεις-ποσοστά ή επιπτώσεις-πυκνότητες μεταξύ εκτεθειμένων σ' έναν παράγοντα ατόμων και μη εκτεθειμένων, τότε δεν είναι βέβαιο ότι οι διαφορές στις επιπτώσεις-ποσοστά ή τις επιπτώσεις-πυκνότητες οφείλονται στην έκθεση στο μελετώμενο παράγοντα. Αντίθετα, οι διαφορές αυτές μπορεί να οφείλονται σε άλλα χαρακτηριστικά, ως προς τα οποία διαφέρουν τα εκτεθειμένα και μη εκτεθειμένα άτομα. Είναι δυνατόν να εντοπιστούν ορισμένα από τα χαρακτηριστικά αυτά ως προς τα οποία διαφέρουν τα εκτεθειμένα και μη εκτεθειμένα άτομα, αλλά πάντα υπάρχει ο κίνδυνος να διαφεύγουν ορισμένα. Το ιδανικό θα ήταν τα ίδια ακριβώς άτομα να ανήκουν την ίδια χρονική στιγμή και στις δύο κατηγορίες, δηλαδή τόσο στους εκτεθειμένους όσο και στους μη εκτεθειμένους. Πρόκειται ουσιαστικά για εξομοίωση του κάθε ατόμου με τον ίδιο του τον εαυτό. Σε μια τέτοια περίπτωση, η μόνη διαφορά μεταξύ των δύο κατηγοριών, εκτεθειμένων και μη εκτεθειμένων, θα ήταν ο μελετώμενος παράγοντας κινδύνου, οπότε θα ήταν δυνατόν να υπολογιστούν τα μέτρα αποτελέσματος. Επειδή όμως η κατάσταση αυτή

μελέτη, π.χ., στην οποία αντικείμενο είναι να υπολογιστεί ο λόγος των επιπτώσεων-πυκνοτήτων του καρκίνου του πνεύμονα σε εκείνους που καπνίζουν σε σχέση με εκείνους που δεν καπνίζουν, η εμπειρική τιμή του λόγου που προκύπτει αποτελεί εκτίμηση της πραγματικής (θεωρητικής, αφηρημένης) τιμής του λόγου των επιπτώσεων-πυκνοτήτων (εκτίμηση της παραμέτρου της σχέσης).<sup>1</sup> Όσο πλησιέστερα είναι η τιμή που προκύπτει από τη μελέτη προς την πραγματική τιμή, τόσο μικρότερο είναι το σφάλμα που υπεισέρχεται και τόσο πιο *αυθεντική\** (accurate) είναι η μελέτη. Αντίθετα, εάν η τιμή που προκύπτει από τη μελέτη διαφέρει σημαντικά από την πραγματική τιμή, τότε η μελέτη θεωρείται μη αυθεντική (inaccurate). Το πρόβλημα που προκύπτει είναι ότι δεν είναι γνωστή η πραγματική (θεωρητική) τιμή του λόγου των επιπτώσεων-πυκνοτήτων του καρκίνου του πνεύμονα στις δύο κατηγορίες (ενδεικτική και αναφοράς) του θεωρητικού τομέα. Το μόνο γνωστό είναι το αποτέλεσμα μιας συγκεκριμένης (σε τόπο και χρόνο) μελέτης. Προφανώς, αυτό δεν ισχύει μόνο στη συγκεκριμένη περι-

πτωση, καθώς οι πραγματικές τιμές των μέτρων σχέσης, αλλά και των μέτρων συχνότητας, είναι πάντα άγνωστες. Επειδή, ακριβώς, οι πραγματικές τιμές είναι άγνωστες, δεν είναι δυνατόν να καθοριστεί με ακρίβεια το συστηματικό σφάλμα σε μια μελέτη. Για το λόγο αυτόν, ο σχεδιασμός των επιδημιολογικών μελετών απαιτεί ιδιαίτερη προσοχή ώστε να περιορίζονται, όσο το δυνατόν περισσότερο, τα συστηματικά σφάλματα και να διασφαλίζεται η εγκυρότητα της μελέτης.

Σε μια επιδημιολογική μελέτη υπάρχουν δύο είδη σφάλματος: (α) το τυχαίο σφάλμα και (β) το συστηματικό σφάλμα.<sup>4,7-10</sup> Αν και υπάρχουν αρκετά είδη συστηματικών σφαλμάτων, παραδοσιακά κατατάσσονται σε τρεις μεγάλες κατηγορίες: (α) στα *συστηματικά σφάλματα επιλογής*, (β) στα *συστηματικά σφάλματα πληροφορίας* και (γ) στη *σύγχυση*.<sup>1,5,10,11</sup>

## 2. ΣΥΣΤΗΜΑΤΙΚΑ ΣΦΑΛΜΑΤΑ ΕΠΙΛΟΓΗΣ

Είναι προφανές ότι, σε μια επιδημιολογική μελέτη, δεν είναι δυνατόν να συμπεριληφθούν όλα τα άτομα του πληθυσμού που πάσχουν από ένα συγκεκριμένο νόσημα ή όλα τα άτομα που ανήκουν σε μια συγκεκριμένη κατηγορία του προσδιοριστή. Έτσι, επιλέγονται ορισμένα μόνο άτομα, τα οποία αποτελούν το μελετώμενο πληθυσμό (study population).

Τα *συστηματικά σφάλματα επιλογής*<sup>1,10,12</sup> (systematic selection biases)\*\* σε μια μελέτη προέρχονται από τις διαδικασίες που χρησιμοποιούνται για την επιλογή των συμμετεχόντων και από τους παράγοντες που επηρεάζουν τη συμμετοχή στη μελέτη. Υφίστανται όταν η σχέση μεταξύ του μελετώμενου προσδιοριστή\*\*\* και του νοσή-

είναι ανέφικτη, είναι γνωστή ως *υποθετικές προτάσεις του μη πραγματικού* (counterfactuals).<sup>2-6</sup> Στην ιδανική περίπτωση των υποθετικών προτάσεων του μη πραγματικού, κάθε άτομο ανήκει την ίδια ακριβώς χρονική στιγμή στην κατηγορία τόσο των εκτεθειμένων όσο και των μη εκτεθειμένων. Έτσι, είναι δυνατόν να συγκριθούν οι επιπτώσεις-ποσοστά μεταξύ εκτεθειμένων και μη εκτεθειμένων. Η διαφορά στις επιπτώσεις-ποσοστά αντικατοπτρίζει το αποτέλεσμα της έκθεσης στο μελετώμενο παράγοντα. Εάν π.χ. 100 άτομα εκτεθούν σ' έναν παράγοντα κινδύνου και τα 25 εμφανίσουν τη νόσο σε χρονικό διάστημα ενός έτους, τότε η επίπτωση-ποσοστό είναι 25/100=0,25 σε ένα έτος. Αντικείμενο της μελέτης είναι να συγκριθεί αυτή η επίπτωση-ποσοστό με εκείνη που θα υπολογιζόταν αν τα ίδια εκατό άτομα, την ίδια ακριβώς χρονική περίοδο και κάτω από τις ίδιες ακριβώς συνθήκες, δεν ήταν εκτεθειμένα στο μελετώμενο παράγοντα κινδύνου. Εάν στην κατηγορία των μη εκτεθειμένων εμφανιζαν τη νόσο 10 άτομα, τότε η επίπτωση-ποσοστό θα ήταν 10/100=0,10 σε ένα έτος. Αυτή η διαφορά 0,25-0,10=0,15 στις επιπτώσεις-ποσοστά -ή, διαφορετικά, 25-10=15 περιπτώσεις νοσήματος σε σύνολο 100 ατόμων στη διάρκεια ενός έτους- αντικατοπτρίζει το αποτέλεσμα της έκθεσης στο μελετώμενο παράγοντα. Τα *μέτρα αποτελέσματος* (measures of effect) υπολογίζονται στις περιπτώσεις εκείνες, κατά τις οποίες η έκθεση και η μη έκθεση σε έναν προσδιοριστή αφορούν στον ίδιο πληθυσμό, ενώ εκτιμούν το αποτέλεσμα της έκθεσης (σε έναν προσδιοριστή) στη συχνότητα του νοσήματος. Αντιθέτως, τα *μέτρα σχέσης* (measures of association) υπολογίζονται στις περιπτώσεις εκείνες κατά τις οποίες η έκθεση και η μη έκθεση σε έναν προσδιοριστή αφορούν σε δύο διαφορετικούς πληθυσμούς, ενώ εκτιμούν τη σχέση μεταξύ έκθεσης (σε έναν προσδιοριστή) και συχνότητας του νοσήματος. Ο υπολογισμός των μέτρων αποτελέσματος απαιτεί τα μελετώμενα άτομα να βρίσκονται ταυτόχρονα τόσο στην εκτεθειμένη όσο και στη μη εκτεθειμένη ομάδα. Αυτό βέβαια είναι αδύνατο και γι' αυτό σε μια μελέτη υπολογίζονται πάντοτε τα μέτρα σχέσης.

\* Η *αυθεντικότητα* (accuracy) αναφέρεται σε ποσοτικά (μετρικά) και μόνο μεγέθη (ή έννοιες) και σημαίνει την απουσία τόσο τυχαίων όσο και συστηματικών σφαλμάτων. Είναι λάθος να χαρακτηρίζεται το αποτέλεσμα μιας εργαστηριακής δοκιμασίας ως αυθεντικό (accurate), όταν εκφράζεται ως θετικό ή αρνητικό (ποιοτικό χαρακτηριστικό ή έννοια).

\*\* Με τον όρο μεροληψία έχει αποδοθεί ο όρος bias. Ο όρος δεν πρέπει να συγχέεται με την προκατάληψη, που αποδίδει τον όρο prejudice.

\*\*\* *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 της μελετώμενης έκβασης.<sup>8,13</sup> 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).<sup>13</sup> Η επιλογή της κατηγορίας αναφοράς του προσδιοριστή είναι μείζονος σημασίας στο σχεδιασμό της μελέτης. Είναι κρίσιμη σημασίας να γίνει αντιληπτό ότι η κατηγορία αναφοράς δεν είναι συμφύες χα-

ματος είναι διαφορετική σ' αυτούς που συμμετέχουν στη μελέτη σε σχέση μ' εκείνους που δεν συμμετέχουν σ' αυτή, όταν δηλαδή παρουσιάζεται επιλεκτική διαφοροποίηση του μελετώμενου πληθυσμού σε σχέση με τον πληθυσμό-πηγή.\* Η σχέση μεταξύ του προσδιοριστή και του νοσήματος στους μη συμμετέχοντες σε μια μελέτη είναι, συνήθως, άγνωστη, με αποτέλεσμα η παρουσία του συστηματικού σφάλματος επιλογής, συνήθως, να συμπεραίνεται λογικά παρά να παρατηρείται.

## 2.1. Συστηματικό σφάλμα εθελοντικής συμμετοχής ή αυτοεπιλογής

Το *συστηματικό σφάλμα της εθελοντικής συμμετοχής* (volunteer bias) σε μια μελέτη ή το *συστηματικό σφάλμα της αυτοεπιλογής*<sup>10,12,14</sup> (self-selection bias) συμβαίνει, όταν οι εθελοντές και οι μη εθελοντές διαφέρουν σε ορισμένα χαρακτηριστικά (όπως φύλο, ηλικία, κοινωνικοοικονομική κατάσταση κ.ά.) που σχετίζονται με τη συχνότητα του μελετώμενου νοσήματος. Συχνά, οι εθελοντές έχουν καλύτερη υγειονομική συνείδηση σε σχέση με τους μη εθελοντές, ενώ είναι δυνατόν ορισμένοι από αυτούς να συμμετέχουν σε μια μελέτη επειδή ανησυχούν ιδιαίτερα για τον κίνδυνο εμφάνισης του μελετώμενου νοσήματος.<sup>14</sup>

Μια νέα δοκιμασία διαλογής (screening test) για τη διάγνωση του καρκίνου του παχέος εντέρου εφαρμόζεται σε μια κοινότητα για μια καταρχήν εκτίμηση της

---

ρακτριτικό 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. 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 χρησιμοποιείται (α) ως μια ειδική (ενδεικτική) κατηγορία ενός (δυναμικού) 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 (δυναμικού) της μελετώμενης έκβασης, (γ) ως περιβαλλοντικός προσδιοριστής και (δ) εσφαλμένα ως προσδιοριστής (συγγενής, περιβαλλοντικός ή συμπεριφορικός) ή ως ενδεικτική κατηγορία ενός προσδιοριστή.

\* *Πληθυσμός-πηγή* (source population) είναι ο πληθυσμός από τον οποίο προέρχονται οι περιπτώσεις που περιλαμβάνονται στη μελέτη. Αντιπροσωπεύει τον υποθετικό μελετώμενο πληθυσμό, στον οποίο θα μπορούσε να είχε διεξαχθεί η μελέτη.<sup>1</sup> Ο πληθυσμός-πηγή είναι ο πληθυσμός από τον οποίο προέρχεται (είτε με επιλογή είτε με ορισμό) ο μελετώμενος πληθυσμός. Η πραγματοποίηση της μελέτης γίνεται στο μελετώμενο πληθυσμό, στο σύνολο, δηλαδή, των ατόμων που παρατηρούνται από τον ερευνητή.

