

ΕΦΑΡΜΟΣΜΕΝΗ ΙΑΤΡΙΚΗ ΕΡΕΥΝΑ APPLIED MEDICAL RESEARCH

Προτυποποίηση#

1. Εισαγωγή
2. Η έννοια του συγχυτή
3. Η έννοια της προτυποποίησης
4. Βασικές αρχές της προτυποποίησης
5. «Άμεση» και «έμμεση» προτυποποίηση
6. Προτυποποίηση έναντι διαστρωμάτωσης
7. Σύνοψη

ΑΡΧΕΙΑ ΕΛΛΗΝΙΚΗΣ ΙΑΤΡΙΚΗΣ 2007, 24(5):471-486
ARCHIVES OF HELLENIC MEDICINE 2007, 24(5):471-486

Π. Γαλιάνης,##
Λ. Σπάρος

Εργαστήριο Κλινικής Επιδημιολογίας,
Τμήμα Νοσηλευτικής, Πανεπιστήμιο
Αθηνών, Αθήνα

Standardization

Abstract at the end of the article

Λέξεις ευρετηρίου

Διαστρωμάτωση
Προτυποποίηση
Πρότυπος πληθυσμός
Σύγχυση
Συγχυτής

Υποβλήθηκε 15.5.2006
Εγκρίθηκε 7.7.2006

1. ΕΙΣΑΓΩΓΗ

Η *Επιδημιολογία* (Epidemiology) είναι ο επιστημονικός κλάδος που μελετά την κατανομή της συχνότητας των νοσημάτων, καθώς και τα χαρακτηριστικά (συγγενή, περιβαλλοντικά, συμπεριφορικά) των ατόμων από τα οποία εξαρτάται η συχνότητα αυτή και τα οποία καλούνται παράγοντες κινδύνου ή εκθέσεις ή προσδιοριστές, όπως τελικά επικράτησε να λέγονται σήμερα.¹⁻³ 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* (determinant) 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 επιλογής του ερευνητή, άρα συνέπεια του σχεδιασμού της μελέτης. Από την επιλογή της κατηγορίας αναφοράς εξαρτάται η ερμηνεία των αποτελεσμάτων μιας μελέτης. Παραδείγματος χάριν, 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».

Σκοπός μιας επιδημιολογικής μελέτης είναι η διερεύνηση της σχέσης μεταξύ ενός προσδιοριστή και της μελετώμενης έκβασης. Στις περισσότερες μελέτες, ωστόσο, εξαιτίας εσφαλμένου σχεδιασμού, υπάρχουν ορισμένα χαρακτηριστικά των ατόμων που καλούνται συγχυτές και τα οποία συγχέουν τη σχέση μεταξύ προσδιοριστή και έκβασης, οδηγώντας, έτσι, σε εσφαλμένα συμπεράσματα. Ο σχεδιασμός των μελετών περιλαμβάνει το

* Ο όρος «έκβαση» στην Επιδημιολογία χρησιμοποιείται για να δηλώσει άλλοτε την εμφάνιση ενός νοσήματος (health outcome) και άλλοτε την ίαση ή το θάνατο ή την εμφάνιση καταλοίπων από ένα νόσημα (illness outcome).

Μερική χρηματοδότηση: Ειδικός Λογαριασμός Κονδυλίων Έρευνας Πανεπιστημίου Αθηνών

Ο Π. Γαλιάνης είναι υπότροφος του Ιδρύματος Α.Σ. ΩΝΑΣΗ

σχεδιασμό του αντικειμένου και της μεθόδου κάθε μελέτης. Το αντικείμενο της εφαρμοσμένης ιατρικής έρευνας είναι και στις τρεις ερευνητικές κατηγορίες (αιτιόγνωση, διάγνωση και πρόγνωση) μια συνάρτηση συχνότητας (occurrence function). Η μέθοδος είναι ο τρόπος προσέγγισης του αντικειμένου της μελέτης. Σκοπός του σχεδιασμού μιας επιδημιολογικής μελέτης είναι η μεγιστοποίηση ή αριστοποίηση της καθολικής (συνολικής, ολικής) ποιότητας της μελέτης. Το επιθυμητό (desideratum), κατά το σχεδιασμό του αντικειμένου της μελέτης, είναι η επιστημονική αποδοχή (αντικειμενικότητα, ορθολογικότητα) και η συνάφεια του αντικειμένου με τη φροντίδα υγείας, ενώ τα επιθυμητά (desiderata) κατά το σχεδιασμό της μεθόδου της μελέτης είναι η εγκυρότητα (απουσία μεροληψίας) και η αποδοτικότητα. Η εγκυρότητα της μεθόδου προηγείται της εγκυρότητας του αποτελέσματος ή, με άλλη διατύπωση, η εγκυρότητα του αποτελέσματος προϋποθέτει την εγκυρότητα της μεθόδου.

Μεροληψία των μεθόδων (bias in methods)³ είναι το σφάλμα (ή άλλη ποιοτική ατέλεια) που επηρεάζει τις διάμεσες τιμές των «επανελημμένων κατανομών» (replication distributions) των εμπειρικών τιμών των παραμέτρων που μελετώνται στη συνάρτηση συχνότητας. Το σφάλμα αυτό προκαλεί απόκλιση των διαμέσων από τις αντίστοιχες θεωρητικές τιμές (οι αποκλίσεις αυτές παριστάνουν μεροληψίες στα αποτελέσματα).

Οι τύποι των μεροληψιών³ (bias) ή, με άλλη διατύπωση, των συστηματικών σφαλμάτων είναι οι εξής:

- *Μεροληψία επιλογής* (selection bias) ή μεροληψία της «βάσης μελέτης»* (study-base bias) ή μεροληψία αποδοχής (admission bias) ή μεροληψία εγγραφής (enrollment bias). Στις περιπτώσεις αυτές, το αποτέλεσμα είναι αναπόφευκτα μεροληπτικό ακόμη και όταν η μελέτη αφορά σε περιγραφικούς σκοπούς.
- *Μεροληψία πληροφορίας* (information, observation ή documentation bias). Η συλλογή των δεδομένων ή και η σύνοψή τους είναι τέτοια, που το αποτέλεσμα

της μελέτης είναι μεροληπτικό ακόμη και όταν η μελέτη αφορά σε περιγραφικούς σκοπούς.

- *Συγχυτική μεροληψία* (σε αιτιακή μελέτη). Ο ατελής έλεγχος της σύγχυσης που υπάρχει στη «βάση μελέτης», ιδιαίτερα η συστηματική (αναμενόμενη, μη τυχαία) σύγχυση, που καθιστά μη ερμηνεύσιμα ακόμη και τα περιγραφικά ορθά αποτελέσματα.
- *Μεροληψία υπερπαραμετροποίησης* (overparametrization bias) σε περιγραφικές μελέτες. Μεγάλος αριθμός ανεξάρτητων παραμέτρων στο μοντέλο εξάρτησης (regression model) σε σχέση με τον αριθμό των «μονάδων παρατήρησης» ή, με άλλη διατύπωση, τον αριθμό των μελετώμενων ατόμων, που καθιστά μεροληπτική την εμπειρική τιμή (fitting-based value) της εξαρτημένης παραμέτρου (πολύ υψηλή όταν είναι υψηλή ή πολύ μικρή όταν είναι μικρή).

Μια καλά σχεδιασμένη μελέτη απαιτεί την εξουδετέρωση των συγχυτών, η οποία μπορεί να γίνει με διάφορες μεθόδους,** όπως η διαστρωμάτωση, η πολυπαραμετρική ανάλυση, η προτυποποίηση, η περιοριστική μέθοδος και το πείραμα.

2. Η ΕΝΝΟΙΑ ΤΟΥ ΣΥΓΧΥΤΗ

Η *σύγχυση* (confounding) στις επιδημιολογικές μελέτες ορίζεται ως η απόκλιση ενός μέτρου σχέσης***

** Στις μεθόδους εξουδετέρωσης των συγχυτών ανήκει και η *εξομοίωση* (matching), αλλά πρέπει να τονιστεί ότι δεν συμβάλλει, όπως παλαιότερα πιστευόταν, στην εγκυρότητα των ονομαζόμενων μελετών «ασθενών-μαρτύρων», παρά μόνο στην εγκυρότητα των προγνωστικών μελετών όπου χρησιμοποιούνται κλειστοί πληθυσμοί.

*** 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 στο μελετώμενο προσδιοριστή. 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 προσδιοριστής, οπότε θα ήταν δυνατόν να υπολογιστούν τα μέτρα αποτελέσματος. Επειδή, όμως, η κατάσταση αυτή είναι ανέφικτη, είναι γνωστή ως *υποθετικές προτάσεις του μη πραγματικού* (counterfactuals).^{2,4,5} 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

* Ο *πληθυσμός-πηγή*² (source population) αποτελεί την πηγή των υποκειμένων μιας συγκεκριμένης μελέτης. Με άλλη διατύπωση, 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* (study population), στο σύνολο, δηλαδή, των ατόμων που παρατηρούνται από τον ερευνητή. Ο μελετώμενος πληθυσμός μπορεί να οριστεί σωστότερα ως ο πληθυσμός, η εμπειρία του οποίου διαμορφώνει τη «βάση μελέτης» (study-base).^{2,3} Με το δυσκολομετάφραστο αυτόν όρο νοείται η εμπειρία του μελετώμενου πληθυσμού, όπως αυτός κινείται στο χρόνο. Ουσιαστικά, η «βάση μελέτης» αποτελεί την τάξη αναφοράς (πληθυσμο-χρόνο) του αντικειμένου της μελέτης.

που προκαλείται από έναν εξωγενή προσδιοριστή και συμβαίνει μόνο μέσα στα πλαίσια μιας συγκεκριμένης μελέτης. Τα κυριότερα μέτρα σχέσης σε μια αιτιολογική μελέτη είναι (α) η διαφορά των επιπτώσεων-πυκνότητων (IDe-IDo) και (β) ο λόγος των επιπτώσεων-πυκνότητων (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 σύγκριση δεν είναι ένα χαρακτηριστικό του κάθε εξωγενούς 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.^{2,6-11} Ο συγχυτής (confounder) πρέπει να έχει τα παρακάτω χαρακτηριστικά:^{3-5,12-14}

- Να σχετίζεται με τη νόσο (να αποτελεί αιτία ή απλώς προσδιοριστή της νόσου, αλλά όχι αποτέλεσμα της νόσου)

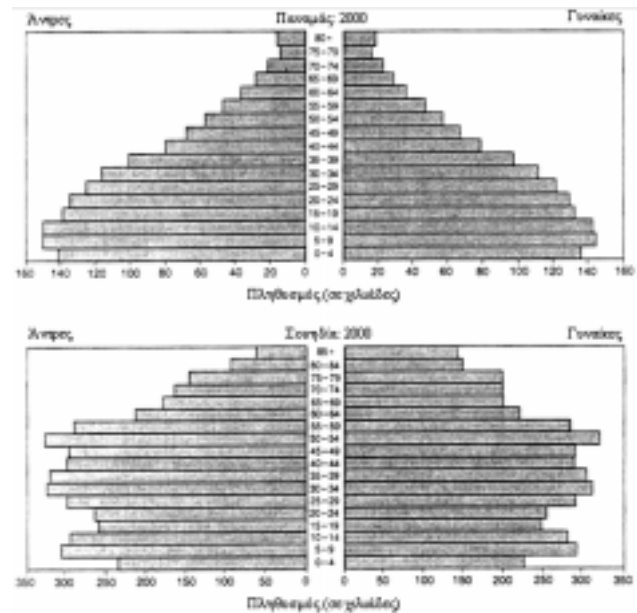
επιπτώσεις-ποσοστά αντικατοπτρίζει το αποτέλεσμα της έκθεσης στο μελετώμενο παράγοντα. Εάν, π.χ., 100 άτομα είναι εκτεθειμένα και τα 25 εμφανίσουν τη νόσο σε χρονικό διάστημα ενός έτους, τότε η επίπτωση-ποσοστό είναι $25/100=0,25$ σε ένα έτος. Αντικείμενο της μελέτης είναι η σύγκριση αυτής της επίπτωσης-ποσοστού με εκείνη που θα υπολογιζόταν αν τα ίδια 100 άτομα, την ίδια ακριβώς χρονική περίοδο και κάτω από τις ίδιες ακριβώς συνθήκες, δεν ήταν εκτεθειμένα. Εάν στην κατηγορία των μη εκτεθειμένων εμφανίζαν 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 του μελετώμενου προσδιοριστή. 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. Αντίθετα, 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 αδύνατον και γι' αυτό σε μια μελέτη υπολογίζονται πάντοτε τα μέτρα σχέσης. Το αποτέλεσμα (effect) στις εμπειρικές επιστήμες αποτελεί τη μεταβολή της παραμέτρου της έκβασης που προκαλείται από μια αιτία ή ενός προληπτικού μέτρου παρέμβασης, ενώ στη στατιστική αποτελεί την τιμή της παραμέτρου της σχέσης.

* Οι εξωγενείς προσδιοριστές (extraneous determinants) καλούνται δυνητικοί συγχυτές (potential confounders),³ καθώς σχετίζονται με τη συχνότητα εμφάνισης της μελετώμενης έκβασης. Όταν οι εξωγενείς αυτοί προσδιοριστές ανισοκατανέμονται ήδη στις κατηγορίες του μελετώμενου προσδιοριστή, τότε καλούνται πραγματικοί συγχυτές (actual confounders).

** Οι συγχυτές είναι συστηματικά σφάλματα που μπορούν, εφόσον διαπιστωθούν, να εξουδετερωθούν (ή ελεγχθούν) κατά την ανάλυση των δεδομένων, κάτι που δεν μπορεί να γίνει στα συστηματικά σφάλματα επιλογής ή πληροφορίας.

- 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 πλούσια σε κορεσμένα λίπη διαίτα, 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 (πλούσια σε κορεσμένα λίπη διαίτα). Η αυξημένη LDL, ωστόσο, δεν είναι συγχυτής της σχέσης μεταξύ διαίτας και αθηρωματικής νόσου, γιατί αποτελεί αποτέλεσμα του μελετώμενου προσδιοριστή, καθώς η πλούσια σε κορεσμένα λίπη διαίτα προκαλεί αύξηση της LDL.

Στην εικόνα 1⁵ φαίνεται η ηλικιακή κατανομή των πληθυσμών της Σουηδίας και του Παναμά το 2000. Σύμφωνα με την κοινή λογική, η θνησιμότητα*** στη Σουηδία θα πρέπει να είναι μικρότερη από τη θνησιμότητα στον Παναμά. Και αυτό γιατί η ποιότητα ζωής και η



Εικόνα 1. Ηλικιακή κατανομή των πληθυσμών του Παναμά και της Σουηδίας, το 2000.⁵

*** Όταν η μελετώμενη έκβαση είναι ο θάνατος, τότε η επίπτωση-πυκνότητα του θανάτου καλείται θνησιμότητα (mortality).^{5,15-17} 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.

φροντίδα υγείας στη Σουηδία είναι καλύτερες σε σχέση με τον Παναμά. Το αξιοσημείωτο, όμως, είναι ότι η αδρή γενική θνησιμότητα, κάθε χρόνο, είναι μεγαλύτερη στη Σουηδία απ' ό,τι στον Παναμά, γεγονός που έρχεται σε αντίθεση με την κοινή λογική. Το γεγονός αυτό οφείλεται, κατά κύριο λόγο, στη διαφορετική ηλικιακή κατανομή των δύο χωρών. Παρατηρώντας, στην εικόνα 1, την ηλικιακή κατανομή των κατοίκων του Παναμά, εύκολα διαπιστώνει κάποιος ότι ο αριθμός των κατοίκων μειώνεται σημαντικά από τις νεότερες στις γηραιότερες ομάδες, γεγονός που υποδηλώνει ότι οι περισσότεροι Παναμέζοι ανήκουν στις νεότερες ηλικιακές ομάδες. Δεν συμβαίνει, όμως, κάτι ανάλογο και με τον πληθυσμό της Σουηδίας. Πιο συγκεκριμένα, μέχρι την ηλικία των 60 ετών, ο αριθμός των Σουηδών στις διάφορες ηλικιακές ομάδες δεν παρουσιάζει εντυπωσιακές αυξομειώσεις και είναι περίπου σταθερός, ενώ μετά από την ηλικία των 60 ετών παρουσιάζει προοδευτική ελάττωση. Όπως γίνεται σαφές από την εικόνα 1, η ηλικιακή κατανομή των Σουηδών είναι γηραιότερη από εκείνη των Παναμέζων, οπότε ο πληθυσμός της Σουηδίας είναι περισσότερο γηρασμένος από τον πληθυσμό του Παναμά. Η ηλικιακή αυτή διαφορά στις δύο χώρες εξηγεί γιατί η αδρή γενική θνησιμότητα στη Σουηδία είναι μεγαλύτερη από την αδρή γενική θνησιμότητα στον Παναμά. Έτσι, η ηλικία αποτελεί συγχυτή της σχέσης μεταξύ της χώρας διαμονής (προσδιοριστή) και της συχνότητας εμφάνισης του θανάτου (έκβασης). Πρέπει, εξάλλου, να σημειωθεί ότι οι ειδικές θνησιμότητες σε κάθε ηλικιακή ομάδα είναι μικρότερες στους Σουηδούς σε σχέση με τους Παναμέζους.

Η σύγκριση της αδρής γενικής θνησιμότητας (crude overall mortality)* στις δύο χώρες δεν μπορεί να οδηγήσει σε ασφαλή συμπεράσματα, καθώς απαιτείται η εξουδετέρωση της συγχυτικής επίδρασης, τουλάχιστον, της ηλικίας. Όπως προαναφέρθηκε, η εξουδετέρωση των συγχυτών μιας μελέτης μπορεί να γίνει με διάφορες μεθόδους, όπως η διαστρωμάτωση, η πολυμεταβλητή ανάλυση, η προτυποποίηση, η περιοριστική μέθοδος και το πείραμα.^{2,3,5,12,13,18-20} Η παρούσα ανασκόπηση θα περιοριστεί στην ανάλυση της μεθόδου της προτυποποίησης.

* Τονίζεται ότι η γενική (overall) θνησιμότητα (ή νοσηρότητα) αντιδιαστέλλεται προς την ειδική (specific) θνησιμότητα (π.χ. θνησιμότητα κατά φύλο ή ηλικία), ενώ η αδρή γενική (crude overall) θνησιμότητα αντιδιαστέλλεται προς τη διορθωμένη ή προσαρμοσμένη (adjusted) γενική θνησιμότητα. Η προτυποποιημένη γενική θνησιμότητα είναι η γενική θνησιμότητα διορθωμένη με τρόπο ώστε να έχει τα ίδια «βάρη» με μία ή πολλές άλλες γενικές θνησιμότητες (είτε αδρές είτε διορθωμένες).

3. Η ΕΝΝΟΙΑ ΤΗΣ ΠΡΟΤΥΠΟΠΟΙΗΣΗΣ

Η *προτυποποίηση* (standardization) χρησιμοποιείται για την εξουδετέρωση της συγχυτικής επίδρασης της ηλικίας και του φύλου, και ενδεχομένως και άλλων χαρακτηριστικών, κατά τη σύγκριση των μέτρων συχνότητας που υπολογίζονται σε δύο ή περισσότερους πληθυσμούς. Ο William Farr** ήταν ο πρώτος που ασχολήθηκε συστηματικά με τη μέθοδο της προτυποποίησης, διακρίνοντας ότι η σύγκριση αδρών θνησιμοτήτων δεν οδηγεί στην εξαγωγή ασφαλών συμπερασμάτων και εισάγοντας, μάλιστα, και την έννοια του πρότυπου πληθυσμού.

Η προτυποποίηση*** είναι μια μέθοδος που συνδυάζει τις συχνότητες των ειδικών κατηγοριών σε μία και μοναδική τιμή, λαμβάνοντας υπόψη ένα σταθμισμένο μέσο όρο των συχνοτήτων αυτών.^{3,5,12,22-25} Η προτυποποίηση**** σταθμίζει τις συχνότητες των ειδικών κατηγοριών χρησιμοποιώντας «βάρη» (weights) που προέρχονται από έναν «πρότυπο» (standard) πληθυσμό. Τα

** Ο Farr²¹ (1807-1883) θεωρείται ως ένας από τους θεμελιωτές της σύγχρονης Επιδημιολογίας, μελετώντας, κυρίως, τους κλινικούς πίνακες επιβίωσης και τη στατιστική των γάμων, των γεννήσεων και των θανάτων. Συνέκρινε τις θνησιμότητες μεταξύ διαφορετικών γεωγραφικών περιοχών, εξουδετερώνοντας ταυτόχρονα τη συγχυτική επίδραση της ηλικίας και του φύλου, καθώς επίσης και τις θνησιμότητες διαφορετικών επαγγελματικών ομάδων, βοηθώντας, έτσι, στην κατανόηση των κινδύνων που οφείλονται σε επαγγελματικούς προσδιοριστές. Ο Farr ήταν εκείνος που μετέφρασε, ουσιαστικά, το αγγλικό σύστημα καταγραφής των θανάτων σ' ένα εργαλείο μέτρησης της υγειονομικής κατάστασης της Αγγλίας, ενώ η συνεργασία του με τη Florence Nightingale (1820-1910) συνέβαλε αποφασιστικά στην προαγωγή της υγείας και στη βελτίωση της υγιεινής. Μελέτησε τους κλινικούς πίνακες επιβίωσης για να εκτιμήσει την αποτελεσματικότητα των προληπτικών μέτρων στην αύξηση του προσδόκιμου ζωής, ενώ ήταν ο πρώτος που περιέγραψε τις επιδημικές καμπύλες (epidemic curves). Τέλος, στην επιδημία της χολέρας, το 1854, προειδοποίησε τις αγγλικές εταιρείες ύδρευσης για την πιθανότητα παροχής μολυσμένου νερού.

*** Πρέπει να τονιστεί η διαφορά μεταξύ προτύπωσης και προτυποποίησης, καθώς προτύπωση καλείται η προεικόνιση ή, αλλιώς, ο προηγούμενος σχηματισμός, ενώ προτυποποίηση καλείται (α) η μετατροπή σε πρότυπο ή, αλλιώς, η προβολή ενός πράγματος ως προτύπου, (β) η παραγωγή, διανομή και καθιέρωση προτύπων αγαθών και (γ) ο προσδιορισμός των προδιαγραφών, οι οποίες πρέπει να τηρούνται κατά την παραγωγή προϊόντος.²⁶

**** Σύμφωνα με τον Τριχόπουλο,²⁵ είναι σκόπιμος ο υπολογισμός ενός συνιστάμενου (σύνθετου) δείκτη για καθέναν από τους συγκρινόμενους πληθυσμούς, που να μην εξαρτάται από τη διαφορετική κατά ηλικία σύνθεση των πληθυσμών αυτών, κάτι που επιτυγχάνεται με την τεχνική της προτύπωσης ή προτυποποίησης κατά ηλικία. Ο Κατσουγιαννόπουλος²³ αναφέρει ότι η προτυποποίηση ή προτύπωση ή σταθεροποίηση ή ρύθμιση ή προσαρμογή αποτελεί μια βιοστατιστική τεχνική εξισορρόπησης και, παράλληλα, άρσης της επίδρασης της διαφορετικής σύνθεσης κατά φύλο και ηλικία των συγκρινόμενων ομάδων. Επιπλέον, οι Παπαευσταθίου και Τσίμπος²² αναφέρονται στην «άμεση» και «έμμεση» προτυποποίηση, καθώς και στον υπολογισμό των προτυποποιημένων δεικτών και ιδιαίτερα του προτυποποιημένου δείκτη θνησιμότητας.