αποτελεσματικότητας της δοκιμασίας. Η εκτίμηση γίνεται με τη σύγκριση της επίπτωσης-πυκνότητας\*\* του καρκίνου μεταξύ αυτών που εθελοντικά υποβλήθηκαν και εκείνων που δεν υποβλήθηκαν στη δοκιμασία. Ο ερευνητής υποθέτει ότι στη σύγκριση αυτή υπεισέρχεται συστηματικό σφάλμα επιλογής. Πιο συγκεκριμένα, αναμένεται, λογικά, να υπάρχει διαφορά στην επίπτωση-πυκνότητα του καρκίνου του παχέος εντέρου στις δύο συγκρινόμενες ομάδες ανεξάρτητα από το εάν η δοκιμασία διαλογής έχει κάποιο αποτέλεσμα. Τα άτομα που εθελοντικά υποβλήθηκαν στον έλεγχο αναμένεται να ενδιαφέρονται περισσότερο για την υγεία τους σε σχέση με τα άτομα που δεν υποβλήθηκαν στη δοκιμασία. Εκείνοι που ενδιαφέρονται περισσότερο για την υγεία τους είναι πιθανότερο να προσέχουν περισσότερο, π.χ. τη διατροφή τους, με αποτέλεσμα να έχουν μικρότερη πιθανότητα εμφάνισης καρκίνου του παχέος εντέρου. Έτσι, εκείνοι που υποβλήθηκαν εθελοντικά στη δοκιμασία αναμένεται, λογικά, να έχουν μικρότερη επίπτωση-πυκνότητα για λόγους που δεν σχετίζονται με τον έλεγχο. Η διαφορά αυτή υποδηλώνει την παρουσία συστηματικού σφάλματος στην εκτίμηση του αποτελέσματος της διαλογής.

Είναι δυνατόν, εντούτοις, να συμβαίνει και το ακριβώς αντίθετο. Πιο συγκεκριμένα, μπορεί ορισμένα άτομα να υποβλήθηκαν εθελοντικά στη δοκιμασία επειδή ανησυχούσαν για την εμφάνιση καρκίνου του παχέος εντέρου εξαιτίας, π.χ., επιβαρυσμένου οικογενειακού ιστορικού. Έτσι, οι εθελοντές αυτοί έχουν μεγαλύτερη πιθανότητα εμφάνισης καρκίνου του παχέος εντέρου σε σχέση με τους μη εθελοντές.

Επομένως, ορισμένοι εθελοντές έχουν αυξημένο και ορισμένοι άλλοι μειωμένο κίνδυνο σε σχέση με τους μη εθελοντές. Σε κάθε περίπτωση, πάντως, στη μελέτη υπάρχει συστηματικό σφάλμα επιλογής. Τα δύο αυτά σφάλματα αντιτίθενται το ένα στο άλλο, αλλά επειδή είναι πολύ δύσκολο να ποσοτικοποιηθούν, το καθαρό σφάλ-

---

\*\* Η *επίπτωση-πυκνότητα* (incidence-density) ποσοτικοποιεί την εμφάνιση ενός συμβάντος με τάξη αναφοράς μια πεπερασμένη ποσότητα πληθυσμο-χρόνου (population-time).<sup>8</sup> 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 μάλιστα σημειακών συμβάντων (point events) και όχι καταστάσεων. Η επίπτωση-πυκνότητα δεν είναι καθαρός αριθμός, αλλά αφενός έχει αριθμητική τιμή και αφετέρου μονάδα μέτρησης που είναι το αντίστροφο του χρόνου [π.χ. (έτος)<sup>-1</sup>].

μα είναι άγνωστο. Η ύπαρξη των συστηματικών σφαλμάτων επιλογής αποτελεί τον κυριότερο λόγο για τον οποίο η αποτελεσματικότητα των περισσότερων προγραμμάτων διαλογής εκτιμάται με τις τυχαιοποιημένες κλινικές δοκιμές (randomized clinical trials). Αν και μια τυχαιοποιημένη κλινική δοκιμή κοστίζει περισσότερο και είναι περισσότερο σύνθετη μεθοδολογικά, η τυχαιοποίηση εξασφαλίζει ότι οι δύο σειρές είναι συγκρίσιμες στο χρόνο  $T_0$ , αρκεί, βεβαίως, ο αριθμός των μελετώμενων ατόμων να είναι επαρκώς μεγάλος.

## 2.2. «Το φαινόμενο του υγιή εργάτη»

Το συστηματικό σφάλμα επιλογής στο παραπάνω παράδειγμα προέρχεται από αυτοεπιλογή, επειδή τα μελετώμενα άτομα επέλεξαν μόνα τους να υποβληθούν στον έλεγχο διαλογής. Συστηματικά σφάλματα επιλογής μπορούν να προκύψουν και από επιλογές που προέρχονται από τον ερευνητή. Για παράδειγμα, πολλές μελέτες που αφορούν στην υγεία εργαζομένων σε διάφορα επαγγέλματα συγκρίνουν τη θνησιμότητα των εργαζομένων σ' ένα συγκεκριμένο επάγγελμα με τη θνησιμότητα στο «γενικό πληθυσμό».\* Στη σύγκριση αυτή υπεισέρχεται συστηματικό σφάλμα επιλογής, επειδή ο «γενικός πληθυσμός» περιλαμβάνει και άτομα που δεν μπορούν να εργαστούν εξαιτίας κάποιου προβλήματος υγείας. Έτσι, η γενική θνησιμότητα (overall rate of mortality) των εργαζομένων είναι συνήθως μικρότερη από εκείνη του «γενικού πληθυσμού». Το συστηματικό αυτό σφάλμα, συχνά, αναφέρεται ως «φαινόμενο του υγιή εργάτη» (healthy worker effect).<sup>1,5,15-19</sup> Ένας τρόπος για να εξαλειφθεί το σφάλμα αυτό είναι να συγκριθούν οι εργαζόμενοι σ' ένα συγκεκριμένο επαγγελματικό χώρο με τους εργαζόμενους σ' έναν άλλο επαγγελματικό χώρο, ώστε οι δύο μελετώμενες ομάδες εργαζομένων να διαφέρουν ως προς τις επαγγελματικές εκθέ-

σεις. Εάν όλα τα άτομα που περιλαμβάνονται στη σύγκριση είναι εργαζόμενοι, τότε αποφεύγεται το σφάλμα που οφείλεται στο φαινόμενο του υγιή εργάτη.

Στον πίνακα 1 φαίνεται ένα υποθετικό παράδειγμα που αφορά στο φαινόμενο του υγιή εργάτη.<sup>1</sup> Η θνησιμότητα<sup>20</sup> των εργαζομένων που είναι εκτεθειμένοι σ' ένα συγκεκριμένο περιβαλλοντικό προσδιοριστή (ή έκθεση) ισούται με 0,05 (έτος)<sup>-1</sup>. Η θνησιμότητα στο «γενικό πληθυσμό», εξάλλου, ισούται με 0,07 (έτος)<sup>-1</sup>. Επομένως, η θνησιμότητα στο «γενικό πληθυσμό» είναι μεγαλύτερη από τη θνησιμότητα των εργαζομένων. Πιο συγκεκριμένα, η θνησιμότητα των εργαζομένων είναι το 71% της θνησιμότητας του «γενικού πληθυσμού». Ο «γενικός πληθυσμός», όμως, αποτελείται από δύο επιμέρους ομάδες. Η μια ομάδα, η αριθμητικά μεγαλύτερη, αποτελείται από άτομα που εργάζονται, ενώ η άλλη ομάδα, η μικρότερη, αποτελείται από άτομα με προβλήματα υγείας, που δεν εργάζονται. Στο «γενικό πληθυσμό», η θνησιμότητα στην ομάδα των μη εργαζομένων [0,25 (έτος)<sup>-1</sup>] είναι πολύ μεγαλύτερη από τη θνησιμότητα στην ομάδα των εργαζομένων [0,05 (έτος)<sup>-1</sup>]. Στο υποθετικό αυτό παράδειγμα, οι εργαζόμενοι στο «γενικό πληθυσμό» έχουν την ίδια θνησιμότητα με τους εργαζόμενους που είναι εκτεθειμένοι σ' ένα συγκεκριμένο προσδιοριστή (ή έκθεση), αλλά επειδή η θνησιμότητα των μη εργαζομένων στο «γενικό πληθυσμό» είναι πέντε φορές μεγαλύτερη από εκείνη των εκτεθειμένων εργαζομένων, η γενική θνησιμότητα στο «γενικό πληθυσμό» είναι μεγαλύτερη από τη θνησιμότητα των εκτεθειμένων εργαζομένων. Έτσι, όταν συγκρίνεται η θνησιμότητα μιας ομάδας εργαζομένων με τη θνησιμότητα του «γενικού πληθυσμού», οι εκτεθειμένοι εργαζόμενοι θα έχουν μικρότερη θνησιμότητα ως αποτέλεσμα του συστηματικού σφάλματος επιλογής που εισάγεται από το φαινόμενο του υγιή εργάτη.

**Πίνακας 1.** Το «φαινόμενο του υγιή εργάτη» αποτελεί συστηματικό σφάλμα επιλογής, το οποίο υποεκτιμά τη θνησιμότητα, που οφείλεται σε επαγγελματικές εκθέσεις, όπως φαίνεται από τις παρακάτω θνησιμότητες σε εκτεθειμένους σ' ένα συγκεκριμένο παράγοντα κινδύνου και στο γενικό πληθυσμό.<sup>1</sup>

	Γενικός πληθυσμός			Σύνολο
	Εκτεθειμένοι εργαζόμενοι	Εργαζόμενοι	Μη εργαζόμενοι	
Θάνατοι	50	4500	2500	7000
Πληθυσμο-χρόνος	1000	90000	10000	100000
Θνησιμότητα (θάνατοι/έτος)	0,05	0,05	0,25	0,07

\* Τονίζεται ότι το περιεχόμενο του όρου «γενικός πληθυσμός» (general population) είναι ασαφές και χρησιμοποιείται με διαφορετικό τρόπο από τους ερευνητές. Η έννοια καθαυτή δεν είναι επιστημονικά χρήσιμη.

### 3. ΣΥΣΤΗΜΑΤΙΚΑ ΣΦΑΛΜΑΤΑ ΠΛΗΡΟΦΟΡΙΑΣ

Τα *συστηματικά σφάλματα πληροφορίας* (systematic information biases)<sup>1,5,15,19</sup> σε μια μελέτη οφείλονται στην εσφαλμένη πληροφορία που συλλέγεται για τα μελετώμενα άτομα ή στην εσφαλμένη πληροφορία που προέρχεται από τα ίδια τα μελετώμενα άτομα. Η εσφαλμένη αυτή πληροφορία οδηγεί σε *δυσταξινόμηση* (misclassification), όταν ο μελετώμενος προσδιοριστής είναι ποιοτικός και το σφάλμα έχει ως αποτέλεσμα ένα άτομο να τοποθετείται σε λανθασμένη κατηγορία ή τάξη. Π.χ., ένας «βαρύς» καπνιστής που ταξινομείται ως «ελαφρύς» οδηγεί σε δυσταξινόμηση. Τυπικά, οι δύο σημαντικότερες μεταβλητές\* σε μια επιδημιολογική μελέτη είναι ο προσδιοριστής και το νόσημα. Η δυσταξινόμηση των ατόμων, είτε αφορά στον προσδιοριστή είτε στο νόσημα, μπορεί να είναι *διαφορική* (differential) ή *μη διαφορική* (non-differential). Οι δύο αυτοί όροι αναφέρονται στο μηχανισμό της δυσταξινόμησης. Αναφορικά με τη δυσταξινόμηση του προσδιοριστή, η δυσταξινόμηση είναι μη διαφορική εάν δεν σχετίζεται με τη συχνότητα εμφάνισης ή την παρουσία του νοσήματος. Εάν η δυσταξινόμηση του προσδιοριστή είναι διαφορετική για τους νοσούντες και τους μη νοσούντες, τότε καλείται διαφορική. Παρομοίως, η δυσταξινόμηση του νοσήματος είναι μη διαφορική εάν δεν σχετίζεται με τον προσδιοριστή, ενώ σε αντίθετη περίπτωση είναι διαφορική.

#### 3.1. Διαφορική δυσταξινόμηση

**3.1.1. Συστηματικό σφάλμα ανάκλησης.** Ένα συχνό συστηματικό σφάλμα πληροφορίας είναι το *σφάλμα ανάκλησης* (recall bias)<sup>1,5</sup> που συμβαίνει σε μελέτες όπου η πληροφορία για τον προσδιοριστή στηρίζεται στη συνέντευξη των ατόμων μετά από την εμφάνιση του νοσήματος. Για παράδειγμα, σε ορισμένες μελέτες με παιδιά που γεννιούνται με συγγενείς ανωμαλίες, οι απαιτούμενες πληροφορίες λαμβάνονται με συνέντευξη από τις μητέρες τους μετά από τη γέννηση.<sup>1</sup> Λογικά, οι μητέρες που αποκτούν παιδιά με συγγενή ανωμαλία ανακαλούν\*\* με μεγαλύτερη ακρίβεια διάφορους προσδιοριστές (όπως τη λήψη φαρμάκων, την ύπαρξη πυρετού κ.ά.) στους οποίους εκτέθηκαν κατά τη διάρκεια της κύησης. Θεωρείται ότι η εμφάνιση συγγενούς ανωμαλίας στο νεογνό δρα ως ισχυρό κίνητρο στη μητέρα να σκεφθεί συστη-

ματικά για πιθανές αιτίες που ευθύνονται για τη γέννηση του παιδιού με συγγενή ανωμαλία. Αντιθέτως, το κίνητρο αυτό δεν υφίσταται στις μητέρες που απέκτησαν παιδιά χωρίς συγγενή ανωμαλία, με αποτέλεσμα οι μητέρες αυτές να δυσκολεύονται να ανακαλέσουν προσδιοριστές, όπως π.χ. τη λήψη φαρμάκων, την ύπαρξη πυρετού κ.ά. Η διαφορά αυτή στην ανάκληση αποτελεί μια ιδιαίτερη περίπτωση του σφάλματος ανάκλησης, που είναι γνωστή ως *σφάλμα ανάκλησης της μητέρας* (maternal recall bias).<sup>1,21,22</sup> Η περίπτωση αυτή αποτελεί χαρακτηριστικό παράδειγμα διαφορικής δυσταξινόμησης και πρέπει να διαχωρίζεται από το γενικότερο πρόβλημα της ανάμνησης και της αναφοράς προσδιοριστών που επηρεάζει όλα τα άτομα σε κάποιο βαθμό και το οποίο τείνει να ανήκει στην κατηγορία της μη διαφορικής δυσταξινόμησης.

Ένας τρόπος αποφυγής του σφάλματος ανάκλησης είναι ο καλύτερος σχεδιασμός των ερωτήσεων, έτσι ώστε να επιτυγχάνεται ακριβέστερη ανάκληση. Αυξάνοντας την ακρίβεια της ανάκλησης μειώνεται το σφάλμα ανάκλησης. Ένας άλλος τρόπος είναι να επιλεγεί μια τελείως διαφορετική «βασική σειρά», έτσι ώστε να μην υπόκειται σε μη ολοκληρωμένη ανάκληση. Για παράδειγμα, μητέρες που απέκτησαν παιδιά με διαμαρτία της διάπλασης αποτελούν τη «σειρά των περιπτώσεων»\*\*\* (case-series), ενώ η «βασική σειρά» (base series) αποτελείται από μητέρες που απέκτησαν παιδιά με οποιαδήποτε συγγενή ανωμαλία εκτός από τη μελετώμενη. Οι μητέρες που αποτελούν τη «βασική σειρά» είναι πιθανότερο να ανακαλέσουν στη μνήμη τους πρόσφατους προσδιοριστές σε σχέση με τις μητέρες που απέκτησαν φυσιολογικά παιδιά. Τέλος, το σφάλμα ανάκλησης είναι δυνατόν να αποφευχθεί εάν οι συλλεγόμενες πληροφορίες δεν προέρχονται από προσωπικές συνεντεύξεις, αλλά από ιατρικά αρχεία που καταγράφονται πριν από τη γέννηση των παιδιών.

**3.1.2. Συστηματικό σφάλμα παρακολούθησης.** Το σφάλμα ανάκλησης αποτελεί παράδειγμα διαφορικής δυσταξινόμησης, επειδή η πληροφορία που αφορά στον προσδιοριστή ταξινομείται διαφορετικά στους νοσούντες και στους μη νοσούντες. Ανάλογο παράδειγμα διαφορικής δυσταξινόμησης είναι και το *συστηματικό σφάλμα παρακολούθησης* (follow-up bias).<sup>1,5</sup> Στην περίπτωση αυτή, στα μη εκτεθειμένα άτομα υποδιαγιγνώσκεται το νόση-

\* Το νόσημα και ο προσδιοριστής δεν είναι μεταβλητές (variates), αλλά μετατρέπονται από τον ερευνητή σε (στατιστικές) μεταβλητές.

\*\* Ανακαλώ στη μνήμη σημαίνει επαναφέρω κάτι από το παρελθόν στη μνήμη μου. Συνώνυμα του ανακαλώ είναι το αναθυμούμαι και το αναπολώ.

\*\*\* Η εγκυρότητα της στοιχειώδους σειράς, δηλαδή της «σειράς των περιπτώσεων», απαιτεί (προϋποθέτει) ότι η κατάσταση (state) του προσδιοριστή καθαυτή δεν σχετίζεται με τη διάγνωση των περιπτώσεων. Επίσης, η κατάσταση του προσδιοριστή δεν πρέπει να σχετίζεται με τη δειγματοληψία της βάσης μελέτης (base element).

μα σε μεγαλύτερο βαθμό από τα εκτεθειμένα άτομα. Για παράδειγμα, ένας ερευνητής χρησιμοποιεί έναν κλειστό πληθυσμό για να εκτιμήσει την επίδραση του καπνίσματος στη συχνότητα εμφάνισης του εμφυσήματος. Η συλλεγόμενη πληροφορία βασίζεται στις ιατρικές διαγνώσεις και δεν περιλαμβάνει το αποτέλεσμα εργαστηριακών ελέγχων για την επιβεβαίωση ή όχι της διάγνωσης. Είναι πιθανόν το εμφύσημα, η διάγνωση του οποίου συχνά διαφεύγει από τους γιατρούς, να διαγιγνώσκεται συχνότερα στους καπνιστές σε σχέση με τους μη καπνιστές. Και αυτό γιατί οι γιατροί, εξαιτίας του καπνίσματος, εξετάζουν λεπτομερέστερα τους καπνιστές για την ύπαρξη νοσημάτων του αναπνευστικού. Έτσι, η διάγνωση του εμφυσήματος διαφεύγει συχνότερα στους μη καπνιστές, οδηγώντας σε διαφορετική δυσταξινόμηση του νοσήματος. Ακόμη και αν το κάπνισμα δεν σχετίζεται με το εμφύσημα, οι καπνιστές φαίνεται να έχουν μεγαλύτερη επίπτωση-πυκνότητα του εμφυσήματος σε σχέση με τους μη καπνιστές, εξαιτίας του ότι το εμφύσημα έχει μεγαλύτερη πιθανότητα να μη διαγνωστεί σ' ένα μη καπνιστή. Το σφάλμα αυτό παρακολούθησης μπορεί να αποφευχθεί με την πραγματοποίηση ειδικού εργαστηριακού ελέγχου για τη διάγνωση του εμφυσήματος.

### 3.2. Μη διαφορική δυσταξινόμηση

Το σφάλμα ανάκλισης και το σφάλμα παρακολούθησης αποτελούν παραδείγματα διαφορικής δυσταξινόμησης, όταν είτε ο προσδιοριστής ταξινομείται διαφορετικά σύμφωνα με την κατάσταση ενός ατόμου όσον αφορά στο νόσημα είτε το νόσημα ταξινομείται διαφορετικά σύμφωνα με την κατάσταση ενός ατόμου όσον αφορά στον προσδιοριστή. Η διαφορική δυσταξινόμηση μπορεί είτε να υπερεκτιμήσει είτε να υποεκτιμήσει τα μέτρα σχέσης.<sup>1,5,10</sup>

Η μη διαφορική δυσταξινόμηση επηρεάζει άλλοτε σε μικρό και άλλοτε σε μεγάλο βαθμό κάθε επιδημιολογική μελέτη. Στη μη διαφορική δυσταξινόμηση, είτε ο προσδιοριστής είτε το νόσημα (ή και τα δύο) δυσταξινομούνται, αλλά η δυσταξινόμηση δεν εξαρτάται από την κατάσταση ενός ατόμου ως προς την άλλη μεταβλητή. Παραδείγματος χάριν, σε μια μελέτη που ερευνά τη σχέση μεταξύ της κατανάλωσης κρασιού και της συχνότητας εμφάνισης του πνευμονικού εμφυσήματος δεν είναι λογικό να υποτεθεί ότι τα άτομα που καταναλώνουν περισσότερο ή λιγότερο κρασί έχουν μεγαλύτερη ή μικρότερη πιθανότητα, αντίστοιχα, να διαγνωστούν ότι πάσχουν από εμφύσημα εφόσον πράγματι πάσχουν από αυτό. Έτσι, ορισμένα άτομα πάσχουν από εμφύσημα που δεν έχει διαγνωστεί. Το ποσοστό, όμως, των ατόμων με

μη διαγνωσμένο εμφύσημα αναμένεται να είναι το ίδιο γι' αυτούς που καταναλώνουν και για εκείνους που δεν καταναλώνουν κρασί. Η «υποδιάγνωση» αποτελεί μερική δυσταξινόμηση του εμφυσήματος, αλλά επειδή η τάση αυτή για «υποδιάγνωση» είναι η ίδια για εκτεθειμένους και μη, η δυσταξινόμηση της νόσου είναι μη διαφορική. Στο παράδειγμα αυτό υποτίθεται ότι η κατανάλωση κρασιού δεν σχετίζεται με το κάπνισμα.

Παρομοίως, εάν μια έκθεση δυσταξινομείται με τρόπο που δεν επηρεάζεται από την ύπαρξη ή μη του νοσήματος, τότε η δυσταξινόμηση του προσδιοριστή είναι μη διαφορική.

Η μη διαφορική δυσταξινόμηση οδηγεί σε περισσότερο προβλέψιμα σφάλματα απ' ό,τι η διαφορική δυσταξινόμηση. Η μη διαφορική δυσταξινόμηση ενός διχοτόμου προσδιοριστή οδηγεί σε εκτιμήσεις του μέτρου σχέσης που είναι πλησιέστερο στη μηδενική τιμή (null value) ή τιμή της απουσίας αποτελέσματος (no-effect value) παρά στο πραγματικό αποτέλεσμα. Εάν δεν υπάρχει αποτέλεσμα, τότε η μη διαφορική δυσταξινόμηση του προσδιοριστή δεν εισάγει συστηματικό σφάλμα στην εκτίμηση του αποτελέσματος.

Η απλούστερη περίπτωση μη διαφορικής δυσταξινόμησης είναι εκείνη στην οποία ο προσδιοριστής είναι διχοτόμος. Σε μια μελέτη που ερευνά τη σχέση μεταξύ δίαιτας πλούσιας σε λιπαρά και της συχνότητας εμφάνισης εμφράγματος του μυοκαρδίου, τα μελετώμενα άτομα ταξινομούνται ως εκτεθειμένα στη δίαιτα ή ως μη εκτεθειμένα σύμφωνα μ' ένα αυθαίρετα επιλεγμένο διαχωριστικό όριο. Η ταξινόμηση αυτή, λογικά, δεν είναι ακριβής επειδή είναι σχεδόν αδύνατον να αποφευχθούν τα σφάλματα στις μετρήσεις. Έτσι, είναι δυνατόν ορισμένα άτομα που λαμβάνουν δίαιτα πλούσια σε λιπαρά να ταξινομούνται εσφαλμένα στην κατηγορία των μη εκτεθειμένων και το αντίστροφο. Εάν οι δυσταξινόμησης αυτές δεν σχετίζονται με το αν το άτομο θα εμφανίσει έμφραγμα του μυοκαρδίου, τότε η δυσταξινόμηση είναι μη διαφορική.

Στον πίνακα 2 φαίνεται ένα υποθετικό παράδειγμα μη διαφορικής δυσταξινόμησης ενός διχοτόμου προσδιοριστή.<sup>1</sup> Στο αριστερό μέρος του πίνακα 2 φαίνονται τα δεδομένα της μελέτης εφόσον έχει πραγματοποιηθεί σωστή ταξινόμηση των ατόμων όσον αφορά στην πλούσια σε λιπαρά δίαιτα. Ο λόγος των επιπτώσεων-πυκνοτήτων (που εσφαλμένα φέρεται με την έκφραση «λόγος των οτζ»)\* ισούται με 5, υποδηλώνοντας ότι η συχνότητα

\* Η εσφαλμένη έκφραση «λόγος των οτζ» αντιστοιχεί στο λόγο των ονομαστικών επιπτώσεων-πυκνοτήτων (quasi incidence-density).