«βάρη», ουσιαστικά, καθορίζουν τον πρότυπο πληθυσμό. Παραδείγματος χάριν, για τη στάθμιση των συχνοτήτων στις ειδικές ηλικιακές κατηγορίες ενός πληθυσμού Α με βάση έναν πρότυπο πληθυσμό επιλέγεται ως πρότυπος ο πληθυσμός των ΗΠΑ το 2000. Η επιλογή αυτή συνεπάγεται ότι τα «βάρη» που θα χρησιμοποιηθούν για τον υπολογισμό του σταθμισμένου μέσου όρου των ειδικών συχνοτήτων θα αντανακλούν την κατανομή του πληθυσμού των ΗΠΑ το 2000. Έτσι, η προτυποποίηση είναι μια διαδικασία στάθμισης των συχνοτήτων σε δύο ή περισσότερες κατηγορίες χρησιμοποιώντας μια συγκεκριμένη ομάδα «βαρών». Σ' έναν πληθυσμό Α, η θνησιμότητα για τους άνδρες είναι $\frac{10}{1000}(\text{έτη})^{-1}$, ενώ για τις γυναίκες είναι $\frac{5}{1000}(\text{έτη})^{-1}$. Είναι δυνατόν να σταθμίσουν οι δύο αυτές συχνότητες των επιμέρους κατηγοριών, χρησιμοποιώντας οποιοδήποτε πρότυπο πληθυσμό. Ένας πρότυπος πληθυσμός θα μπορούσε να είναι εκείνος που θα στάθμιζε ισοδύναμα τους άνδρες και τις γυναίκες. Έτσι, ο σταθμισμένος μέσος όρος των δύο

επιμέρους συχνοτήτων θα ήταν $\left(\frac{1}{2} \times \frac{10}{1000}\right) + \left(\frac{1}{2} \times \frac{5}{1000}\right) = \frac{7,5}{1000}(\text{έτη})^{-1}$. Εάν υποθεθεί ότι ο πληθυσμός Α είναι το νοσηλευτικό προσωπικό στο οποίο το 95% είναι γυναίκες, τότε ένας πρότυπος πληθυσμός θα στάθμιζε τους άνδρες μ' ένα «βάρος» της τάξης του 5% και τις γυναίκες μ' ένα «βάρος» της τάξης του 95%. Επομένως, ο σταθμισμένος μέσος όρος των δύο επιμέρους συχνοτήτων ή, με άλλη διατύπωση, η προτυποποιημένη γενική συχνότητα θα ήταν $\left(\frac{5}{100} \times \frac{10}{1000}\right) + \left(\frac{95}{100} \times \frac{5}{1000}\right) = \frac{5,25}{1000}(\text{έτη})^{-1}$.

Εάν οι συχνότητες στις ειδικές κατηγορίες είναι παρόμοιες, τότε η επιλογή των «βαρών» έχει μικρή σημασία. Εάν υποθεθεί ότι στον πληθυσμό Α οι ειδικές θνησιμότητες για τους άνδρες και τις γυναίκες ήταν $\frac{8}{1000}(\text{έτη})^{-1}$, τότε η προτυποποιημένη γενική θνησιμότητα, έπειτα από τη στάθμιση για το φύλο, θα ήταν $\frac{8}{1000}(\text{έτη})^{-1}$.

Στην περίπτωση αυτή, η επιλογή του πρότυπου πληθυσμού δεν είναι ιδιαίτερα σημαντική. Όταν, όμως, οι ειδικές συχνότητες διαφέρουν σημαντικά στις διάφορες κατηγορίες, τότε η επιλογή των «βαρών», και κατ' επέκταση η επιλογή του πρότυπου πληθυσμού, μπορεί να επηρεάσει σημαντικά το συνολικό αποτέλεσμα. Εάν ο πρότυπος πληθυσμός συνδυάζει μεγάλα «βάρη» με κατηγο-

ρίες στις οποίες οι συχνότητες έχουν υψηλές τιμές, τότε η προτυποποιημένη συχνότητα θα είναι μεγάλη, ενώ, αντίθετα, εάν ο πρότυπος πληθυσμός συνδυάζει μεγάλα βάρη με κατηγορίες στις οποίες οι συχνότητες έχουν μικρές τιμές, τότε η προτυποποιημένη συχνότητα θα είναι μικρή. Ορισμένοι ερευνητές προτείνουν να μην υπολογίζεται η προτυποποιημένη συχνότητα εφόσον η τιμή της εξαρτάται, σε μεγάλο βαθμό, από την επιλογή των «βαρών». Αρκετοί, ωστόσο, υποστηρίζουν ότι είναι χρήσιμος ο υπολογισμός της προτυποποιημένης συχνότητας, αρκεί όμως να αναφέρεται με σαφήνεια ο πρότυπος πληθυσμός που «δάνεισε» τα «βάρη» του. Σε κάθε περίπτωση, πάντως, η προτυποποίηση χρησιμοποιείται κατά κύριο λόγο για την πραγματοποίηση συγκρίσεων, οπότε υπάρχουν δύο τουλάχιστον ή περισσότερες ομάδες συχνοτήτων που πρόκειται να σταθμιστούν. Στην περίπτωση που πρόκειται να συγκριθούν συχνότητες σε εκτεθειμένα και μη εκτεθειμένα άτομα, θα πρέπει και οι δύο ομάδες να σταθμιστούν ως προς τον ίδιο πρότυπο πληθυσμό.

4. ΒΑΣΙΚΕΣ ΑΡΧΕΣ ΤΗΣ ΠΡΟΤΥΠΟΠΟΙΗΣΗΣ

Η προτυποποίηση είναι μια μέθοδος που χρησιμοποιείται, κυρίως, για την εξουδετέρωση της συγχυτικής επίδρασης της ηλικίας σε μια συγκριτική μελέτη.

Για παράδειγμα, αντικείμενο μιας μελέτης είναι η σύγκριση των θνησιμοτήτων στη Σουηδία και τον Παναμά το 2000. Οι πληθυσμοί, το 2000, των δύο χωρών ήταν ανοικτοί,* οπότε οι θνησιμότητες υπολογίστηκαν ως επιπτώσεις-πυκνότητες. Ο μέσος πληθυσμός της Σουηδίας ήταν 10.000.000 άτομα και στη διάρκεια του 2000

* Ανοικτός (open) ή δυναμικός (dynamic) πληθυσμός είναι ο πληθυσμός μιας πόλης ή χώρας, οι νοσηλεύόμενοι σ' ένα νοσοκομείο, τα μέλη μιας ασφαλιστικής εταιρείας κ.ά.^{1-3,5,12,27} Τα μέλη του ανοικτού πληθυσμού εναλλάσσονται στο χρόνο. Η ιδιότητα αυτή δικαιολογεί και το χαρακτηρισμό του ως ανοικτού. Ο ανοικτός πληθυσμός δεν γράφεται. Η μέση ηλικία, π.χ., των νοσηλευόμενων στον «Ευαγγελισμό» ή των κατοίκων της Αθήνας παραμένει διαχρονικά σταθερή, παρά τη συχνή ανανέωση των μελών τους. Το ίδιο ισχύει και για άλλα χαρακτηριστικά του πληθυσμού, όπως η κατανομή κατά φύλο. Εκείνο όμως που βαθύτερα χαρακτηρίζει τον ανοικτό πληθυσμό και αποτελεί το κύριο στοιχείο του ορισμού του, είναι ότι η ιδιότητα του μέλους 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 σε αυτή. Χάνει, όμως, την ιδιότητα του μέλους εφόσον απομακρυνθεί από αυτήν. Στον ανοικτό πληθυσμό επιτρέπεται τόσο η είσοδος όσο και η έξοδος μελών κατά τη διάρκεια της κίνησης των πληθυσμών στο χρόνο (δυναμικός πληθυσμός). Η μέτρηση της συχνότητας ενός νοσήματος ή του θανάτου σ' έναν ανοικτό πληθυσμό γίνεται με το μέτρο που φέρεται ως επίπτωση (incidence rate) ή επίπτωση-πυκνότητα (incidence density).

συνέβησαν 100.000 θάνατοι, οπότε η αδρή γενική θνησιμότητα ήταν $\frac{100.000}{10.000.000}=0,01$ (έτη)⁻¹. Ο μέσος πληθυσμός του Παναμά ήταν 1.500.000 άτομα και στη διάρκεια του 2000 συνέβησαν 10.000 θάνατοι, οπότε η αδρή γενική θνησιμότητα ήταν $\frac{10.000}{1.500.000}=0,0066$ (έτη)⁻¹. Η αδρή γενική θνησιμότητα στη Σουηδία ήταν 1,5 φορές μεγαλύτερη από την αδρή γενική θνησιμότητα στον Παναμά. Με βάση το αποτέλεσμα αυτό, θα μπορούσε κάποιος να ισχυριστεί ότι η ζωή στη Σουηδία, το 2000, ήταν περισσότερο «επικίνδυνη» από τη ζωή στον Παναμά. Το συμπέρασμα αυτό, βέβαια, είναι αντίθετο στην κοινή λογική, καθώς η ποιότητα ζωής και η φροντίδα υγείας στη Σουηδία, όχι μόνο το 2000 αλλά και γενικότερα, είναι πολύ καλύτερες απ' ό,τι στον Παναμά. Για το λόγο αυτόν, είναι απαραίτητο να ελεγχθεί η τυχόν συγχυτική επίδραση της ηλικιακής κατανομής των πληθυσμών των δύο χωρών στη σχέση μεταξύ της χώρας διαμονής και της συχνότητας του θανάτου. Στον πίνακα 1 φαίνονται οι θνησιμότητες στη Σουηδία και τον Παναμά, το 2000, ανάλογα με την ηλικία, η οποία χωρίστηκε σε τρεις ομάδες (άτομα ηλικίας 0–29 ετών, άτομα ηλικίας 30–59 ετών και άτομα >60 ετών).

Για τα άτομα <60 ετών η θνησιμότητα ήταν αρκετά μεγαλύτερη στον Παναμά, ενώ για τα άτομα >60 ετών η θνησιμότητα στη Σουηδία ήταν περίπου η ίδια με τη θνησιμότητα στον Παναμά. Το συμπέρασμα αυτό είναι ακριβώς αντίθετο με το συμπέρασμα που προέκυψε από τη σύγκριση των αδρών γενικών θνησιμοτήτων στις δύο χώρες. Από τον πίνακα 1 εύκολα διαπιστώνει κάποιος ότι οι ηλικιακές κατανομές στις δύο χώρες διαφέρουν σημαντικά μεταξύ τους. Πιο συγκεκριμένα, το 69% των Παναμέζων ήταν <30 ετών, ενώ το αντίστοιχο ποσοστό για τους Σουηδούς ήταν μόλις 42%. Επιπλέον, το 17,3% των Σουηδών ήταν >60 ετών, ενώ το αντίστοιχο ποσοστό για τους Παναμέζους ήταν μόλις 5,5%. Είναι προφανές ότι ο πληθυσμός του Παναμά, το 2000, ήταν μικρότερος ηλικιακά από εκείνον της Σουηδίας. Η ηλικιακή αυτή διαφορά στις δύο χώρες εξηγεί και τη μικρότερη αδρή γενική θνησιμότητα στον Παναμά σε σχέση με τη Σουηδία, καθώς τα νεότερα άτομα τείνουν να έχουν μικρότερη θνησιμότητα από τα άτομα μεγαλύτερης ηλικίας. Έτσι, η σύγκριση μεταξύ των δύο χωρών αποτελεί, ουσιαστικά, σύγκριση μεταξύ νεότερων (Παναμέζων) και γηραιότερων (Σουηδών) ατόμων.

Όταν ο συνολικός αριθμός θανάτων σ' έναν πληθυσμό διαιρείται με το συνολικό πληθυσμο-χρόνο παρακολούθησης όλων των ατόμων του πληθυσμού, τότε η επίπτωση-πυκνότητα θανάτου που υπολογίζεται καλείται

Πίνακας 1. Υποθετικές θνησιμότητες σε τρεις ηλικιακές ομάδες στη Σουηδία και τον Παναμά, το 2000.

	Ηλικία (έτη)		
	0–29	30–59	>60
<i>Πληθυσμός</i>			
Σουηδία	4.200.000	4.070.000	1.730.000
Παναμάς	1.034.000	383.720	82.279
<i>Θάνατοι</i>			
Σουηδία	4.790	14.857	80.353
Παναμάς	5.480	1.097	3.423
<i>Θνησιμότητα [(έτη)⁻¹]</i>			
Σουηδία	0,0011	0,0036	0,046
Παναμάς	0,0053	0,0028	0,0416

αδρή γενική θνησιμότητα. Ένα αδρό μέτρο συχνότητας μετρά την πραγματική εμπειρία ενός πληθυσμού. Στο παράδειγμα με τη Σουηδία και τον Παναμά υπολογίστηκε η αδρή γενική επίπτωση-πυκνότητα θανάτου ή, με άλλη διατύπωση, η αδρή γενική θνησιμότητα. Όπως φάνηκε και στο παράδειγμα αυτό, η σύγκριση αδρών θνησιμοτήτων (ή, γενικότερα, οποιωνδήποτε αδρών μέτρων συχνότητας) είναι δυνατόν να οδηγήσει σε εσφαλμένα συμπεράσματα, καθώς μπορεί να επηρεάζεται από διαφορές σε άλλους παράγοντες (π.χ. ηλικία). Η σύγκριση των αδρών γενικών θνησιμοτήτων στη Σουηδία και τον Παναμά υποδηλώνει ότι η αδρή γενική θνησιμότητα στη Σουηδία, το 2000, ήταν 50% μεγαλύτερη από εκείνη του Παναμά, αλλά, όπως έγινε εύκολα κατανοητό, η σύγκριση αυτή επηρεάζεται από το ότι ο πληθυσμός του Παναμά ήταν νεαρότερος από εκείνον της Σουηδίας. Η εξουδετέρωση της συγχυτικής επίδρασης της ηλικίας επιτυγχάνεται με διάφορες μεθόδους, μια από τις οποίες είναι η προτυποποίηση.

Η βασική αρχή της προτυποποίησης είναι ο υπολογισμός των υποθετικών αδρών συχνοτήτων για κάθε ομάδα από τις συγκρινόμενες με τη χρήση της ίδιας «τεχνητής» κατανομής για το συγχυτή. Η «τεχνητή» κατανομή είναι γνωστή ως «πρότυπη». Για την ασφαλή σύγκριση των γενικών θνησιμοτήτων στη Σουηδία και τον Παναμά, το 2000, εξουδετερώνοντας τη συγχυτική επίδραση της ηλικίας, απαιτείται η επιλογή ενός «πρότυπου πληθυσμού» όσον αφορά στην ηλικία. Ο πρότυπος αυτός πληθυσμός μπορεί να έχει την ηλικιακή κατανομή ενός από τους συγκρινόμενους πληθυσμούς (π.χ. της Σουηδίας ή του Παναμά), μπορεί να έχει την ηλικιακή κατανομή του πληθυσμού που προκύπτει από το συνδυασμό των ηλικιακών κατανομών των συγκρινόμενων πληθυσμών ή μπορεί να έχει την ηλικιακή κατανομή ενός

οποιοδήποτε αυθαίρετα επιλεγόμενου πληθυσμού. Η επιλογή του πρότυπου πληθυσμού εξαρτάται από το σκοπό της μελέτης. Π.χ., όταν συγκρίνονται εκτεθειμένοι σ' ένα συγκεκριμένο προσδιοριστή σε σχέση με μη εκτεθειμένους, τότε το πλέον λογικό είναι ο πρότυπος πληθυσμός να χρησιμοποιήσει την ηλικιακή κατανομή των εκτεθειμένων, επειδή η μελετώμενη έκβαση συμβαίνει συχνότερα στην ομάδα των εκτεθειμένων.

Η προτυποποιημένη συχνότητα ορίζεται ως ο σταθμισμένος μέσος όρος των συχνοτήτων στις ειδικές κατηγορίες με «βάρη» που λαμβάνονται από τον επιλεγόμενο πρότυπο πληθυσμό. Αντικείμενο είναι η προτυποποίηση των γενικών θνησιμοτήτων στη Σουηδία και τον Παναμά και ο πρότυπος πληθυσμός που επιλέγεται έχει τα ηλικιακά «βάρη» που φαίνονται στον πίνακα 2.

Πολλαπλασιάζεται καθένα από τα ηλικιακά «βάρη» που φαίνονται στον πίνακα 2 με τη θνησιμότητα κάθε χώρας για την αντίστοιχη ηλικιακή ομάδα. Έπειτα, προστίθενται τα επιμέρους γινόμενα, οπότε προκύπτει η προτυποποιημένη γενική συχνότητα ή θνησιμότητα στη συγκεκριμένη περίπτωση. Έτσι, στην περίπτωση της Σουηδίας προκύπτουν τα εξής:

Ηλικία (έτη)	Θνησιμότητα [(έτη) ⁻¹]	«Βάρη» (ποσοστό)	Γινόμενο [(έτη) ⁻¹]
0-29	0,0011	0,35	0,00039
30-59	0,0036	0,35	0,00126
>60	0,046	0,30	0,0138

Επομένως, η προτυποποιημένη, ως προς την ηλικία, γενική θνησιμότητα στη Σουηδία, χρησιμοποιώντας τον πρότυπο πληθυσμό Α, είναι (0,00039+0,00126+0,0138)=0,015 (έτη)⁻¹. Η θνησιμότητα αυτή είναι η γενική προτυποποιημένη θνησιμότητα που θα είχε η Σουηδία, εάν η ηλικιακή κατανομή των Σουηδών μετατρεπόταν από εκείνη που πραγματικά ήταν το 2000 στην υποθετική ηλικιακή κατανομή του πρότυπου πληθυσμού Α. Η προτυποποιημένη γενική θνησιμότητα είναι μεγαλύτερη από την αβρή γενική θνησιμότητα, γιατί στον πρότυπο πληθυσμό η βαρύτητα (0,30) που δόθηκε στα άτομα >60 ετών ήταν μεγαλύτερη από τη βαρύτητα (0,17) που είχε η ίδια ηλικιακή ομάδα στον πραγματικό πληθυσμό της

Πίνακας 2. Τα «βάρη» που αντιστοιχούν στις διάφορες ηλικιακές ομάδες ενός πρότυπου πληθυσμού Α.

Ηλικία (έτη)	«Βάρη» (ποσοστό)
0-29	0,35
30-59	0,35
>60	0,30

Σουηδίας. Η προτυποποιημένη γενική θνησιμότητα είναι η υποθετική τιμή της αβρής γενικής θνησιμότητας, εάν η ηλικιακή κατανομή του πραγματικού πληθυσμού ήταν ίδια με την ηλικιακή κατανομή του πρότυπου πληθυσμού Α.

Στην περίπτωση του Παναμά, χρησιμοποιώντας τα «βάρη» του πρότυπου πληθυσμού Α προκύπτουν τα εξής:

Ηλικία (έτη)	Θνησιμότητα [(έτη) ⁻¹]	«Βάρη» (ποσοστό)	Γινόμενο [(έτη) ⁻¹]
0-29	0,0053	0,35	0,00186
30-59	0,0028	0,35	0,001
>60	0,0416	0,30	0,01248

Έτσι, η προτυποποιημένη, ως προς την ηλικία, γενική θνησιμότητα στον Παναμά, χρησιμοποιώντας τον πρότυπο πληθυσμό Α, είναι (0,00186+0,001+0,01248)=0,01534 (έτη)⁻¹.

Η διαφορά των προτυποποιημένων γενικών θνησιμοτήτων στις δύο χώρες είναι 0,01534-0,015=0,00034 (έτη)⁻¹, γεγονός που υποδηλώνει ότι η προτυποποιημένη γενική θνησιμότητα στον Παναμά είναι μεγαλύτερη από την προτυποποιημένη γενική θνησιμότητα στη Σουηδία. Επιπλέον, ο λόγος των προτυποποιημένων γενικών θνησιμοτήτων είναι $\frac{0,01534}{0,015} = 1,02$, γεγονός που υποδηλώνει ότι η προτυποποιημένη γενική θνησιμότητα στον Παναμά είναι 2% μεγαλύτερη από την αντίστοιχη στη Σουηδία. Το συμπέρασμα που προκύπτει από τη χρήση των προτυποποιημένων γενικών θνησιμοτήτων, έπειτα από την εξουδετέρωση της συγχυτικής επίδρασης της ηλικίας, είναι τελείως διαφορετικό από το συμπέρασμα που προκύπτει από τη χρήση των αβρών γενικών θνησιμοτήτων, γεγονός που φανερώνει ότι η εξαγωγή ασφαλών συμπερασμάτων απαιτεί την εξουδετέρωση όσο το δυνατόν περισσότερων συγχυτών.

Πρέπει, πάντως, να σημειωθεί ότι η επιλογή του πρότυπου πληθυσμού επηρεάζει ως ένα βαθμό τη σύγκριση. Σε καμία περίπτωση, όμως, δεν είναι δυνατόν να οδηγήσει σε εκ διαμέτρου αντίθετα συμπεράσματα. Εάν, για παράδειγμα, για τον υπολογισμό των προτυποποιημένων γενικών θνησιμοτήτων στη Σουηδία και τον Παναμά χρησιμοποιηθεί ως πρότυπος ο πληθυσμός του Παναμά, τότε τα βάρη στις διάφορες ηλικιακές ομάδες είναι τα εξής:

Ηλικία (έτη)	«Βάρη» (ποσοστό)
0-29	0,69
30-59	0,26
>60	0,05

Έτσι, χρησιμοποιώντας ως πρότυπο τον πληθυσμό του Παναμά, η προτυποποιημένη γενική θνησιμότητα στη Σουηδία είναι $0,004$ (έτη)⁻¹, ενώ στον Παναμά είναι $0,0066$ (έτη)⁻¹. Η προτυποποιημένη γενική θνησιμότητα στον Παναμά είναι ίδια με την αδρή γενική θνησιμότητα, κάτι αναμενόμενο αφού χρησιμοποιήθηκε ως πρότυπος ο πληθυσμός του Παναμά. Η διαφορά των προτυποποιημένων γενικών θνησιμοτήτων είναι πλέον μεγαλύτερη [$(0,0066-0,004=0,0026$ (έτη)⁻¹]. Επιπλέον, και ο λόγος των προτυπωμένων θνησιμοτήτων είναι μεγαλύτερος $\left(\frac{0,0066}{0,004} = 1,65\right)$, γεγονός που υποδηλώνει ότι η προτυποποιημένη γενική θνησιμότητα στον Παναμά είναι 65% μεγαλύτερη από την αντίστοιχη στη Σουηδία.

Ο πρότυπος πληθυσμός Α (πίν. 2) που χρησιμοποιήθηκε στην πρώτη περίπτωση έδινε παρόμοια βαρύτητα (0,35, 0,35 και 0,30) και στις τρεις ηλικιακές ομάδες, ενώ όταν χρησιμοποιήθηκε ως πρότυπος ο πληθυσμός του Παναμά, αυτός έδινε πολύ μεγαλύτερη βαρύτητα (0,69) στα άτομα 0-29 ετών και πολύ μικρή βαρύτητα (0,05) στα άτομα >60 ετών. Και ακριβώς επειδή η ειδική θνησιμότητα των Σουηδών <30 ετών είναι πολύ μικρότερη από την ειδική θνησιμότητα των Παναμέζων <30 ετών, η διαφορά αυτή αντικατοπτρίζεται στη σύγκριση των προτυποποιημένων γενικών θνησιμοτήτων όταν χρησιμοποιούνται διαφορετικοί πρότυποι πληθυσμοί.

Όπως προαναφέρθηκε, όταν ο πληθυσμός του Παναμά χρησιμοποιείται ως πρότυπος πληθυσμός, τότε η προτυποποιημένη γενική θνησιμότητα του Παναμά είναι ίση με την αδρή γενική θνησιμότητα. Όπως η προτυποποιημένη γενική θνησιμότητα, έτσι και η αδρή γενική θνησιμότητα είναι ένας σταθμισμένος μέσος όρος των ειδικών θνησιμοτήτων στις διάφορες κατηγορίες με «βάρη» που προέρχονται από την ηλικιακή κατανομή του πραγματικού πληθυσμού.