**Πίνακας 2.** Μη διαφορική δυσταξινόμηση σε μια υποθετική μελέτη «ασθενών-μαρτύρων».<sup>1</sup>

	Σωστή ταξινόμηση		Μη διαφορική δυσταξινόμηση			
			20% Όχι <input type="checkbox"/> Ναι		20% Όχι <input type="checkbox"/> Ναι 20% Ναι <input type="checkbox"/> Όχι	
	Δίαιτα πλούσια σε λιπαρά		Δίαιτα πλούσια σε λιπαρά		Δίαιτα πλούσια σε λιπαρά	
	Όχι	Ναι	Όχι	Ναι	Όχι	Ναι
«Ασθενείς»	450	250	360	340	410	290
«Μάρτυρες»	900	100	720	280	740	260
Λόγος επιπτώσεων-πυκνοτήτων	5		2,4		2	

τα εμφάνισης του εμφράγματος του μυοκαρδίου είναι πέντε φορές μεγαλύτερη σε εκείνους που λαμβάνουν δίαιτα πλούσια σε λιπαρά. Η μεσαία στήλη δείχνει το αποτέλεσμα όταν το 20% εκείνων που δεν λαμβάνουν δίαιτα πλούσια σε λιπαρά ταξινομούνται εσφαλμένα στην κατηγορία των ατόμων που λαμβάνουν δίαιτα πλούσια σε λιπαρά, στην κατηγορία δηλαδή των εκτεθειμένων. Το επίπεδο αυτό (20%) της δυσταξινόμησης είναι υψηλότερο από εκείνο που συνήθως αναμένεται, ακόμη και για μια έκθεση που δεν είναι ιδιαίτερα δύσκολο να μετρηθεί με ακρίβεια, όπως η κατανάλωση λιπαρών, αλλά περιλαμβάνει ένα μικρό μόνο ποσοστό των μελετώμενων ατόμων. Μετακινώντας το 20% εκείνων από τη στήλη του «όχι» στη στήλη του «ναι» προκύπτει η μεσαία στήλη του πίνακα 2, σύμφωνα με την οποία ο λόγος των επιπτώσεων-πυκνοτήτων ισούται, πλέον, με 2,4, λιγότερο από το μισό που προέκυψε όταν τα άτομα ταξινομήθηκαν σωστά. Στην αριστερή στήλη των σωστά ταξινομημένων ατόμων, ο πλεονάζων λόγος των επιπτώσεων-πυκνοτήτων ισούται με 4 (=5-1, καθώς η μηδενική τιμή που αντιστοιχεί στην απουσία αποτελέσματος της δίαιτας πλούσια σε λιπαρά ισούται με τη μονάδα). Στη μεσαία στήλη, που προέκυψε έπειτα από τη μη διαφορική δυσταξινόμηση των μελετώμενων ατόμων, ο πλεονάζων λόγος των επιπτώσεων-πυκνοτήτων ισούται με 1,4 (=2,4-1). Έτσι, η μη διαφορική δυσταξινόμηση οδηγεί σε εκτιμήσεις του αποτελέσματος που πλησιάζουν στη μηδενική τιμή. Στη συγκεκριμένη περίπτωση, «μηδενίστηκαν» περίπου τα 2/3 του αποτελέσματος που προέκυψε από τα σωστά ταξινομημένα άτομα της αριστερής στήλης. Η δυσταξινόμηση είναι μη διαφορική, καθώς μεταφέρθηκε το 20% τόσο των «ασθενών» όσο και των «μαρτύρων». Μη διαφορική δυσταξινόμηση του προσδιοριστή υποδηλώνει ότι τα ποσοστά των δυσταξινομημένων ατόμων μεταξύ των «ασθενών» και των «μαρτύρων» είναι ίσα.

Εάν τα ποσοστά αυτά διαφέρουν μεταξύ τους, τότε η δυσταξινόμηση είναι διαφορική.

Στη δεξιά στήλη του πίνακα 2 προστίθεται επιπλέον μη διαφορική δυσταξινόμηση, καθώς, εκτός από το 20% τόσο των «ασθενών» όσο και των «μαρτύρων» που μεταφέρεται από τη στήλη του «όχι» στη στήλη του «ναι», μεταφέρεται, επιπλέον, και ένα 20% τόσο των «ασθενών» όσο και των «μαρτύρων» από τη στήλη του «ναι» στη στήλη του «όχι». Έτσι, το 20% εκείνων που ελάμβαναν δίαιτα πλούσια σε λιπαρά ταξινομήθηκε εσφαλμένα στην κατηγορία εκείνων που δεν ελάμβαναν δίαιτα πλούσια σε λιπαρά, στην κατηγορία δηλαδή των μη εκτεθειμένων. Και πάλι η δυσταξινόμηση είναι μη διαφορική, καθώς μεταφέρθηκε το ίδιο ακριβώς ποσοστό (20%) τόσο για τους «ασθενείς» όσο και για τους «μάρτυρες». Στην περίπτωση αυτή, ουσιαστικά, υπάρχουν δύο μη διαφορικές δυσταξινομήσεις που πραγματοποιούνται προς αντίθετες κατευθύνσεις. Όπως φαίνεται στη δεξιά στήλη, μετά και την επιπλέον μη διαφορική δυσταξινόμηση, ο λόγος των επιπτώσεων-πυκνοτήτων μειώθηκε ακόμη περισσότερο, φθάνοντας στο 2. Έτσι, η επιπλέον δυσταξινόμηση οδήγησε σ' ένα αποτέλεσμα που πλησιάζει ακόμη περισσότερο τη μηδενική τιμή που ισούται με τη μονάδα. Στην περίπτωση αυτή της επιπλέον δυσταξινόμησης «μηδενίστηκαν» τα 3/4 του αποτελέσματος που προέκυψε όταν τα άτομα είχαν ταξινομηθεί σωστά (αριστερή στήλη).

Η μη διαφορική δυσταξινόμηση ενός διχοτόμου προσδιοριστή εισάγει πάντοτε σφάλμα στην εκτίμηση ενός μέτρου σχέσης, εφόσον βεβαίως υπάρχει αποτέλεσμα, μετακινώντας την τιμή του μέτρου σχέσης προς τη μηδενική τιμή. Εάν η έκθεση δεν είναι δυαδική, τότε εισάγεται και πάλι σφάλμα, αλλά δεν είναι γνωστή εκ των προτέρων η κατεύθυνση της μετακίνησης της τιμής του

αποτελέσματος, καθώς εξαρτάται από τις κατηγορίες στις οποίες δυσταξινομούνται τα άτομα.<sup>4,5</sup>

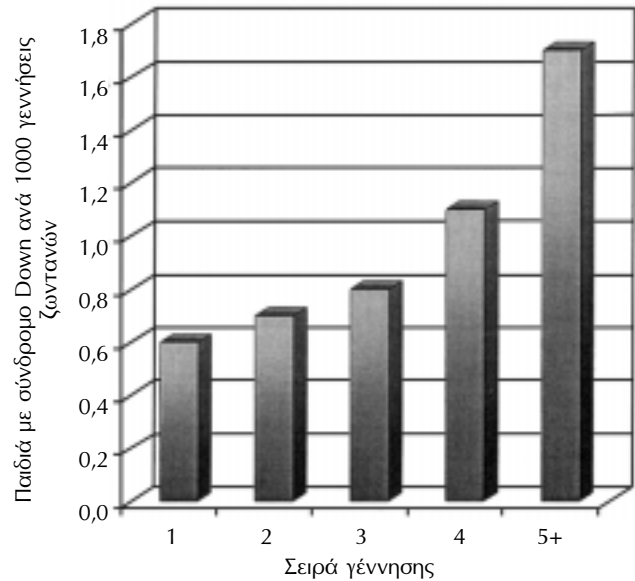
#### 4. ΣΥΓΧΥΣΗ

Η ύπαρξη σύγχυσης<sup>1,2,5,10,19,23</sup> είναι εξαιρετικά σημαντική σε μια επιδημιολογική μελέτη. Ένας απλός ορισμός της σύγχυσης θα μπορούσε να είναι η ανάμιξη ή, αλλιώς, το ανακάτεμα των αποτελεσμάτων. Ο ορισμός αυτός υποδηλώνει ότι το αποτέλεσμα του μελετώμενου προσδιοριστή αναμειγνύεται με το αποτέλεσμα ενός εξωγενούς προσδιοριστή,\* οδηγώντας έτσι σε συστηματικό σφάλμα.

Η σύγχυση (confounding)<sup>1,4,5,23-26</sup> 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 σχέσης 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 κατηγορία των εκτεθειμένων, ενώ με IDo συμβολίζεται η επίπτωση-πυκνότητα στους μη εκτεθειμένους. Η σύγχυση δεν αποτελεί χαρακτηριστικό του κάθε εξωγενούς προσδιοριστή. Έτσι, ένας εξωγενής προσδιοριστής που είναι συγχυτής μιας σχέσης σε μια μελέτη, μπορεί να μην είναι συγχυτής της ίδιας σχέσης σε μια άλλη μελέτη.<sup>2,4,5,23</sup>

Κλασικό παράδειγμα για την κατανόηση του φαινομένου της σύγχυσης είναι αυτό της σχέσης μεταξύ της σειράς γέννησης και της συχνότητας εμφάνισης του συνδρόμου Down. Στην εικόνα 1 φαίνονται τα δεδομένα μιας μελέτης<sup>27</sup> που αφορά στη σχέση της σειράς γέννησης και της συχνότητας εμφάνισης του συνδρόμου Down. Από τα δεδομένα αυτά προκύπτει μια αξιοσημείωτη αύξηση του επιπολασμού\*\* του συνδρόμου Down καθώς αυξάνεται η σειρά γέννησης. Η σχέση αυτή ενδεχομένως να οφείλεται στο αποτέλεσμα της σειράς γέννησης στη συχνότητα εμφάνισης του συνδρόμου Down. Το αποτέλεσμα της σειράς γέννησης, ωστόσο, είναι μια ανάμιξη του αποτελέσματος της σειράς γέννησης αυτής καθαυτής και του αποτελέσματος ενός εξωγενούς προσδιοριστή που συσχετίζεται σε υψηλό βαθμό

\* Οι εξωγενείς προσδιοριστές (extraneous determinants) καλούνται δυνητικοί συγχυτές (potential confounders),<sup>13</sup> καθώς σχετίζονται με τη συχνότητα εμφάνισης της μελετώμενης έκβασης. Όταν οι εξωγενείς αυτοί προσδιοριστές ανισοκατανέμονται κίχλας στις κατηγορίες του μελετώμενου προσδιοριστή, τότε καλούνται πραγματικοί συγχυτές (actual confounders).

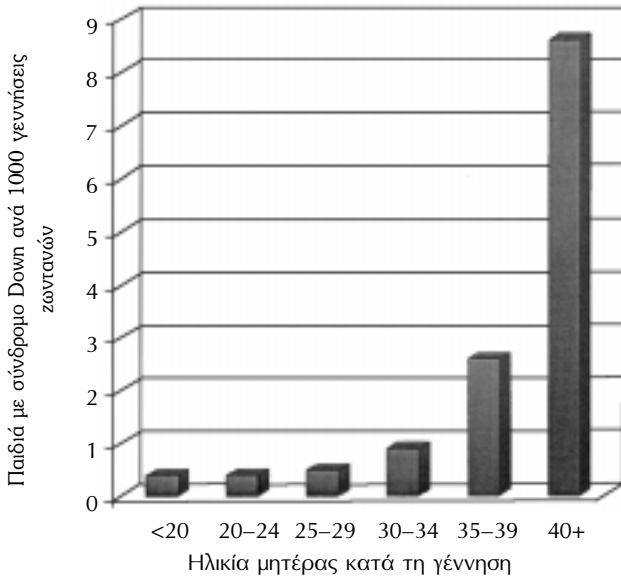


Εικόνα 1. Επιπολασμός του συνδρόμου Down σύμφωνα με τη σειρά γέννησης.<sup>27</sup>

με τη σειρά γέννησης. Ο εξωγενής προσδιοριστής που συχθεί τη σχέση μεταξύ της σειράς γέννησης και της συχνότητας εμφάνισης του συνδρόμου Down είναι η ηλικία της μητέρας κατά τη γέννηση. Στην εικόνα 2 φαίνεται η σχέση μεταξύ της ηλικίας της μητέρας και της συχνότητας εμφάνισης του συνδρόμου Down, με βάση και πάλι τα δεδομένα της ίδιας μελέτης. Η εικόνα 2 φανερώνει μια πολύ ισχυρή σχέση μεταξύ της ηλικίας της μητέρας και της συχνότητας εμφάνισης του συνδρόμου Down, καθώς όσο αυξάνεται η ηλικία της μητέρας τόσο αυξάνεται και η συχνότητα εμφάνισης του συνδρόμου Down. Η σχέση αυτή, μάλιστα, είναι πολύ πιο ισχυρή από τη σχέση μεταξύ της σειράς γέννησης και της συχνότητας εμφάνισης του συνδρόμου Down. Πιο συγκεκριμένα, στην εικόνα 1, ο επιπολασμός από 0,6 περιπτώσεις/1000 γεννήσεις ζωντανών παιδιών στην πρώτη γέννηση αυξάνεται σε 1,7 περιπτώσεις/1000 γεννήσεις ζωντανών παιδιών στην πέμπτη (ή και μεγαλύτερη γέννηση), ενώ στην εικόνα 2 ο επιπολασμός από 0,2 περιπτώσεις/1000 γεννήσεις ζωντανών παι-

\*\* Ο επιπολασμός<sup>20</sup> αποτελεί μέτρο συχνότητας των νοσημάτων. Ο επιπολασμός ενός συγκεκριμένου χαρακτηριστικού αντιστοιχεί στο ποσοστό του πληθυσμού που έχει το συγκεκριμένο χαρακτηριστικό σε ένα ορισμένο σημείο στο χρόνο. Είναι ποσοστό, οπότε λαμβάνει τιμές από 0-1, ενώ δεν έχει χρονική διάσταση, καθώς αναφέρεται σε μια συγκεκριμένη χρονική στιγμή. Για παράδειγμα, εάν ο πληθυσμός των γυναικών που διέμεναν στην Αθήνα, την 1η Ιανουαρίου 2000, ήταν 1.500.000 και 5.000 γυναίκες αυτού του πληθυσμού έπασχαν, τη συγκεκριμένη χρονική στιγμή, από καρκίνο του μαστού, τότε ο επιπολασμός του καρκίνου του μαστού, την 1η Ιανουαρίου 2000, στην Αθήνα ήταν 0,0033 (5.000/1.500.000).





**Εικόνα 2.** Επιπολασμός του συνδρόμου Down σύμφωνα με την ηλικία της μητέρας κατά την κύηση.<sup>27</sup>

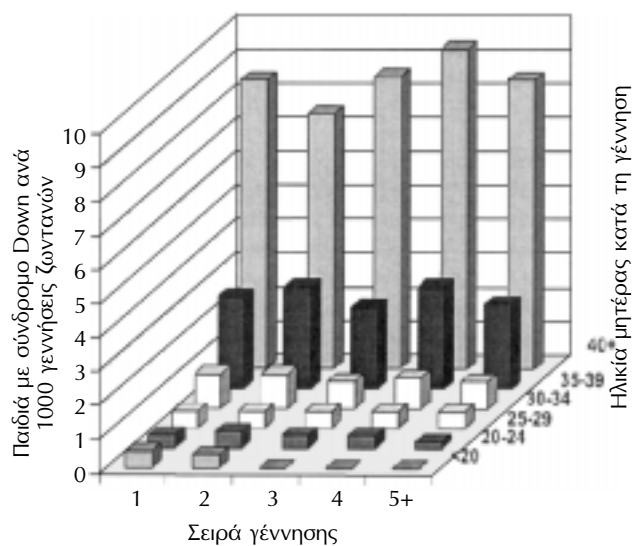
διών όταν η ηλικία της μητέρας είναι <20 ετών (στη μικρότερη δηλαδή ηλικιακή ομάδα) αυξάνεται σε 8,5 περιπτώσεις/1000 γεννήσεις ζωντανών παιδιών όταν η ηλικία της μητέρας είναι >40 ετών (στη μεγαλύτερη δηλαδή ηλικιακή ομάδα).

Η σειρά γέννησης συσχετίζεται σε υψηλό βαθμό με την ηλικία της μητέρας, οπότε αναμένεται οι μητέρες που αποκτούν πέμπτο παιδί να είναι σημαντικά μεγαλύτερης ηλικίας από τις μητέρες που αποκτούν πρώτο παιδί. Έτσι, η σύγκριση παιδιών που ανήκουν σε μεγαλύτερη σειρά γέννησης με παιδιά που ανήκουν σε μικρότερη σειρά γέννησης είναι, ως ένα βαθμό, μια σύγκριση μεταξύ παιδιών που γεννήθηκαν από μεγαλύτερες σε ηλικία μητέρες με παιδιά που γεννήθηκαν από μικρότερης ηλικίας μητέρες. Η σύγκριση, επομένως, της σειράς γέννησης στην εικόνα 1 αναμιγνύει το αποτέλεσμα της ηλικίας της μητέρας με το αποτέλεσμα της σειράς γέννησης. Ο βαθμός της ανάμιξης εξαρτάται από το βαθμό στον οποίο η ηλικία της μητέρας συσχετίζεται με τη σειρά γέννησης. Η ανάμιξη αυτή των αποτελεσμάτων καλείται σύγχυση. Στο συγκεκριμένο παράδειγμα, το αποτέλεσμα της σειράς γέννησης στη συχνότητα εμφάνισης του συνδρόμου Down, όπως φαίνεται στην εικόνα 1, συγχέεται από το αποτέλεσμα της ηλικίας της μητέρας, όπως φαίνεται στην εικόνα 2.

Στο παραπάνω παράδειγμα, βεβαίως, θα μπορούσε να διατυπωθεί και το αντίστροφο ερώτημα, αν δηλαδή η σχέση μεταξύ της ηλικίας της μητέρας και της συχνότητας εμφάνισης του συνδρόμου Down συγχέεται από τη

σειρά γέννησης. Ενδεχομένως αυτό να ισχύει, κάτι που εξαρτάται όμως από το αν η σειρά γέννησης ασκεί κάποια επίδραση συνολικά ή όχι. Η επίδραση της ηλικίας της μητέρας, όπως φαίνεται στην εικόνα 2, είναι πολύ πιο ισχυρή από την επίδραση της σειράς γέννησης, όπως φαίνεται στην εικόνα 1, με αποτέλεσμα η σειρά γέννησης να μην μπορεί να εξηγήσει πλήρως την επίδραση της ηλικίας της μητέρας, ενώ υπάρχει πιθανότητα η ηλικία της μητέρας να εξηγήσει πλήρως την επίδραση της σειράς γέννησης. Ένας απλός τρόπος για να κατανοηθεί ο βαθμός στον οποίο η επίδραση μιας μεταβλητής εξηγεί το αποτέλεσμα μιας άλλης είναι να εξεταστούν και τα δύο αποτελέσματα ταυτόχρονα. Στην εικόνα 3 φαίνεται ο επιπολασμός του συνδρόμου Down σύμφωνα με τη σειρά γέννησης και την ηλικία της μητέρας ταυτόχρονα.

Στην εικόνα 3 φαίνεται καθαρά ότι σε κάθε κατηγορία της σειράς γέννησης, κοιτάζοντας από εμπρός προς τα πίσω, υπάρχει μια σημαντική αύξηση του επιπολασμού του συνδρόμου Down καθώς αυξάνεται η ηλικία της μητέρας. Αντιθέτως, σε κάθε κατηγορία της ηλικίας της μητέρας, κοιτάζοντας από αριστερά προς τα δεξιά, δεν υπάρχει ευδιάκριτη τάση στον επιπολασμό του συνδρόμου Down σε σχέση με τη σειρά γέννησης. Έτσι, η σχέση μεταξύ της σειράς γέννησης και του επιπολασμού, όπως φαίνεται στην εικόνα 1, οφείλεται αποκλειστικά στη συγχυτική επίδραση της ηλικίας της μητέρας. Δεν υπάρχει, ωστόσο, σύγχυση προς την αντίθετη κατεύθυνση. Και αυτό, γιατί η σειρά γέννησης δεν συγχέει τη σχέση μεταξύ της ηλικίας της μητέρας και της



**Εικόνα 3.** Επιπολασμός του συνδρόμου Down σύμφωνα με τη σειρά γέννησης και την ηλικία της μητέρας.<sup>27</sup>

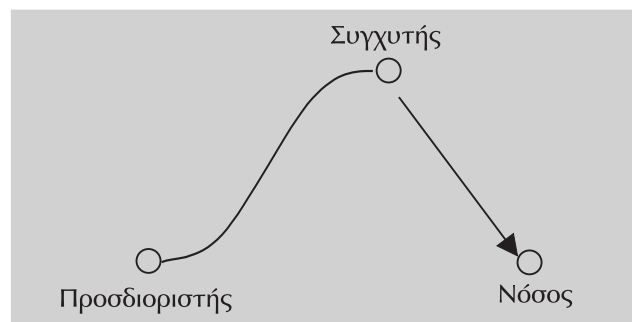
συχνότητας εμφάνισης του συνδρόμου Down, καθώς η σειρά γέννησης δεν έχει αποτέλεσμα. Το αποτέλεσμα της σειράς γέννησης στην εικόνα 1 καλείται αδρό αποτέλεσμα (crude effect). Στην εικόνα 3 φαίνεται ότι στις διάφορες κατηγορίες της ηλικίας της μητέρας δεν υφίσταται επίδραση της σειράς γέννησης.

Είναι, ωστόσο, και η επίδραση της ηλικίας της μητέρας, όπως φαίνεται στην εικόνα 2, αποτέλεσμα της σύγχυσης; Όπως προαναφέρθηκε, η επίδραση της ηλικίας της μητέρας δεν συγχέεται από τη σειρά γέννησης, αλλά συγχέεται από άλλους παράγοντες. Η ύπαρξη άλλων συγχυτών είναι βέβαιη, καθώς η ηλικία δεν μπορεί, αυτή καθαυτή, να προκαλέσει βιολογικές αλλαγές που να οδηγούν στην εμφάνιση του συνδρόμου Down. Ενδεχομένως, με την αύξηση της ηλικίας να επέρχονται διαταραχές βιολογικών παραμέτρων που συμβάλλουν σημαντικά στην αύξηση της συχνότητας εμφάνισης του συνδρόμου Down σε παιδιά που γεννιούνται από μητέρες μεγαλύτερης ηλικίας. Έτσι, η ηλικία της μητέρας είναι ένα «χαρακτηριστικό» (marker) που εκφράζει μη γνωστές βιολογικές διαταραχές, που, ουσιαστικά, ευθύνονται για την αυξημένη συχνότητα εμφάνισης του συνδρόμου Down στις μητέρες μεγαλύτερης ηλικίας. Όταν τα βιολογικά αυτά γεγονότα καθοριστούν με ακρίβεια, τότε θα φανεί ότι η ηλικία της μητέρας δεν έχει κάποια επίδραση στην εμφάνιση του συνδρόμου Down, αλλά είναι αποτέλεσμα της σύγχυσης που προκαλούν διάφορα, άγνωστα μέχρι σήμερα, βιολογικά γεγονότα.

Η διαδικασία αναγνώρισης και εξουδετέρωσης των συγχυτών παρομοιάζεται με προσπάθεια εξόδου από ένα λαβύρινθο. Στο παραπάνω παράδειγμα, η σχέση της σειράς γέννησης και της συχνότητας εμφάνισης του συνδρόμου Down εξηγείται πλήρως από τη συγχυτική επίδραση της ηλικίας της μητέρας. Επίσης, η επίδραση της ηλικίας της μητέρας μπορεί να εξηγηθεί από άλλους παράγοντες, όπως οι βιολογικοί, που δεν έχουν αναγνωριστεί ακόμη. Με την προοδευτική ανακάλυψη και εξουδετέρωση ολοένα και περισσότερων συγχυτών αποκαλύπτεται ο βαθύτερος υποκείμενος αιτιολογικός μηχανισμός.

*Χαρακτηριστικά ενός συγχυτή.* Συμπερασματικά, όπως προαναφέρθηκε, η σύγχυση μπορεί να θεωρηθεί ως «ανάμιξη των αποτελεσμάτων». Ένας συγχυτής, επομένως, πρέπει να έχει κάποιο αποτέλεσμα και επίσης θα πρέπει να ανισοκατανέμεται στις δύο, ή περισσότερες, κατηγορίες του μελετώμενου προσδιοριστή. Αναλυτικότερα, ένας εξωγενής προσδιοριστής, για να θεωρηθεί πραγματικός συγχυτής, πρέπει να έχει τα ακόλουθα χαρακτηριστικά (εικ. 4).<sup>1,2,5,10,19</sup>

- *Να σχετίζεται με τη συχνότητα εμφάνισης της μελετώμενης έκβασης<sup>16</sup> (που, συνήθως, είναι η εμφάνιση ενός νοσήματος) ή, με άλλη διατύπωση, να ασκεί κάποια επίδραση στην εμφάνιση της μελετώμενης έκβασης.* Θεωρητικά, ένας συγχυτής θα πρέπει να αποτελεί αιτία ενός νοσήματος, αλλά στην πράξη μπορεί, απλά, να είναι «χαρακτηριστικό» (marker) μιας άγνωστης αιτίας. Αυτό συμβαίνει στην περίπτωση της ηλικίας της μητέρας, όπως προαναφέρθηκε, η οποία αυτή καθαυτή δεν αποτελεί την αιτία της εμφάνισης του συνδρόμου Down, αλλά λειτουργεί ως «ενδεικτικό σήμα» (marker) ή δείκτης άγνωστων βιολογικών γεγονότων που συσσωρεύονται με το πέρασμα του χρόνου. Είτε αποτελεί αιτία είτε «δείκτη» μιας αιτίας, ο συγχυτής θα πρέπει να προβλέπει τη συχνότητα εμφάνισης ενός νοσήματος.
- *Να σχετίζεται με το μελετώμενο προσδιοριστή ή, με άλλη διατύπωση, να κατανέμεται ανισοτόμα στις δύο, ή περισσότερες, κατηγορίες του μελετώμενου προσδιοριστή.* Εάν υποθεθεί ότι η ηλικία, όπως συνήθως συμβαίνει, είναι ένας προσδιοριστής ενός συγκεκριμένου νοσήματος (π.χ. της στεφανιαίας νόσου) ή, καλύτερα, της συχνότητας εμφάνισης ενός νοσήματος, τότε η ηλικία θα αποτελέσει πραγματικό συγχυτή μόνο εάν η κατανομή της ηλικίας των μελετώμενων ατόμων στις διάφορες κατηγορίες (άνδρες, γυναίκες) του μελετώμενου προσδιοριστή (π.χ. φύλο) είναι διαφορετική. Εάν σε κάθε κατηγορία του μελετώμενου προσδιοριστή περιλαμβάνονται άτομα, των οποίων η κατανομή της ηλικίας είναι ίδια με την κατανομή στις άλλες κατηγορίες του προσδιοριστή, τότε η σύγκριση των μέτρων συχνότητας στις διάφορες κατηγορίες του προσδιοριστή δεν επηρεάζεται



Εικόνα 4. Σχηματική απεικόνιση των χαρακτηριστικών ενός συγχυτή.<sup>15</sup>

\* Στην αιτιογνωστική έρευνα, έκβαση (outcome) αποτελεί η εμφάνιση ή μη του νοσήματος. Η έκφραση «σχετίζεται με την έκβαση» σημαίνει σχέση με τη συχνότητα του νοσήματος και όχι με το νόσημα καθαυτό.