Η *προτυποποιημένη συχνότητα* (standardized rate, SR) προκύπτει από την παρακάτω ισότητα:⁵

$$SR = \frac{\sum w_i R_i}{\sum w_i} \quad (1)$$

Με R_i συμβολίζεται η ειδική συχνότητα στην κατηγορία i , ενώ με w_i συμβολίζεται το «βάρος» για την κατηγορία i που προκύπτει από τον πρότυπο πληθυσμό. Διαιρώντας με $\sum w_i$, μετατρέπεται ένα οποιοδήποτε σύνολο «βαρών» (w_i) σ' ένα σύνολο ποσοστών που προστιθέμενα πλησιάζουν στη μονάδα. Εάν τα «βάρη» του πρότυπου πληθυσμού επιλεγούν με τέτοιο τρόπο, ώστε το άθροισμά τους να είναι η μονάδα, τότε η ισότητα 1 δεν

είναι απαραίτητο να περιλαμβάνει το $\sum w_i$ ως παρονομαστή, επειδή το $\sum w_i$ στην περίπτωση αυτή θα ισούται με τη μονάδα. Το R_i προκύπτει από τη διαίρεση του αριθμού των περιπτώσεων (y_i) με τον πληθυσμο-χρόνο παρακολούθησης (n_i), οπότε η ισότητα 1 μετατρέπεται ως εξής:

$$SR = \frac{\sum w_i \frac{y_i}{n_i}}{\sum w_i} \quad (2)$$

Εάν ο πραγματικός πληθυσμός χρησιμοποιηθεί ως πρότυπος πληθυσμός, όπως συνέβη στην περίπτωση του Παναμά παραπάνω, τότε ο πρότυπος πληθυσμός, ο οποίος είναι το σύνολο των «βαρών» (w_i), είναι παρόμοιος με την κατανομή του πληθυσμο-χρόνου παρακολούθησης του πραγματικού πληθυσμού, που είναι το σύνολο των παρονομαστών (n_i). Εάν $w_i = n_i$ για όλες τις κατηγορίες i , τότε η ισότητα 2 γίνεται ως εξής:

$$SR = \frac{\sum w_i \frac{y_i}{n_i}}{\sum n_i} = \frac{\sum y_i}{\sum n_i} = \text{αδρή συχνότητα} \quad (3)$$

Έτσι, επιβεβαιώνεται η εγκυρότητα της άποψης που διατυπώθηκε παραπάνω, ότι η προτυποποιημένη συχνότητα μπορεί να ερμηνευτεί ως η υποθετική αδρή συχνότητα εάν ο πραγματικός πληθυσμός είχε την κατανομή του πρότυπου πληθυσμού. Εάν ο πραγματικός πληθυσμός χρησιμοποιηθεί ως πρότυπος, τότε η υποθετική αδρή συχνότητα είναι η πραγματική αδρή συχνότητα.

5. «ΑΜΕΣΗ» ΚΑΙ «ΕΜΜΕΣΗ» ΠΡΟΤΥΠΟΠΟΙΗΣΗ

Η μέθοδος της προτυποποίησης που αναφέρθηκε εκτενώς παραπάνω καλείται συχνά «άμεση» (direct) προτυποποίηση, με σκοπό να διαχωριστεί από εναλλακτικές μεθόδους προτυποποίησης, η πλέον γνωστή από τις οποίες καλείται «έμμεση» (indirect) προτυποποίηση.^{3,12,22,23,25,28,29} Η «άμεση» προτυποποίηση είναι μια μέθοδος που σταθμίζει μια ομάδα παρατηρηθεισών ειδικών συχνοτήτων στις διάφορες κατηγορίες σύμφωνα με τα «βάρη» ενός πρότυπου πληθυσμού. Η «έμμεση» προτυποποίηση υποτίθεται ότι είναι μια διαφορετική μέθοδος, όπου ο πρότυπος πληθυσμός δεν παρέχει τα «βάρη» με τα οποία γίνεται η στάθμιση, αλλά παρέχει ένα σύνολο πρότυπων συχνοτήτων, που στη συνέχεια σταθμίζονται με τους πληθυσμο-χρόνους παρακολούθησης στις διάφορες κατηγορίες του μελετώμενου πληθυσμού. Όπως θα γίνει ευκολότερα κατανοητό στη συνέχεια, η «άμεση» προτυποποίηση δεν είναι διαφορετική μέθοδος από

την «έμμεση», αλλά απλά χρησιμοποιούνται οι όροι αυτοί για να διακρίνονται ευκολότερα οι δύο αυτές διαφορετικές διαδικασίες. Η «έμμεση» προτυποποίηση δεν μπορεί να χαρακτηρίσει από μόνη της την εμπειρία ενός νοσήματος στο μελετώμενο πληθυσμό, επειδή, έως τώρα, δεν παρέχει πληροφορίες για τη συχνότητα εμφάνισης του νοσήματος στο μελετώμενο πληθυσμό. Έτσι, η «έμμεση» προτυποποίηση χρησιμοποιείται για τον υπολογισμό της «αναμενόμενης» (expected) προτυποποιημένης συχνότητας ή της αναμενόμενης τιμής για την αδρή συχνότητα ή του αναμενόμενου αριθμού περιπτώσεων (expected number of cases) στο μελετώμενο πληθυσμό. Η σύγκριση μεταξύ του μελετώμενου και του πρότυπου πληθυσμού πραγματοποιείται με το *προτυπωμένο πηλίκο νοσηρότητας* (standardized morbidity ratio, SMR)^{12,22,25,28,29} ή *προτυπωμένο πηλίκο θνησιμότητας* (standardized mortality ratio, SMR). Το SMR προκύπτει από τη διαίρεση του παρατηρούμενου αριθμού των περιπτώσεων (observed number of cases) που συμβαίνουν στο μελετώμενο πληθυσμό προς τον αναμενόμενο αριθμό των περιπτώσεων που υπολογίζονται έπειτα από τη χρήση των ειδικών συχνοτήτων στις διάφορες κατηγορίες του πρότυπου πληθυσμού.

Για λόγους ευκολίας, ο μελετώμενος πληθυσμός στο εξής θα αναφέρεται ως εκτεθειμένος πληθυσμός. Ο παρατηρούμενος αριθμός περιπτώσεων είναι ο συνολικός αριθμός περιπτώσεων που συμβαίνουν στον εκτεθειμένο πληθυσμό. Ο αναμενόμενος αριθμός περιπτώσεων υπολογίζεται από τον πολλαπλασιασμό των ειδικών συχνοτήτων σε κάθε κατηγορία του μη εκτεθειμένου πληθυσμού με τους πληθυσμο-χρόνους παρακολούθησης σε κάθε κατηγορία του εκτεθειμένου πληθυσμού. Στον εκτεθειμένο πληθυσμό, η συχνότητα στην κατηγορία i συμβολίζεται με R_{1i} , ο αριθμός των περιπτώσεων (αριθμητής) συμβολίζεται με y_{1i} και ο πληθυσμο-χρόνος παρακολούθησης (παρονομαστής) συμβολίζεται με n_{1i} . Στο μη εκτεθειμένο πληθυσμό, η συχνότητα στην κατηγορία i συμβολίζεται με R_{0i} , ο αριθμός των περιπτώσεων (αριθμητής) συμβολίζεται με y_{0i} και ο πληθυσμο-χρόνος παρακολούθησης (παρονομαστής) συμβολίζεται με n_{0i} . Έτσι, το SMR υπολογίζεται ως εξής:^{12,30}

$$SMR = \frac{\sum y_{1i}}{\sum n_{1i} \frac{y_{0i}}{n_{0i}}} = \frac{\text{παρατηρούμενος αριθμός περιπτώσεων}}{\text{αναμενόμενος αριθμός περιπτώσεων}} \quad (4)$$

Η ισότητα 4 είναι ισοδύναμη με το λόγο της αδρής συχνότητας στον εκτεθειμένο πληθυσμό προς την «έμ-

μεση» προτυποποιημένη συχνότητα. Η «έμμεση» προτυποποίηση σταθμίζει τις ειδικές συχνότητες στις διάφορες κατηγορίες του μη εκτεθειμένου πληθυσμού με τα «βάρη» που προέρχονται από τις αντίστοιχες κατηγορίες του εκτεθειμένου πληθυσμού:¹²

$$SMR = \frac{\sum y_{1i}}{\sum n_{1i} \frac{y_{0i}}{n_{0i}}} \quad (5)$$

Η ισότητα 5 προκύπτει από την ισότητα 4, αν διαιρεθεί αριθμητής και παρονομαστής με το $\sum n_{1i}$. Έτσι, το SMR είναι ο λόγος δύο προτυπωμένων συχνοτήτων που έχουν προτυπωθεί ως προς την κατανομή του εκτεθειμένου πληθυσμού. Ο αριθμητής του SMR είναι η αδρή συχνότητα του εκτεθειμένου πληθυσμού, ενώ ο παρονομαστής είναι ο σταθμισμένος μέσος όρος των ειδικών συχνοτήτων στις διάφορες κατηγορίες του μη εκτεθειμένου πληθυσμού. Οι ειδικές αυτές συχνότητες σταθμίζονται με βάση την κατανομή του εκτεθειμένου πληθυσμού. Με τον τρόπο αυτόν, οι ομάδες των ειδικών συχνοτήτων στις διάφορες κατηγορίες, τόσο στον εκτεθειμένο όσο και στο μη εκτεθειμένο πληθυσμό, σταθμίζονται με βάση την κατανομή του εκτεθειμένου πληθυσμού.

Η «έμμεση» προτυποποίηση που περιγράφηκε παραπάνω δεν διαφέρει, ουσιαστικά, από την «άμεση». Και στην «έμμεση» η διαδικασία περιλαμβάνει τον υπολογισμό ενός σταθμισμένου μέσου όρου των ειδικών συχνοτήτων στις διάφορες κατηγορίες. Στην περίπτωση του SMR, τα «βάρη» για τη στάθμιση των συχνοτήτων προκύπτουν πάντοτε από τον εκτεθειμένο πληθυσμό, τον πληθυσμό δηλαδή από τον οποίο προκύπτουν οι παρατηρηθείσες περιπτώσεις. Έτσι, ο πρότυπος πληθυσμός είναι πάντοτε ο εκτεθειμένος πληθυσμός.

Το συχνότερο πρόβλημα στην «έμμεση» προτυποποίηση είναι ότι ορισμένοι θεωρούν το μη εκτεθειμένο πληθυσμό ως πρότυπο πληθυσμό. Έτσι, προβαίνουν στη σύγκριση δύο SMR, τα οποία προκύπτουν από τη χρήση δύο διαφορετικών πληθυσμών ως προτύπων. Είναι προφανές ότι η σύγκριση αυτή των δύο SMR οδηγεί σε εσφαλμένα συμπεράσματα. Σ' ένα εργοστάσιο, οι εργάτες ταξινομούνται σε δύο διαφορετικούς εκτεθειμένους πληθυσμούς ανάλογα με την έκθεσή τους σ' έναν τοξικό παράγοντα κινδύνου. Για τους δύο εκτεθειμένους πληθυσμούς των εργατών υπολογίζονται δύο διαφορετικά SMR. Και για τα δύο SMR, ως μη εκτεθειμένος

πληθυσμός χρησιμοποιείται ο γενικός πληθυσμός. Τα δύο SMR, ωστόσο, έχουν διαφορετικούς πρότυπους πληθυσμούς, καθώς έχουν διαφορετικούς εκτεθειμένους πληθυσμούς και, όπως προαναφέρθηκε, ο πρότυπος πληθυσμός είναι πάντοτε ο εκτεθειμένος πληθυσμός. Ακριβώς επειδή τα δύο SMR χρησιμοποιούν διαφορετικούς πρότυπους πληθυσμούς είναι εσφαλμένη η σύγκριση μεταξύ τους και, επομένως, δεν μπορούν να χρησιμοποιηθούν για την εκτίμηση του σχετικού αποτελέσματος των δύο διαφορετικών εκθέσεων στον τοξικό παράγοντα. Η εκτίμηση αυτή μπορεί να πραγματοποιηθεί εύκολα αν και τα δύο SMR υπολογιστούν με βάση τον ίδιο πρότυπο πληθυσμό.