από το αποτέλεσμα της ηλικίας. Εάν, όμως, η ηλικία (ή, ορθότερα, η κατανομή της ηλικίας) ανισοκατανέμεται στις κατηγορίες του μελετώμενου προσδιοριστή, τότε η σύγκριση μιας κατηγορίας του προσδιοριστή με μια άλλη περιλαμβάνει άτομα, όπου οι κατανομές των ηλικιών τους διαφέρουν. Στην περίπτωση αυτή, το αποτέλεσμα του μελετώμενου προσδιοριστή (φύλου) συγχέεται με το αποτέλεσμα του πραγματικού συγχυτή (ηλικίας) και το μέγεθος της σύγχυσης εξαρτάται από τη δύναμη της σχέσης μεταξύ της ηλικίας και του νοσήματος, καθώς επίσης και από το βαθμό, στον οποίο η ηλικία ανισοκατανέμεται στις κατηγορίες του προσδιοριστή.

- *Να μην αποτελεί αποτέλεσμα του μελετώμενου προσδιοριστή, δηλαδή να μην αποτελεί ενδιάμεσο στάδιο στον αιτιακό μηχανισμό που συνδέει προσδιοριστή και έκβαση.* Για παράδειγμα, η πλούσια σε κορεσμένα λίπη δίαιτα προκαλεί αύξηση των επιπέδων της λιποπρωτεΐνης χαμηλής πυκνότητας (LDL) στο αίμα και η υψηλή LDL αυξάνει την πιθανότητα εμφάνισης της αθηρωματικής νόσου. Έτσι, τα υψηλά επίπεδα LDL σχετίζονται τόσο με τη συχνότητα της αθηρωματικής νόσου όσο και με τον προσδιοριστή (την πλούσια σε κορεσμένα λίπη δίαιτα). Εντούτοις, η αυξημένη LDL δεν είναι συγχυτής της σχέσης δίαιτας και αθηρωματικής νόσου, γιατί αποτελεί αποτέλεσμα του προσδιοριστή, καθώς η πλούσια σε κορεσμένα λίπη δίαιτα προκαλεί αύξηση της LDL.

Το 1970, διεξήχθη μια κλινική δοκιμή με αντικείμενο την εκτίμηση της αποτελεσματικότητας τριών θεραπευτικών αγωγών για την πρόληψη των θανατηφόρων επιπλοκών του σακχαρώδους διαβήτη.<sup>28</sup> Στον πίνακα 3 φαίνονται τα αδρά δεδομένα για τη σύγκριση μιας από τις θεραπευτικές αγωγές (λήψη τολβουταμίδης) σε σύγκριση με τη λήψη ανενεργούς ή αδρανούς ουσίας (εικονικό φάρμακο, placebo\*). Για τη σύγκριση αυτή υπολογίστηκε η επίπτωση-ποσοστό του θανάτου (θνητότητα) σε χρονικό διάστημα 7 ετών. Η θνητότητα<sup>20</sup> ήταν μεγαλύτερη στους ασθενείς που ελάμβαναν τολβουτα-

μίδα (0,147) σε σχέση με εκείνους που ελάμβαναν το εικονικό φάρμακο (0,102). Το εύρημα αυτό ήταν αντίθετο με την κοινή λογική και απαιτούσε περαιτέρω διερεύνηση. Η διαφορά θνητοτήτων στις δύο ομάδες ήταν  $0,147 - 0,102 = 0,045$ , που σήμαινε ότι οι ασθενείς που ελάμβαναν τολβουταμίδα αντιμετώπιζαν πρόσθετο κίνδυνο της τάξης του 4,5% να πεθάνουν στο διάστημα των 7 ετών σε σχέση με τους ασθενείς που ελάμβαναν το εικονικό φάρμακο.

Αν και η μελέτη ανήκε στις τυχαίοποιημένες κλινικές δοκιμές, η τυχαίοποίηση δεν εξασφάλισε την ίση κατανομή των εξωγενών προσδιοριστών στις δύο συγκρινόμενες ομάδες του μελετώμενου προσδιοριστή. Στη συγκεκριμένη μελέτη, η τυχαίοποίηση δεν εξουδετέρωσε τη συγχυτική επίδραση της ηλικίας, καθώς οι ασθενείς που ελάμβαναν τολβουταμίδα ήταν μεγαλύτερης ηλικίας από τους ασθενείς που ελάμβαναν το εικονικό φάρμακο. Επιπλέον, η ηλικία, ως γνωστόν, εμφανίζει ισχυρή συσχέτιση με τον κίνδυνο θανάτου, με αποτέλεσμα στη μελέτη αυτή η ηλικία να αποτελεί ισχυρό συγχυτή της σχέσης που μελετάται. Στο παράδειγμα με το σύνδρομο Down που προαναφέρθηκε, η εξουδετέρωση της σύγχυσης πραγματοποιήθηκε με τον έλεγχο του αποτελέσματος της σειράς γέννησης στις διάφορες κατηγορίες της ηλικίας της μητέρας. Η διαδικασία αυτή εξουδετέρωσης των συγχυτών μιας μελέτης καλείται *διαστρωμάτωση* (stratification). Έτσι, και στο παράδειγμα με τους διαβητικούς ασθενείς, για την εξουδετέρωση της συγχυτικής επίδρασης της ηλικίας πραγματοποιήθηκε διαστρωμάτωση των δεδομένων με βάση την ηλικία των συμμετεχόντων (πίν. 4).

Στον πίνακα 4 φαίνεται ότι 98 από τους 204 (δηλαδή το 48%) ασθενείς που ελάμβαναν τολβουταμίδα ήταν >55 ετών, ενώ 85 από τους 205 (δηλαδή το 41%) ασθενείς που ελάμβαναν εικονικό φάρμακο ήταν >55 ετών. Η διαφορά αυτή στην κατανομή της ηλικίας μπορεί να μη φαίνεται εντυπωσιακή, αλλά η θνητότητα εκείνων που ήταν >55 ετών ήταν σημαντικά μεγαλύτερη σε σχέση με τη θνητότητα εκείνων που ήταν <55 ετών. Εφόσον η ηλικία συσχετίζεται σε υψηλό βαθμό με τον κίνδυνο θανάτου, η διαφορά στις θνητότητες μεταξύ των δύο συγκρινόμενων κατηγοριών του μελετώμενου προσδιοριστή ενδεχομένως οφείλεται στη συγχυτική επίδραση της ηλικίας. Για να λυθεί το πρόβλημα αυτό πρέπει

**Πίνακας 3.** Θνητότητα μεταξύ διαβητικών ασθενών που ελάμβαναν τολβουταμίδα ή ανενεργό ουσία (placebo).<sup>28</sup>

	Τολβουταμίδα	Placebo
Θάνατοι	30	21
Ζωντανοί	174	184
Σύνολο	204	205
Επίπτωση-ποσοστό θανάτου (θνητότητα)	0,147	0,102

\* Πλασέμπο (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.<sup>29</sup>

**Πίνακας 4.** Θνητότητα μεταξύ διαβητικών ασθενών που ελάμβαναν τολβουταμίδη ή εικονικό φάρμακο (placebo) έπειτα από τη διαστρωμάτωση, με βάση την ηλικία.<sup>28</sup>

	Ηλικία <55 ετών		Ηλικία ≥55 ετών	
	Τολβουταμίδη	Placebo	Τολβουταμίδη	Placebo
Θάνατοι	8	5	22	16
Ζωντανοί	98	115	76	69
Σύνολο	106	120	98	85
Επίπτωση-ποσοστό θανάτου (θνητότητα)	0,076	0,042	0,224	0,188
Διάφορα θνητοτήτων	0,034		0,036	

να υπολογιστεί η διαφορά θνητοτήτων των δύο συγκρινόμενων ομάδων στις δύο επιμέρους ομάδες ηλικιών που προέκυψαν από τη διαστρωμάτωση των δεδομένων με βάση την ηλικία. Και στις δύο ομάδες ηλικιών, οι ασθενείς που ελάμβαναν τολβουταμίδη αντιμετώπιζαν έναν πρόσθετο κίνδυνο θανάτου της τάξης περίπου του 3,5% στο διάστημα των 7 ετών σε σχέση με τους ασθενείς που ελάμβαναν το εικονικό φάρμακο. Σύμφωνα με τον πίνακα 4, η διαφορά των θνητοτήτων ήταν 3,4% για τους ασθενείς <55 ετών και 3,6% για τους ασθενείς >55 ετών (ως ένα συγκεντρωτικό μέτρο υπολογίζεται ο μέσος όρος των δύο επιμέρους διαφορών στις θνητότητες, οπότε προκύπτει η μέση συνολική διαφορά, που ήταν 3,5%). Μαθηματικά, πάντως, ο μέσος όρος των δύο επιμέρους διαφορών υπολογίζεται σταθμίζοντας κάθε ηλικιακή κατηγορία ανάλογα με τον αριθμό των δεδομένων που έχει. Χωρίς την εξουδετέρωση της συγχυτικής επίδρασης της ηλικίας, η διαφορά θνητοτήτων ήταν 4,5% (πίν. 3), ενώ μετά από την εξουδετέρωση η διαφορά μειώθηκε σε 3,5%. Η τιμή 4,5%, που προέκυψε από τα αδρά δεδομένα της μελέτης, ήταν αποτέλεσμα σύγχυσης και οδήγησε σε υπερεκτίμηση του ανεπιθύμητου αποτελέσματος της τολβουταμίδης. Η τιμή 3,5%, εξάλλου, δεν εκφράζει την πλήρη εξουδετέρωση της συγχυτικής επίδρασης της ηλικίας. Και αυτό, γιατί οι ασθενείς της μελέτης χωρίστηκαν μόνο σε δύο ομάδες, οπότε ενδεχομένως στις δύο αυτές ομάδες να είχαν διατηρηθεί διαφορές στην ηλικιακή κατανομή. Έπειτα από τη διαστρωμάτωση, πάντως, η εκτίμηση του αποτελέσματος ήταν μικρότερη από εκείνη των αδρών δεδομένων. Με βάση τα αδρά δεδομένα υπερεκτιμήθηκε το ανεπιθύμητο αποτέλεσμα της τολβουταμίδης περίπου κατά 30%  $\left( \frac{4,5\% - 3,5\%}{3,5\%} = 0,30 \text{ ή } 30\% \right)$ , καθώς το 4,5% είναι περίπου κατά 30% μεγαλύτερο από το 3,5%.

Από το παραπάνω παράδειγμα προκύπτει ότι η τυχαιοποίηση δεν εξασφαλίζει απόλυτα την εξουδετέρωση της σύγχυσης σε όλες τις περιπτώσεις. Στο παράδειγμα με

τους διαβητικούς ασθενείς η τυχαιοποίηση δεν εμπόδιζε την ανισοκατανομή της ηλικίας στις δύο συγκρινόμενες ομάδες, γεγονός που προκάλεσε μέτριου βαθμού σύγχυση. Στο παράδειγμα με το σύνδρομο Down η σύγχυση ήταν αρκετά μεγαλύτερη, επειδή η σχέση μεταξύ της ηλικίας (σύγχυτή) και της σειράς γέννησης (προσδιοριστή) ήταν ισχυρότερη από τη σχέση μεταξύ της ηλικίας (σύγχυτή) και της θεραπευτικής αγωγής (προσδιοριστή) που προέκυψε από την τυχαιοποίηση στο παράδειγμα με τους διαβητικούς.

Η σύγχυση μπορεί να οδηγήσει είτε σε υπερεκτίμηση του αποτελέσματος, όπως συνέβη στα παραδείγματα με τους διαβητικούς και το σύνδρομο Down, είτε σε υποεκτίμηση. Σε ορισμένες περιπτώσεις, εξάλλου, το σφάλμα που οφείλεται στη σύγχυση μπορεί να οδηγήσει σε ακριβώς αντίθετα συμπεράσματα.

*Σύγχυση λόγω ένδειξης.* Η Φαρμακοεπιδημιολογία είναι κλάδος της Επιδημιολογίας που μελετά την αποτελεσματικότητα των φαρμάκων χρησιμοποιώντας, συχνά, μη πειραματικές μελέτες.<sup>1</sup> Στις μελέτες αυτές συγκρίνεται η διαφορά στην αποτελεσματικότητα των φαρμάκων, ενώ το θεραπευτικό αποτέλεσμα δεν εξαρτάται μόνο από το θεραπευτικό μέτρο, αλλά και από τις θεραπευτικές ενδείξεις. Χωρίς τη διεξαγωγή μιας τυχαιοποιημένης κλινικής δοκιμής, αποτελεί πρόκληση ο σχεδιασμός μιας μελέτης που να επιτρέπει την έγκυρη σύγκριση αυτών που λαμβάνουν ένα φάρμακο σε σχέση μ' εκείνους που δεν λαμβάνουν το φάρμακο αυτό. Το πρόβλημα συνίσταται κυρίως στο φαινόμενο που αναφέρεται ως «σύγχυση λόγω ένδειξης» (confounding by indication). Το πρόβλημα οφείλεται στο γεγονός ότι αυτοί που λαμβάνουν ένα φάρμακο γενικά διαφέρουν από εκείνους που δεν λαμβάνουν το φάρμακο αυτό, σύμφωνα με την ιατρική ένδειξη για την οποία το φάρμακο χορηγείται. Ακόμη και αν η συγκρινόμενη ομάδα περιλαμβάνει ασθενείς με το ίδιο νόσημα που λαμβάνουν μια διαφορετική θεραπεία, ή και καμιά θεραπεία, θα υπάρχουν διαφορές στη σοβαρότητα του νοσήματος ή σε άλλους εξωγενείς

προσδιοριστές μεταξύ των ομάδων που λαμβάνουν διαφορετικές θεραπείες. Οι διαφορές αυτές εισάγουν συστηματικό σφάλμα στη μελέτη, που καλείται σύγχυση λόγω ένδειξης.

Όπως προαναφέρθηκε, η εξουδετέρωση της σύγχυσης λόγω ένδειξης επιτυγχάνεται μόνο με πειραματικές μελέτες, δηλαδή με τις τυχαιοποιημένες κλινικές δοκιμές (randomized clinical trials). Η σύγχυση λόγω ένδειξης είναι ένα συστηματικό σφάλμα, το οποίο οφείλεται στις διαφορές στην πρόγνωση μεταξύ ασθενών που λαμβάνουν διαφορετικές θεραπευτικές αγωγές. Για παράδειγμα, ένα νέο αντιβιοτικό πιστεύεται ότι είναι αποτελεσματικότερο στην αντιμετώπιση των ανθεκτικών βακτηριακών στελεχών που προκαλούν μηνιγγίτιδα, αλλά, επίσης, εμφανίζει ανεπιθύμητες ενέργειες και κοστίζει περισσότερο. Λογικά, το νέο αυτό αντιβιοτικό χορηγείται σε ασθενείς που έχουν μεγαλύτερο κίνδυνο εμφάνισης μιας θανατηφόρας έκβασης. Ακόμη και αν το νέο αντιβιοτικό είναι εξαιρετικά αποτελεσματικό, η θνητότητα αυτών που το λαμβάνουν θα είναι μεγαλύτερη από τη θνητότητα εκείνων που λαμβάνουν τη συνήθη θεραπευτική αγωγή, επειδή εκείνοι που λαμβάνουν το νέο αντιβιοτικό βρίσκονται σε υψηλότερο κίνδυνο εμφάνισης του θανάτου. Η έγκυρη εκτίμηση της θεραπευτικής αποτελεσματικότητας του νέου αντιβιοτικού μπορεί να επιτευχθεί μόνο εάν οι διαφορές όσον αφορά στην πρόγνωση μπορούν να «διορθωθούν» ή, με άλλη διατύπωση, να εξουδετερωθούν στη σύγκριση.

Στις μη πειραματικές μελέτες, ακόμη και όταν η πληροφορία για τους δυνητικούς συγχυτές είναι επαρκής, δεν είναι δυνατόν να εξουδετερωθεί πλήρως η σύγχυση λόγω ένδειξης. Για το λόγο αυτόν, απαιτείται η διεξαγωγή τυχαιοποιημένης κλινικής δοκιμής, στην οποία η τυχαία κατανομή των μελετώμενων ατόμων έχει ως αποτέλεσμα να εξισώνονται οι προγνωστικοί παράγοντες μεταξύ των ομάδων που λαμβάνουν διαφορετικές θεραπείες.

## 5. ΕΞΟΥΔΕΤΕΡΩΣΗ ΤΗΣ ΣΥΓΧΥΣΗΣ

Η εξουδετέρωση των συγχυτών μιας μελέτης μπορεί να γίνει, είτε πριν είτε μετά από τη διεξαγωγή της, με διάφορους τρόπους, όπως η τυχαιοποίηση, η περιοριστική μέθοδος, η εξομοίωση, η πολυμεταβλητή ανάλυση, η προτύπωση και η διαστρωματική ανάλυση.<sup>1,2,5,10,24</sup>

Η *τυχαιοποίηση* (randomization) ή, με άλλη διατύπωση, η τυχαία κατανομή των μελετώμενων ατόμων στις συγκρινόμενες ομάδες μπορεί να πραγματοποιηθεί μόνο

πειραματικά. Στην *περιοριστική μέθοδο* (restriction), τα μελετώμενα άτομα έχουν την ίδια, ή περίπου την ίδια, τιμή για μια μεταβλητή που πιθανολογείται ότι είναι συγχυτής. Παραδείγματος χάριν, για την εξουδετέρωση της συγχυτικής επίδρασης του φύλου η μελέτη γίνεται μόνο στους άνδρες. Η περιοριστική μέθοδος μπορεί να χρησιμοποιηθεί σε οποιαδήποτε επιδημιολογική μελέτη, ανεξάρτητα από το αν είναι πείραμα ή όχι. Στην *εξομοίωση* (matching), η επιλογή των προσωπο-στιγμών που αποτελούν τις δύο βασικές σειρές (base series) των παρονομαστών των οιονεί επιπτώσεων-πυκνοτήτων γίνεται με τη διαδικασία της «δειγματοληψίας» του πληθυσμο-χρόνου ως προς ένα συγχυτή. Τονίζεται, όμως, ότι η εξομοίωση δεν συμβάλλει, όπως παλαιότερα πιστευόταν, στην εγκυρότητα των ονομαζόμενων μελετών «ασθενών-μαρτύρων» παρά μόνο στην εγκυρότητα προγνωστικών μελετών, όπου χρησιμοποιούνται κλειστοί πληθυσμοί.

Στα πειράματα όπου ο ερευνητής καθορίζει την έκθεση στα μελετώμενα άτομα, η *τυχαιοποίηση* παρουσιάζει σημαντικά πλεονεκτήματα. Με την τυχαιοποίηση, σ' ένα μελετώμενο πληθυσμό επαρκή αριθμητικά, προκύπτουν δύο ή περισσότερες μελετώμενες ομάδες με την ίδια περίπου κατανομή των διαφόρων χαρακτηριστικών. Η ομοιόμορφη κατανομή όλων των μεταβλητών υποδηλώνει ότι οι συγκρινόμενες ομάδες είναι όμοιες όσον αφορά στους δυνητικούς συγχυτές. Η τυχαιοποίηση, ωστόσο, δεν εξασφαλίζει με βεβαιότητα την απουσία σύγχυσης, καθώς μια τυχαία διαδικασία μπορεί να οδηγήσει σε ανισοκατανομή ορισμένων χαρακτηριστικών στις συγκρινόμενες ομάδες, όπως συνέβη με την ανισοκατανομή της ηλικίας, που φαίνεται στον πίνακα 4. Η πιθανότητα μεγάλης ανισοκατανομής, εξάλλου, μειώνεται όσο αυξάνεται ο αριθμός των μελετώμενων ατόμων που τυχαιοποιούνται. Το μεγαλύτερο πλεονέκτημα, ενδεχομένως, της τυχαιοποίησης είναι ότι προλαμβάνει τη σύγχυση που προκαλείται όχι μόνο από τους γνωστούς, αλλά και από τους άγνωστους συγχυτές. Έτσι, ακόμη και άγνωστοι δυνητικοί συγχυτές δεν εμφανίζουν πραγματική δράση σε μια τυχαιοποιημένη κλινική δοκιμή, εφόσον, βεβαίως, ο αριθμός των συμμετεχόντων είναι επαρκώς μεγάλος.

Η *περιοριστική μέθοδος*, σε αντίθεση με την τυχαιοποίηση, δεν είναι δυνατόν να εξουδετερώσει τους άγνωστους συγχυτές.<sup>1</sup> Είναι περισσότερο αξιόπιστη, ωστόσο, για την πρόληψη της σύγχυσης που προκαλείται από τους παράγοντες εκείνους για τους οποίους η περιοριστική μέθοδος εφαρμόζεται. Για παράδειγμα, σε μια μελέτη που διερευνά τη σχέση μεταξύ κατανάλωσης ονο-

πνεύματος και καρκίνου του λάρυγγα, η καπνισματική συνήθεια ενδεχομένως αποτελεί συγχυτή. Και αυτό, γιατί το κάπνισμα αφενός αποτελεί προσδιοριστή της συχνότητας εμφάνισης καρκίνου του λάρυγγα και αφετέρου τα άτομα που καταναλώνουν περισσότερο οινόπνευμα καπνίζουν, συνήθως, και περισσότερο από εκείνα που καταναλώνουν λιγότερο οινόπνευμα. Εάν η μελέτη περιοριστεί στους μη καπνιστές, τότε η καπνισματική συνήθεια δεν μπορεί να έχει συγχυτική επίδραση. Παρομοίως, εάν η ηλικία θεωρηθεί συγχυτής σε μια μελέτη, η σύγχυση αυτή μπορεί να προληφθεί με τη συμμετοχή ατόμων που είναι στην ίδια περίπου ηλικία. Με τον τρόπο αυτό, βεβαίως, τα αποτελέσματα μιας μελέτης περιορίζονται σε μια συγκεκριμένη ηλικιακή ομάδα. Εάν όλα τα μελετώμενα άτομα έχουν την ίδια, ή περίπου την ίδια, τιμή για μια μεταβλητή, τότε ο παράγοντας αυτός δεν διαφέρει, πλέον, στις μελετώμενες ομάδες και μετατρέπεται σε μια σταθερά. Για την ύπαρξη σύγχυσης, ένας συγχυτής πρέπει να ανισοκατανέμεται στις διάφορες κατηγορίες του μελετώμενου προσδιοριστή, αλλά εάν ένας παράγοντας είναι σταθερός, τότε δεν είναι δυνατόν να ανισοκατανέμεται.

Η *περιοριστική μέθοδος* μπορεί να χρησιμοποιηθεί ακόμη και στα πειράματα, επιπλέον της τυχαιοποίησης, έτσι ώστε ο ερευνητής να είναι βέβαιος για την εξουδετέρωση της σύγχυσης που προκαλείται από ορισμένους, τουλάχιστον, γνωστούς συγχυτές. Χρησιμοποιείται, επίσης, στις βασικές επιστήμες σε πειράματα με ζώα για την πρόληψη της σύγχυσης και την εξασφάλιση της εγκυρότητας. Έτσι, ένας ερευνητής που πειραματίζεται με ποντίκια, επιθυμεί τα ποντίκια αυτά να προέρχονται από το ίδιο εργαστήριο, ώστε οι διατροφικές συνθήκες να είναι οι ίδιες και, επιπλέον, να έχουν τον ίδιο γονότυπο, την ίδια ηλικία και το ίδιο φύλο.

Το αξιοσημείωτο, όμως, είναι ότι η περιοριστική μέθοδος δεν χρησιμοποιείται συχνά στην επιδημιολογική έρευνα. Μια εξήγηση είναι ότι αρκετοί ερευνητές πιστεύουν ότι μια επιδημιολογική μελέτη, είτε είναι πειραματική είτε όχι, θα πρέπει να περιλαμβάνει άτομα, τα χαρακτηριστικά των οποίων είναι αντιπροσωπευτικά του πληθυσμού-πηγή στον οποίο τα αποτελέσματα της μελέτης αναφέρονται. Η επιδίωξη της «αντιπροσωπευτικότητας» (representativeness) είναι αντίθετη προς την περιοριστική μέθοδο, στην οποία ο μελετώμενος πληθυσμός είναι ομοιογενής και, έτσι, όχι παρόμοιος με τον πληθυσμό-πηγή.<sup>1</sup> Η «αντιπροσωπευτικότητα», πάντως, είναι μια πλάνη που δημιουργεί προβλήματα στις επιδημιολογικές μελέτες από αρκετές δεκαετίες. Η αντίληψη ότι η «αντιπροσωπευτικότητα» είναι χρήσιμη οφείλεται, κυρίως, στις μελέτες εκείνες στις οποίες λαμβάνεται ένα δείγμα

από ένα μεγαλύτερο (πεπερασμένο) πληθυσμό (πληθυσμό-στόχο), έτσι ώστε να μειωθεί το οικονομικό κόστος και να αποφευχθούν τα προβλήματα που επιφέρει η μελέτη ενός ολόκληρου πληθυσμού. Το στατιστικό συμπέρασμα που στηρίζεται σε δείγματα είναι επιφανειακά μόνο όμοιο με το επιστημονικό συμπέρασμα, το οποίο αποτελεί το αντικείμενο της επιδημιολογικής έρευνας. Στην περίπτωση του επιστημονικού συμπεράσματος, το αντικείμενο δεν είναι να εξαχθεί ένα αποτέλεσμα που να ισχύει μόνο σ' ένα συγκεκριμένο πληθυσμό τοποχρονικά, αλλά αντίθετα να αναφέρεται στον αφηρημένο γενικό πληθυσμό.