Στον πίνακα 3 φαίνονται οι υποθετικές θνησιμότητες σε δύο εκτεθειμένους πληθυσμούς Α και Β και στο γενικό πληθυσμό ανάλογα με την ηλικία, η οποία χωρίστηκε σε δύο ομάδες, νέους και ηλικιωμένους. Οι θνησιμότητες στους νέους και τους ηλικιωμένους είναι οι ίδιες στους εκτεθειμένους πληθυσμούς Α και Β. Στους νέους, η θνησιμότητα στους εκτεθειμένους $[0,005 (\text{έτη})^{-1}]$ είναι δέκα φορές μεγαλύτερη από τη θνησιμότητα $[0,0005 (\text{έτη})^{-1}]$ στο γενικό πληθυσμό (μη εκτεθειμένος πληθυσμός), ενώ στους ηλικιωμένους, η θνησιμότητα $[0,004 (\text{έτη})^{-1}]$ στους εκτεθειμένους πληθυσμούς είναι διπλάσια από τη θνησιμότητα στο γενικό πληθυσμό $[0,002 (\text{έτη})^{-1}]$.

Αρχικά, σταθμίζονται όλες οι θνησιμότητες χρησιμοποιώντας τα «βάρη» που προκύπτουν από το γενικό πληθυσμό, χρησιμοποιώντας δηλαδή ως πρότυπο πληθυσμό το γενικό πληθυσμό. Έτσι, για το γενικό πληθυσμό η προτυποποιημένη γενική θνησιμότητα ισούται με την αδρή γενική θνησιμότητα, καθώς ο γενικός πληθυσμός χρησιμοποιήθηκε ως πρότυπος πληθυσμός, οπότε η προτυποποιημένη γενική θνησιμότητα του γενικού πληθυσμού

$$\text{ισούται με } \frac{50 + 400}{(100.000 + 200.000) \text{ έτη}} = 0,0015 (\text{έτη})^{-1}.$$

Για τον εκτεθειμένο πληθυσμό Α, η προτυποποιημένη γενική θνησιμότητα ισούται με $\frac{1}{3} [0,005 (\text{έτη})^{-1}] + \frac{2}{3} [0,004 (\text{έτη})^{-1}] = 0,0043 (\text{έτη})^{-1}$. Για τον εκτεθειμένο πληθυσμό Β, η προτυποποιημένη γενική θνησιμότητα ισούται με $\frac{1}{3} [0,005 (\text{έτη})^{-1}] + \frac{2}{3} [0,004 (\text{έτη})^{-1}] = 0,0043$.

Ο λόγος της προτυποποιημένης γενικής θνησιμότητας του εκτεθειμένου πληθυσμού Α προς την προτυποποιημένη γενική θνησιμότητα του γενικού πληθυσμού καλείται προτυποποιημένος λόγος θνησιμοτήτων ή, γενικότερα, προτυποποιημένος λόγος συχνοτήτων (standardized rate ratio, SRR) και ισούται με:

$$SRR_1 = \frac{0,0043 (\text{έτη})^{-1}}{0,0015 (\text{έτη})^{-1}} = 2,87$$

Παρομοίως, ο λόγος της προτυποποιημένης γενικής θνησιμότητας στον εκτεθειμένο πληθυσμό Β προς την προτυποποιημένη γενική θνησιμότητα στο γενικό πληθυσμό ισούται με:

$$SRR_2 = \frac{0,0043 (\text{έτη})^{-1}}{0,0015 (\text{έτη})^{-1}} = 2,87$$

Έτσι, $SRR_1 = SRR_2$. Το αποτέλεσμα είναι αναμενόμενο, αν αναλογιστεί κάποιος ότι κατά τον υπολογισμό των SRR_1 και SRR_2 χρησιμοποιήθηκε ο ίδιος πρότυπος πληθυσμός, που ήταν ο γενικός πληθυσμός, καθώς επίσης και το ότι οι θνησιμότητες στις δύο ηλικιακές ομάδες ήταν οι ίδιες στους εκτεθειμένους πληθυσμούς Α και Β.

Στον εκτεθειμένο πληθυσμό Α, το SMR υπολογίζεται ως εξής:

$$SMR_A = \frac{\text{παρατηρούμενος αριθμός περιπτώσεων}}{\text{αναμενόμενος αριθμός περιπτώσεων}} =$$

Πίνακας 3. Υποθετικές θνησιμότητες ανάλογα με την ηλικία για δύο εκτεθειμένους πληθυσμούς Α και Β και το γενικό πληθυσμό.

Ηλικιακή ομάδα	Γενικός πληθυσμός	Εκτεθειμένος πληθυσμός Α	Εκτεθειμένος πληθυσμός Β
<i>Νέοι</i>			
Περιπτώσεις (θάνατοι)	50	50	5
Πληθυσμο-χρόνος (ανθρωπο-έτη)	100.000	10.000	1000
Θνησιμότητα $[(\text{έτη})^{-1}]$	0,0005	0,005	0,005
<i>Ηλικιωμένοι</i>			
Περιπτώσεις (θάνατοι)	400	4	40
Πληθυσμο-χρόνος (ανθρωπο-έτη)	200.000	1000	10.000
Θνησιμότητα $[(\text{έτη})^{-1}]$	0,002	0,004	0,004

$$= \frac{50+4}{\frac{50}{100.000} (10.000) + \frac{400}{200.000} (10.000)}$$

$$SMR_1 = \frac{54}{5+2} = \frac{54}{7} = 7,71$$

Στον εκτεθειμένο πληθυσμό Β, το SMR υπολογίζεται ως εξής:

$$SMR_2 = \frac{\text{παρατηρούμενος αριθμός περιπτώσεων}}{\text{αναμενόμενος αριθμός περιπτώσεων}} =$$

$$= \frac{5+40}{\frac{50}{100.000} (1000) + \frac{400}{200.000} (10.000)}$$

$$SMR_2 = \frac{45}{0,5+20} = \frac{45}{20,5} = 2,20$$

Η σημαντική διαφορά στις τιμές των δύο SMR δεν οφείλεται στο ότι το αποτέλεσμα της έκθεσης Α είναι μεγαλύτερο από το αποτέλεσμα της έκθεσης Β, καθώς τα αποτελέσματα αυτά (με βάση τα SRR_1 και SRR_2) είναι ισοδύναμα, αλλά στο ότι ο εκτεθειμένος πληθυσμός Α είναι νεότερος, ενώ ο εκτεθειμένος πληθυσμός Β είναι περισσότερο ηλικιωμένος. Μεταξύ των νέων ατόμων το σχετικό αποτέλεσμα είναι πολύ μεγαλύτερο και γι' αυτό το SMR_1 είναι πολύ μεγαλύτερο από το SMR_2 . Ο μόνος λόγος στον οποίο οφείλεται η διαφορά μεταξύ των δύο SMR είναι η διαφορετική ηλικιακή κατανομή των δύο εκτεθειμένων πληθυσμών Α και Β. Στην περίπτωση των SMR, παρότι πραγματοποιήθηκε προτυποποίηση για την εξουδετέρωση της συγχυτικής επίδρασης της ηλικίας, δεν ήταν δυνατή η σύγκρισή τους, καθώς το καθένα από αυτά χρησιμοποίησε διαφορετικό πρότυπο πληθυσμό. Έτσι, για τον υπολογισμό του SMR_1 χρησιμοποιήθηκε ο εκτεθειμένος πληθυσμός Α ως πρότυπος πληθυσμός, ενώ για τον υπολογισμό του SMR_2 χρησιμοποιήθηκε ο εκτεθειμένος πληθυσμός Β ως πρότυπος πληθυσμός. Το παράδειγμα αυτό είναι ενδεικτικό του λάθους που κάνουν ορισμένοι στην περίπτωση της «έμμεσης» προτυποποίησης, ως προς το ποιος πληθυσμός είναι ο πρότυπος. Ορισμένοι, λανθασμένα, θα πίστευαν ότι ο πρότυπος πληθυσμός και των δύο SMR είναι ο γενικός πληθυσμός, οπότε είναι δυνατή η σύγκριση των δύο SMR μεταξύ τους. Όπως προαναφέρθηκε, όμως, δεν ισχύει κάτι τέτοιο. Σε κάθε περίπτωση, πάντως, ο υπολογισμός των SMR δεν αποτελεί μια ξεχωριστή μέθοδο προτυποποίησης. Πρόκειται για την ίδια διαδικασία με την «άμεση» προτυποποίηση, αλλά ο εκτεθειμένος πληθυσμός κατά τον υπολογισμό του SMR χρησιμοποιείται ως πρότυπος πληθυσμός. Συγκρίνοντας δύο

SMR ισοδυναμεί με σύγκριση μέτρων που έχουν διαφορετικά πρότυπα και επομένως είναι μη έγκυρη, ακόμη και αν χρησιμοποιείται ο ίδιος μη εκτεθειμένος πληθυσμός. Για το λόγο αυτόν απαιτείται να υπολογίζονται τα SRR και να συγκρίνονται μεταξύ τους. Συμπερασματικά, ο υπολογισμός και η σύγκριση των SMR απαιτεί ιδιαίτερη προσοχή όσον αφορά στην επιλογή του πρότυπου πληθυσμού, έτσι ώστε τα αποτελέσματα να είναι έγκυρα.

6. ΠΡΟΤΥΠΟΠΟΙΗΣΗ ΕΝΑΝΤΙ ΔΙΑΣΤΡΩΜΑΤΩΣΗΣ

Τόσο η προτυποποίηση όσο και η *διαστρωμάτωση* (stratification)^{5,10-12,31,32} χρησιμοποιούνται για την εξουδετέρωση των συγχυτών μιας μελέτης.^{2,5,12} Πρόκειται για δύο διαφορετικές μεθόδους που μπορεί να οδηγήσουν σε διαφορετικά αποτελέσματα και γι' αυτό συχνά τίθεται το ερώτημα ποια από τις δύο είναι προτιμότερη. Και οι δύο περιλαμβάνουν τη σύγκριση ενός σταθμισμένου μέσου όρου, που προκύπτει από τις ειδικές συχνότητες στις διάφορες κατηγορίες. Στη διαστρωμάτωση, τα «βάρη» που αντιστοιχούν στα διάφορα στρώματα δεν γίνονται άμεσα αντιληπτά, καθώς είναι «κρυμμένα» στους τύπους των Mantel-Haenszel. Τα «βάρη» αντανakλούν την πληροφορία που περιέχουν τα δεδομένα των διαφόρων στρωμάτων. Έτσι, στους τύπους των Mantel-Haenszel τα «βάρη» είναι μεγάλα για τα στρώματα που προσφέρουν αρκετή πληροφορία για τα στρώματα, δηλαδή που έχουν μεγάλο αριθμό δεδομένων, ενώ τα «βάρη» είναι μικρά για τα στρώματα που προσφέρουν λιγότερη πληροφορία, τα στρώματα, δηλαδή, που έχουν μικρό αριθμό δεδομένων. Επειδή τα «βάρη» αντανakλούν την ποσότητα της πληροφορίας σε κάθε στρώμα, το αποτέλεσμα της συνένωσης, κατά τη διαστρωμάτωση, είναι μια συνολική εκτίμηση, κάτι επιθυμητό από την πλευρά της στατιστικής αποδοτικότητας. Στην προτυποποίηση, επίσης, προσδιορίζεται ένα ορισμένο «βάρος» σε κάθε στρώμα και υπολογίζεται ένας σταθμισμένος μέσος όρος των αποτελεσμάτων στα διάφορα στρώματα. Στην προτυποποίηση, εξάλλου, σε αντίθεση με τη διαστρωμάτωση, τα «βάρη» δεν σχετίζονται με τον αριθμό των δεδομένων που περιέχουν τα στρώματα. Έτσι, στη διαστρωμάτωση τα «βάρη» καθορίζονται, ουσιαστικά, από τα δεδομένα αυτά καθαυτά, ενώ στην προτυποποίηση τα «βάρη» δεν σχετίζονται με τα δεδομένα, αλλά απλά αντικατοπτρίζουν την κατανομή ενός πρότυπου πληθυσμού, ο οποίος είτε είναι ένας συγκεκριμένος πληθυσμός είτε επιλέγεται αυθαίρετα.^{5,12}

Επιπλέον, η προτυποποίηση διαφέρει από τη διαστρωμάτωση στο ότι η συνένωση που πραγματοποιείται στη

διαστρωμάτωση είναι εφικτή μόνον όταν το αποτέλεσμα της έκθεσης είναι το ίδιο σε όλα τα στρώματα.^{5,12} Η συνένωση, δηλαδή, είναι δυνατή μόνον όταν ισχύει η υπόθεση της ομοιογένειας του αποτελέσματος, οπότε εξάγονται και οι αντίστοιχοι τύποι των Mantel-Haenszel. Ουσιαστικά, η συνένωση είναι εφικτή μόνο εάν φαίνεται απίθανο, είτε από τα δεδομένα μιας συγκεκριμένης μελέτης είτε από τα δεδομένα άλλων μελετών, ο τυχόν συγχυτής να είναι ταυτόχρονα και τροποποιητής του μέτρου σχέσης. Η υπόθεση της ομοιογένειας, πάντως, δεν υποδηλώνει ότι το αποτέλεσμα θα πρέπει να είναι το ίδιο, ή σχεδόν το ίδιο, στα διάφορα στρώματα, αλλά λαμβάνει υπόψη της τη στατιστική μεταβλητότητα στα διάφορα στρώματα. Στην προκειμένη περίπτωση, είναι δυνατόν να διεξαχθεί ο κατάλληλος 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 το ότι το αποτέλεσμα είναι ομοιόμορφο. Πρέπει, πάντως, να σημειωθεί ότι η υπόθεση ότι το αποτέλεσμα είναι ομοιόμορφο είναι πιθανότητα λάθος στις περισσότερες περιπτώσεις. Είναι περισσότερο ρεαλιστικό να θεωρηθεί η υπόθεση αυτή ως μια «διευκόλυνση» που επιτρέπει τον υπολογισμό ενός «συγκεντρωτικού» μέτρου σχέσης. Εκτός και αν τα δεδομένα μιας μελέτης υποδηλώνουν ξεκάθαρα μια μεταβλητότητα που κλονίζει την υπόθεση της ομοιομορφίας του αποτελέσματος στα διάφορα στρώματα, είναι συνήθως λογικό να χρησιμοποιούνται οι τύποι των Mantel-Haenszel. Η προτυποποίηση, ωστόσο, σε αντίθεση με τη διαστρωμάτωση, μπορεί να εφαρμοστεί ακόμη και αν το αποτέλεσμα είναι εντυπωσιακά διαφορετικό στα διάφορα στρώματα. Έτσι, η εφαρμογή της προτυποποίησης επιτρέπει τον υπολογισμό μιας συγκεντρωτικής εκτίμησης του αποτελέσματος στα διάφορα στρώματα, καθώς δεν απαιτείται η ομοιομορφία του αποτελέσματος στα διάφορα στρώματα.

Η προτυποποίηση προτιμάται, έναντι της διαστρωμάτωσης, όχι μόνο γιατί μπορεί να εφαρμοστεί ακόμη και όταν το αποτέλεσμα της έκθεσης είναι σημαντικά διαφορετικό στα διάφορα στρώματα, αλλά και γιατί χρησιμοποιείται ένα καθορισμένο σύνολο «βαρών», με σκοπό να συνδυαστούν τα αποτελέσματα στα διάφορα στρώματα. Το γεγονός αυτό επιτρέπει την καλύτερη σύγκριση των προτυπωμένων αποτελεσμάτων τόσο στην ίδια μελέτη όσο και μεταξύ διαφορετικών μελετών.

Τα δεδομένα μιας μελέτης με κλειστούς* πληθυσμούς, στην οποία υπολογίζονται επιπτώσεις-πυκνότητες, φαίνονται στον πίνακα 4. Είναι δυνατόν να υπολογιστεί ο «συγκεντρωτικός» (ή συνοπτικός) λόγος των επιπτώσεων-πυκνότητων (incidence-density ratio, IDR) ή η «συγ-

Πίνακας 4. Δεδομένα μιας 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

κεντρωτική» διαφορά των επιπτώσεων-πυκνότητων (incidence-density difference, IDD). Πιο συγκεκριμένα, η «συγκεντρωτική» διαφορά των επιπτώσεων-πυκνότητων στα διαστρωματωμένα δεδομένα υπολογίζεται με την παρακάτω ισότητα των Mantel-Haenszel.^{5,12}

$$IDD_{MH} = \frac{\sum_i \frac{a_i PT_{0i} - b_i PT_{1i}}{T_i}}{\sum_i \frac{PT_{1i} PT_{0i}}{T_i}} \quad (6)$$

Ο «συγκεντρωτικός» λόγος των επιπτώσεων-πυκνότητων υπολογίζεται με την παρακάτω ισότητα των Mantel-Haenszel:^{5,12}

$$IDR_{MH} = \frac{\sum_i \frac{a_i PT_{0i}}{T_i}}{\sum_i \frac{b_i PT_{1i}}{T_i}} \quad (7)$$

Στον πίνακα 5 φαίνονται τα υποθετικά δεδομένα μιας μελέτης με κλειστούς πληθυσμούς, όπου υπολογίστηκαν οι επιπτώσεις-πυκνότητες θανάτου (ή θνησιμότητες) για παροντικούς και παρελθοντικούς λήπτες κλοζαπίνης, ενός φαρμάκου που χρησιμοποιείται για τη θεραπεία

* Κλειστός (closed) ή σταθερός (fixed) ή στατικός (static) πληθυσμός ή κοόρτη (cohort) είναι ένα «κλειστό» σύνολο ατόμων, όπου η ιδιότητα του μέλους καθορίζεται από ένα συμβάν (event) σε μια συγκεκριμένη τοποχρονική περιοχή.^{1-3,5,12,17,27} Ο χαρακτηρισμός «κλειστός» σημαίνει ότι «απαγορεύεται» η έξοδος των μελών από τον πληθυσμό αυτόν. Η ιδιότητα του μέλους ενός κλειστού πληθυσμού δεν χάνεται με το θάνατο του μέλους. Η κοόρτη των στρατεύσιμων του 1900 στην Αθήνα εξακολουθεί να κινείται στο χρόνο. Το ίδιο ισχύει και για τον κλειστό πληθυσμό των διδασκόντων, π.χ., του 1930 του Πανεπιστημίου της Αθήνας. Τα κριτήρια του μέλους του κλειστού πληθυσμού εκπληρώνονται σε ένα συγκεκριμένο χρονικό σημείο (χρόνος μηδέν, T_0), που μπορεί να είναι χρονολογικά διαφορετικός στα μέλη του πληθυσμού. Παραδείγματα κλειστού πληθυσμού είναι (α) τα παιδιά εγκύων που εντάχθηκαν σε μια συγκεκριμένη μελέτη, (β) οι στρατεύσιμοι μιας γεωγραφικής περιοχής σ' ένα ημερολογιακό έτος, (γ) τα άτομα που πέθαναν σε μια συγκεκριμένη περιοχή σ' ένα συγκεκριμένο χρονικό διάστημα και (δ) τα άτομα που εξετάστηκαν για διαγνωστικούς δείκτες ή παράγοντες κινδύνου (προσδιοριστές) για στεφανιαία νόσο σε ένα συγκεκριμένο πρόγραμμα διαλογής. Χρόνος T_0 των κλειστών πληθυσμών που αναφέρθηκαν είναι ο χρόνος γέννησης, στράτευσης, θανάτου ή διαλογής, αντίστοιχα. Η μέτρηση της συχνότητας ενός νοσήματος ή του θανάτου σε έναν κλειστό πληθυσμό γίνεται με δύο μέτρα συχνότητας, που είναι η επίπτωση-πυκνότητα και η επίπτωση-ποσοστό (incidence-proportion).

Πίνακας 5. Υποθετικές θνησιμότητες για παροντικούς και παρελθοντικούς λήπτες κλοζαπίνης, ανάλογα με την ηλικία τους και συνολικά.

	Ηλικία (έτη)				Σύνολο	
	10-54		55-94		Τωρινοί	Παρελθόντες
	Τωρινοί	Παρελθόντες	Τωρινοί	Παρελθόντες		
Θάνατοι	270	153	230	216	500	369
Πληθυσμο-χρόνος (ανθρωπο-έτη)	91.000	23.110	9000	4078	100.000	27.188
Θνησιμότητα [$\square 10^{-5}$ (έτη) ⁻¹]	297	662	2556	5297	500	1357
Διαφορά θνησιμοτήτων [$\square 10^{-5}$ (έτη) ⁻¹]	-365		-2.741		-857	
Λόγος θνησιμοτήτων	0,45		0,49		0,37	

ανθεκτικών μορφών σχιζοφρένειας. Οι ερευνητές υπέθεσαν ότι η λήψη κλοζαπίνης μειώνει τη θνησιμότητα στους παροντικούς λήπτες. Για να διερευνηθεί η υπόθεση αυτή, χρησιμοποιήθηκε 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% (9000/100.000) του πληθυσμο-χρόνου ανήκε στη μεγαλύτερη ηλικιακή κατηγορία, ενώ στους παρελθοντικούς λήπτες κλοζαπίνης το 15% (4078/27.188) 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 5, η θνησιμότητα ήταν μικρότερη στους παροντικούς λήπτες. Η αδρή διαφορά των θνησιμοτήτων ήταν $-0,00857$ (έτη)⁻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 6:

$$IDD_{MH} = \frac{\frac{270 \square 23.110 - 153 \square 91.000}{114.110} + \frac{230 \square 4078 - 216 \square 9000}{13.078}}{\frac{91.000 \square 23.110}{114.110} + \frac{9000 \square 4078}{13.078}} = -679 \square 10^{-5} \text{ (έτη)}^{-1}$$

Η τιμή που προκύπτει [$-0,00679$ (έτη)⁻¹], μετά από την εξουδετέρωση της συγχυτικής επίδρασης της ηλικίας, είναι αρκετά μικρότερη από την τιμή που προκύπτει με τη χρήση των αδρών δεδομένων [$-0,00857$ (έτη)⁻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.

Ο «συγκεντρωτικός» λόγος των θνησιμοτήτων για τα δεδομένα του πίνακα 5 υπολογίζεται με βάση την ισότητα 7:

$$IDD_{MH} = \frac{\frac{270 \square 23.110}{114.110} + \frac{230 \square 4078}{13.078}}{\frac{153 \square 91.000}{114.110} + \frac{216 \square 9000}{13.078}} = \frac{54,7+71,7}{122+148,6} = 0,47$$

Η τιμή 0,47 υποδηλώνει ότι, μετά από την εξουδετέρωση της συγχυτικής δράσης της ηλικίας στα δύο επιμέρους ηλικιακά στρώματα, η θνησιμότητα στους παρελθοντικούς λήπτες ήταν διπλάσια σε σχέση με τη θνησιμότητα στους παροντικούς λήπτες.