Το ερώτημα που προκύπτει στην περίπτωση της περιοριστικής μεθόδου είναι αν το αποτέλεσμα που προκύπτει από τη μελέτη μιας συγκεκριμένης ομάδας ατόμων μπορεί να ισχύει και για άλλες ομάδες ατόμων με διαφορετικά χαρακτηριστικά. Για παράδειγμα, ένας ερευνητής που χρησιμοποιεί την περιοριστική μέθοδο για να διερευνήσει τη σχέση μεταξύ της κατανάλωσης κρασιού και της συχνότητας εμφάνισης καρδιαγγειακών νοσημάτων σε άτομα ηλικίας μόνο 60 ετών, μπορεί να είναι βέβαιος ότι το αποτέλεσμα ισχύει και για άτομα ηλικίας 45 ετών; Η αβεβαιότητα αυτή αφήνει περιθώριο για την εξαγωγή ενός εσφαλμένου ή ελλιπούς συμπεράσματος, αλλά αυτή ακριβώς είναι και η φύση της επιστήμης. Είναι, βεβαίως, λανθασμένο να πιστεύει κανείς ότι η θεωρία που απαιτείται για την εφαρμογή των αποτελεσμάτων μιας μελέτης σε μια ομάδα ατόμων με διαφορετικά χαρακτηριστικά μπορεί να αντικατασταθεί από μια μηχανική δειγματοληψία. Εάν ένας ερευνητής πιστεύει ότι το αποτέλεσμα της κατανάλωσης κρασιού στην εμφάνιση καρδιαγγειακών προβλημάτων είναι διαφορετικό για τα άτομα ηλικίας 45 ετών σε σχέση με τα άτομα ηλικίας 60 ετών, τότε θα πρέπει να πραγματοποιήσει μια νέα παρόμοια μελέτη με άτομα ηλικίας 45 ετών. Ο αριθμός των μελετώμενων ατόμων και η ηλικιακή τους κατανομή δεν θα αντικατοπτρίζει την ηλικιακή κατανομή του πληθυσμού-πηγή. Δεν πρέπει, πάντως, τα δημογραφικά χαρακτηριστικά ενός πληθυσμού-πηγή να καθορίζουν και την ηλικιακή κατανομή των μελετώμενων ατόμων. Αντιθέτως, ο ερευνητής είναι εκείνος που πρέπει να καθορίζει την ηλικιακή κατανομή των μελετώμενων ατόμων ανάλογα με τη θεωρία του. Επομένως, επιστημονικά δεν υπάρχει συγκεκριμένος πληθυσμός-πηγή, αλλά μια επιστημονική θεωρία που περιλαμβάνει την κατανάλωση κρασιού, τη συχνότητα εμφάνισης καρδιαγγειακών νοσημάτων και πιθανόν την ηλικία. Η διεξαγωγή μιας έγκυρης μελέτης είναι ο καλύτερος τρόπος για την εξαγωγή ασφαλών συμπερασμάτων και έτσι η περιοριστική μέθοδος, και όχι η «αντιπροσωπευτικότη-

τα», είναι η πλέον επιθυμητή μέθοδος για να επιτευχθεί ο στόχος αυτός.

Η τυχαιοποίηση, η εξομοίωση και η περιοριστική μέθοδος χρησιμοποιούνται για την εξουδετέρωση των συγχυτών πριν από τη διεξαγωγή μιας μελέτης, δηλαδή κατά το σχεδιασμό μιας μελέτης. Η εξουδετέρωση των συγχυτών, όμως, είναι δυνατόν να επιτευχθεί και κατά τη φάση της ανάλυσης των δεδομένων με την εφαρμογή της διαστρωματικής ανάλυσης ή της πολυμεταβλητής ανάλυσης. Για να γίνει, όμως, αυτό θα πρέπει τα δεδομένα μιας μελέτης να παρέχουν επαρκείς πληροφορίες για τους συγχυτές.

## 6. ΣΥΝΟΨΗ

Σε μια επιδημιολογική μελέτη υπάρχουν δύο είδη σφάλματος, το τυχαίο και το συστηματικό.<sup>4,7-10</sup> Αν και υπάρχουν αρκετά είδη συστηματικών σφαλμάτων, παραδοσιακά κατατάσσονται σε τρεις μεγάλες κατηγορίες: (α) στα *συστηματικά σφάλματα επιλογής*, (β) στα *συστηματικά σφάλματα πληροφορίας* και (γ) στη *σύγχυση*.<sup>1,5,10,11</sup>

Στα συστηματικά σφάλματα επιλογής ανήκει το συστηματικό σφάλμα της εθελοντικής συμμετοχής<sup>1,10,12,14</sup> ή, αλλιώς, το συστηματικό σφάλμα της αυτοεπιλογής, καθώς και το συστηματικό σφάλμα που είναι γνωστό ως «φαινόμενο του υγιή εργάτη».<sup>1,5,15-19</sup>

Τα συστηματικά σφάλματα πληροφορίας<sup>1,5,15,19</sup> σε μια μελέτη οφείλονται στην εσφαλμένη πληροφορία που συλλέγεται για τα μελετώμενα άτομα ή στην εσφαλμέ-

νη πληροφορία που προέρχεται από τα ίδια τα μελετώμενα άτομα. Η εσφαλμένη αυτή πληροφορία οδηγεί σε δυσταξινόμηση, όταν ο μελετώμενος προσδιοριστής είναι ποιοτικός και το σφάλμα έχει ως αποτέλεσμα ένα άτομο να τοποθετείται σε λανθασμένη κατηγορία ή τάξη. Για παράδειγμα, ένας «βαρύς» καπνιστής που ταξινομείται ως «ελαφρύς» οδηγεί σε δυσταξινόμηση. Τυπικά, οι δύο σημαντικότερες μεταβλητές σε μια επιδημιολογική μελέτη είναι ο προσδιοριστής και το νόσημα. Η δυσταξινόμηση των ατόμων, είτε αφορά στον προσδιοριστή είτε στο νόσημα, μπορεί να είναι διαφορική ή μη διαφορική.

Τέλος, η ύπαρξη σύγχυσης<sup>1,2,5,10,19,23</sup> είναι εξαιρετικά σημαντική σε μια επιδημιολογική μελέτη. Ένας απλός ορισμός της σύγχυσης θα μπορούσε να είναι η ανάμιξη ή, αλλιώς, το ανακάτεμα των αποτελεσμάτων. Ο ορισμός αυτός υποδηλώνει ότι το αποτέλεσμα του μελετώμενου προσδιοριστή αναμειγνύεται με το αποτέλεσμα ενός εξωγενούς προσδιοριστή, οδηγώντας έτσι σε συστηματικό σφάλμα. Ο πραγματικός συγχυτής θα πρέπει να σχετίζεται με τη συχνότητα εμφάνισης της μελετώμενης έκβασης και να ανισοκατανέμεται στις δύο, ή περισσότερες, κατηγορίες του μελετώμενου προσδιοριστή. Η εξουδετέρωση των συγχυτών, κατά τη φάση του σχεδιασμού των μελετών, μπορεί να πραγματοποιηθεί με την τυχαιοποίηση, την περιοριστική μέθοδο και την εξομοίωση, ενώ κατά την ανάλυση των δεδομένων μπορεί να πραγματοποιηθεί με τη διαστρωμάτωση και τα μοντέλα παλινδρόμησης. Τονίζεται, όμως, ότι η εξομοίωση δεν συμβάλλει, όπως παλαιότερα πιστευόταν, στην εγκυρότητα των ονομαζόμενων μελετών «ασθενών-μαρτύρων», παρά μόνο στην εγκυρότητα προγνωστικών μελετών όπου χρησιμοποιούνται κλειστοί πληθυσμοί.<sup>1</sup>

## ABSTRACT

### Systematic bias in epidemiological studies

P. GALANIS, L. SPAROS

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

*Archives of Hellenic Medicine 2007, 24(4):373-388*

Bias can creep into epidemiologic studies from many directions. Various types of specific bias have been described, but it is helpful to classify bias into three broad categories: selection bias, information bias, and confounding. Selection bias is a systematic bias in a study that stems from the procedures used to select subjects and from factors that influence study participation. It comes about when the association between the determinant under study and the outcome differs for those who participate and those who do not participate in the study. There are two kinds of selection bias, volunteer bias (or self-selection bias) and selection bias, often referred to as the "healthy worker effect". Information bias can arise because the information collected about or from study subjects is erroneous. 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 with the effect of another variable (extraneous determinant), leading to a bias. Extraneous determinants are referred to as "potential

confounders of the conditional relation of interest” because they are predictors of the occurrence of the outcome at issue. When extraneous determinants have different distributions among the compared categories of the determinant under study they constitute actual confounders. Confounding is a systematic error that investigators aim either to prevent or to remove from a study. In study design, there are two common methods to prevent confounding, randomization and restriction. A third method to prevent confounding is matching, which prevents confounding in studies with closed populations, but in case-control studies, surprisingly does not. Two methods that can be used to deal with confounding in the analysis of data are stratification and regression models.

**Key words:** Confounder, Confounding, Misclassification, Systematic information bias, Systematic selection bias

## Βιβλιογραφία

- ROTHMAN KJ. *Epidemiology: An introduction*. Oxford University Press, New York, 2002
- ΣΠΑΡΟΣ ΛΔ, ΓΑΛΑΝΗΣ Π, ΖΑΧΟΣ Ι, ΤΣΙΛΙΔΗΣ Κ. *Επιδημιολογία Ι*. Εκδόσεις ΒΗΤΑ, Αθήνα, 2004
- ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Βασικές αρχές ανάλυσης των επιδημιολογικών δεδομένων. *Αρχ Ελλ Ιατρ* 2005, 22:146–156
- ΓΑΛΑΝΗΣ Π. Αλληλεπίδραση παραγόντων κινδύνου στην επιδημιολογία. Μεταπτυχιακή Διπλωματική Εργασία, Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Νοσηλευτικής, Αθήνα, 2003
- ROTHMAN KJ, GREENLAND S. *Modern epidemiology*. 2nd ed. Lippincott Williams & Wilkins, Philadelphia, 1998
- ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Η έννοια του αποδοτέου κλάσματος στην εφαρμοσμένη ιατρική έρευνα. *Αρχ Ελλ Ιατρ* 2005, 22:157–169
- ΜΕΛΛΟΥ Κ, ΣΠΑΡΟΣ Λ. Τα συστηματικά σφάλματα στις αιτιολογικές επιδημιολογικές μελέτες. *Αρχ Ελλ Ιατρ* 2005, 22:199–207
- ΣΠΑΡΟΣ ΛΔ. *Μετα-επιδημιολογία ή εφαρμοσμένη ιατρική έρευνα*. Εκδόσεις ΒΗΤΑ, Αθήνα, 2001
- ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ ΛΔ. Διαστρωματική ανάλυση δεδομένων. *Αρχ Ελλ Ιατρ* 2004, 21:378–384
- ROTHMAN KJ. *Modern epidemiology*. 1st ed. Little Brown & Co, Boston, 1986
- SACKETT DL. Bias in analytic research. *J Chron Dis* 1979, 32:51–63
- BROWNSON RC, PETITTI DB. *Applied epidemiology. Theory to practice*. Oxford University Press, New York, 1998
- MIETTINEN OS. *Theoretical epidemiology. Principles of occurrence research in medicine*. John Wiley & Sons Inc, New York, 1985
- CALDWELL GG, KELLEY DB. Leukemia among participants in military maneuvers at a nuclear bomb test: A preliminary report. *JAMA* 1980, 244:1575–1578
- AHLBOM A, NORELL S. *Εισαγωγή στη σύγχρονη επιδημιολογία* (ελληνική μετάφραση). Ιατρικές Εκδόσεις Λίτσας, Αθήνα, 1992
- WANG J, MIETTINEN OS. Occupational mortality studies: Principles of validity. *Scand J Work Environ Health* 1982, 8:153–158
- FOX AJ, COLLIER PF. Low mortality rates in industrial cohort studies due to selection for work and survival in the industry. *Br J Prev Soc Med* 1976, 30:225–230
- McMICHAEL AJ. Standardized mortality ratios and the “healthy worker effect”: Scratching beneath the surface. *J Occup Environ Med* 1976, 18:165–168
- ΣΠΑΡΟΣ ΛΔ, ΛΑΜΠΡΟΥ Α, ΜΕΛΛΟΥ Κ. *Επιδημιολογία ΙΙ*. Εκδόσεις ΒΗΤΑ, Αθήνα, 2005
- ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Μέτρα συχνότητας των νοσημάτων. *Αρχ Ελλ Ιατρ* 2005, 22:178–191
- ROCKENBAUER M, OLSEN J, CZEILEL AE, PEDERSEN L, SORENSEN HT, EUROMAP GROUP. Recall bias in a case-control surveillance system on the use of medicine during pregnancy. *Epidemiology* 2001, 12:461–466
- McINTOSH AM, HOLMES S, GLEESON S, BURNS JK, HODGES AK. Maternal recall bias, obstetric history and schizophrenia. *Br J Psychiatry* 2002, 181:520–525
- ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Τροποποίηση του μέτρου αποτελέσματος και σύγκριση στην εφαρμοσμένη ιατρική έρευνα. *Αρχ Ελλ Ιατρ* 2005, 22:170–177
- ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Διαστρωματική ανάλυση δεδομένων. *Αρχ Ελλ Ιατρ* 2005, 21:378–384
- ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Στατιστική αλληλεπίδραση και τροποποίηση του μέτρου αποτελέσματος. *Αρχ Ελλ Ιατρ* 2005, 21:137–147
- ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Βιολογική και στατιστική αλληλεπίδραση. *Αρχ Ελλ Ιατρ* 2004, 21:123–136
- STARK CR, MANTEL N. Effects of maternal age and birth order on the risk of mongolism and leukemia. *J Natl Cancer Inst* 1966, 37:687–698
- 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
- ΜΠΑΜΠΙΝΙΩΤΗΣ Γ. *Λεξικό της νέας ελληνικής γλώσσας*. Κέντρο Λεξικολογίας, Αθήνα, 1998

*Corresponding author:*

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