Στη συνέχεια χρησιμοποιείται η προτυποποίηση για την εξουδετέρωση της συγχυτικής επίδρασης της ηλικίας, οπότε υπολογίζεται η προτυποποιημένη διαφορά θνησιμοτήτων ή προτυποποιημένη διαφορά συχνοτήτων (standardized rate difference, SRD) και ο προτυποποιημένος λόγος θνησιμοτήτων ή, γενικότερα, ο προτυποποιημένος λόγος συχνοτήτων (standardized rate ratio, SRR). Ως πρότυπος επιλέγεται ο πληθυσμός των τωρινών χρηστών κλοζαπίνης, στον οποίο υπήρχαν συνολικά 100.000 ανθρωπο-έτη παρακολούθησης, από τα οποία τα 91.000 (91%) ανήκαν στην ηλικιακή ομάδα 10-54 έτη, ενώ τα 9000 (9%) ανήκαν στην ηλικιακή ομάδα 55-94 έτη. Για την προτυποποίηση της θνησιμότητας

των παρελθόντων χρηστών κλοζαπίνης με βάση την ηλικιακή κατανομή του επιλεγέντος πρότυπου πληθυσμού λαμβάνεται ο σταθμισμένος μέσος όρος των παρελθόντων χρηστών ως εξής:

$$0,91 \times \left[\frac{662}{100.000} (\text{έτη})^{-1} \right] + 0,09 \times \left[\frac{5297}{100.000} (\text{έτη})^{-1} \right] = \frac{1079}{100.000} (\text{έτη})^{-1}$$

Όπως προαναφέρθηκε, η προτυποποιημένη γενική θνησιμότητα των τωρινών χρηστών κλοζαπίνης ισούται με την αδρή γενική θνησιμότητα των τωρινών χρηστών, καθώς ο πληθυσμός των τωρινών χρηστών επιλέχθηκε ως πρότυπος πληθυσμός. Επομένως, η προτυποποιημένη γενική θνησιμότητα των τωρινών χρηστών κλοζαπίνης

$$\text{ισούται με } \frac{(270 + 230) \text{ θάνατοι}}{(91.000 + 9000) \text{ έτη}} = \frac{500}{100.000} (\text{έτη})^{-1}.$$

Έτσι, η προτυποποιημένη διαφορά γενικών θνησιμοτήτων είναι η διαφορά μεταξύ των προτυποποιημένων γενικών θνησιμοτήτων για τους τωρινούς και τους παρελθόντες χρήστες, η οποία ισούται με $\frac{(500 - 1079)}{100.000} (\text{έτη})^{-1} = -\frac{579}{100.000} (\text{έτη})^{-1}$.

Η προτυποποιημένη διαφορά γενικών θνησιμοτήτων $\left[-\frac{579}{100.000} (\text{έτη})^{-1} \right]$ που προέκυψε από την εφαρμογή της προτυποποίησης είναι, κατ' απόλυτη τιμή, λίγο μικρότερη από τη «συγκεντρωτική» διαφορά γενικών θνησιμοτήτων $\left[-\frac{679}{100.000} (\text{έτη})^{-1} \right]$ που προέκυψε από την εφαρμογή της διαστρωματικής ανάλυσης.

Ο προτυποποιημένος λόγος των γενικών θνησιμοτήτων, εξάλλου, προκύπτει από τη διαίρεση της προτυποποιημένης γενικής θνησιμότητας στους τωρινούς χρήστες με την προτυποποιημένη γενική θνησιμότητα στους παρελθόντες χρήστες και ισούται με $\frac{500}{1079} = 0,46$. Ουσιαστικά, ο προτυποποιημένος λόγος θνησιμοτήτων που προέκυψε από την εφαρμογή της προτυποποίησης είναι ίσος με το «συγκεντρωτικό» λόγο των θνησιμοτήτων

που προέκυψε από την εφαρμογή της διαστρωματικής ανάλυσης. Οι ειδικοί λόγοι θνησιμοτήτων στα δύο ηλικιακά στρώματα διαφέρουν πολύ λίγο μεταξύ τους και γι' αυτό είτε χρησιμοποιηθεί η προτυποποίηση είτε η διαστρωματική ανάλυση το αποτέλεσμα είναι το ίδιο.

7. ΣΥΝΟΨΗ

Η προτυποποίηση και η διαστρωμάτωση είναι δύο διαφορετικές μέθοδοι εξουδετέρωσης των συγχυτών μιας μελέτης. Και στις δύο μεθόδους υπολογίζεται ένας σταθμισμένος μέσος όρος των ειδικών αποτελεσμάτων στα διάφορα στρώματα. Η διαφορά έγκειται στον τρόπο προέλευσης των «βαρών» που χρησιμοποιούν. Στη διαστρωμάτωση, τα δεδομένα των διαφόρων στρωμάτων είναι εκείνα που καθορίζουν τα «βάρη», τα οποία παράγονται μαθηματικά, με αποτέλεσμα να προκύπτουν στατιστικά αποδεκτά αποτελέσματα. Με τη διαστρωμάτωση προκύπτουν ακριβή αποτελέσματα (γεγονός που συνδυάζεται και με τα σχετικά στενά διαστήματα εμπιστοσύνης που προκύπτουν), αλλά τα «βάρη» είναι «στατιστικές κατασκευές» που προκύπτουν από τα δεδομένα και δεν μπορούν να καθοριστούν λεπτομερώς. Η προτυποποίηση, σε αντίθεση με τη διαστρωμάτωση, μπορεί να χρησιμοποιεί «βάρη», τα οποία είναι μη αποδοτικά. Αυτό συμβαίνει όταν μεγάλα «βάρη» συνδυάζονται με στρώματα με λίγα δεδομένα και αντίστροφα. Στην περίπτωση της προτυποποίησης, ωστόσο, τα «βάρη» είναι σαφή και εύκολα κατανοητά από τον οποιονδήποτε. Για ακόμη μεγαλύτερη διευκόλυνση, τα «βάρη» που χρησιμοποιούνται στην προτυποποίηση θα πρέπει να παρουσιάζονται αναλυτικά μαζί με τα αποτελέσματα μιας μελέτης, έτσι ώστε να είναι δυνατή η πραγματοποίηση συγκρίσεων και με άλλες μελέτες. Έτσι, η προτυποποίηση μπορεί να είναι λιγότερο αποδοτική, αλλά διευκολύνει σημαντικά την πραγματοποίηση συγκρίσεων. Τέλος, πρέπει να σημειωθεί ότι, ακόμη και αν το αποτέλεσμα της έκθεσης είναι σημαντικά διαφορετικό στα διάφορα στρώματα, η διαστρωματική ανάλυση των δεδομένων και η παρουσίαση των αποτελεσμάτων στα ειδικά στρώματα, χωρίς τον υπολογισμό ενός «συγκεντρωτικού» μέτρου σχέσης, είναι εξαιρετικά χρήσιμη και απαραίτητη για την εξαγωγή ασφαλών συμπερασμάτων.

ABSTRACT

Standardization

P. GALANIS, L. SPAROS

*Laboratory of Clinical Epidemiology, School of Nursing, University of Athens, Athens, Greece**Archives of Hellenic Medicine 2007, 24(5):471–486*

Standardization is a method of controlling the confounding introduced by different age distributions in two or more compared groups. It is a method of combining category-specific rates into a single summary value by taking a weighted average. It weights the category-specific rates using weights derived from a standard population. The weights, in effect, define the standard population. Although a single set of rates can be standardized, the main reason to standardize is to facilitate comparisons and therefore, usually two or more sets of rates are standardized. In order to compare rates for exposed and unexposed people, both groups would be standardized to the same standard. This form of standardization has long been referred to as “direct” standardization, to distinguish it from alternative approaches, the most popular of which is “indirect” standardization. Whereas “direct” standardization is a process involving the weighting of a set of observed category-specific rates according to a standard population, “indirect” standardization is a process in which the standard, instead of supplying the weighting distribution, supplies a standard set of rates, which are then weighted to the distribution of the population under study. The process is used to generate an “expected” rate or an expected number for the crude rate or total number of cases in the population. The comparison between the study population and the standard is generally presented as a standardized morbidity (or mortality) ratio, which is the ratio of the “observed” number of cases to an “expected” number.

Key words: Confounder, Confounding, Standard population, Standardization, Stratification

Βιβλιογραφία

1. ΣΠΑΡΟΣ Λ. *Μετα-επιδημιολογία*. Εκδόσεις ΒΗΤΑ, Αθήνα, 2001
2. ΣΠΑΡΟΣ Λ, ΓΑΛΑΝΗΣ Π, ΖΑΧΟΣ Ι, ΤΣΙΛΙΔΗΣ Κ. *Επιδημιολογία Ι*. Εκδόσεις ΒΗΤΑ, Αθήνα, 2004
3. ΜΙΕΤΤΙΝΕΝ ΟΣ. *Theoretical epidemiology. Principles of occurrence research in medicine*. John Wiley & Sons, New York, 1985
4. ΓΑΛΑΝΗΣ Π. Αλληλεπίδραση παραγόντων κινδύνου στην επιδημιολογία. Μεταπτυχιακή διπλωματική εργασία, Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Νοσηλευτικής, Αθήνα, 2003
5. ROTHMAN KJ. *Epidemiology: An introduction*. Oxford University Press, New York, 2002
6. ΚΑΗΝ ΗΑ, ΣΕΜΠΟΣ ΣΤ. *Statistical methods in epidemiology*. Oxford University Press, Oxford, 1989:105–108
7. CLAYTON D, HILLS M. *Statistical models in epidemiology*. Oxford University Press, Oxford, 1993:261–270
8. STRIKE PW. *Statistical methods in laboratory medicine*. Butterworth-Heinemann, Oxford, 1991:499–502
9. STROM BL. *Pharmacoepidemiology*. 2nd ed. John Wiley & Sons, New York, 1994
10. ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Τροποποίηση του μέτρου αποτελέσματος και σύγχυση στην εφαρμοσμένη ιατρική έρευνα. *Αρχ Ελλ Ιατρ* 2005, 22:170–177
11. ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Στατιστική αλληλεπίδραση και τροποποίηση του μέτρου αποτελέσματος. *Αρχ Ελλ Ιατρ* 2005, 21:137–147
12. ROTHMAN KJ, GREENLAND S. *Modern epidemiology*. 2nd ed. Lippincott Williams & Wilkins, Philadelphia, 1998
13. ΣΠΑΡΟΣ Λ, ΛΑΜΠΡΟΥ Α, ΜΕΛΛΟΥ Κ. *Επιδημιολογία ΙΙ*. Εκδόσεις ΒΗΤΑ, Αθήνα, 2005
14. ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Ανάλυση δεδομένων: Μη μπαγιασιανή προσέγγιση. *Αρχ Ελλ Ιατρ* 2005, 22:377–391
15. SELVIN S. *Statistical analysis of epidemiologic data*. 2nd ed. Oxford University Press, New York, 1996:3–21
16. MORRISON AS. *Screening in chronic disease*. Oxford University Press, New York, 1985:25–29
17. ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Μέτρα συχνότητας των νοσημάτων. *Αρχ Ελλ Ιατρ* 2005, 22:178–191
18. FREEMAN J, GOLDMANN DA, MCGOWAN JE. Confounding and the analysis of multiple variables in hospital epidemiology. *Infect Control* 1987, 8:465–473
19. 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
20. CONSONNI D, BERTAZZI PA, ZOCCHETTI C. Why and how to control for age in occupational epidemiology. *Occup Environ Med* 1997, 54:772–776
21. ASMAR L Farr William. In: Armitage P, Colton T (eds) *Encyclopedia of biostatistics*. John Wiley & Sons, New York, 1998:1507–1509

22. ΠΑΠΑΕΥΑΓΓΕΛΟΥ Γ, ΤΣΙΜΠΟΣ Κ. *Ιατρική δημογραφία και οικογενειακός προγραμματισμός*. Εκδόσεις ΒΗΤΑ, Αθήνα, 1993:23–37
23. ΚΑΤΣΟΥΓΙΑΝΝΟΠΟΥΛΟΣ ΒΧ. *Υγιεινή και Κοινωνική Ιατρική*. 2η έκδοση. Εκδοτικός οίκος Αδελφών Κυριακίδη, Θεσσαλονίκη, 1997:96–102.3
24. ΜΕΜΑΗΟΝ Β, ΤΡΙΧΟΠΟΥΛΟΣ Δ. *Epidemiology. Principles and methods*. 2nd ed. Little, Brown & Co, London, 1996
25. ΤΡΙΧΟΠΟΥΛΟΣ Δ. *Γενική και κλινική επιδημιολογία*. Εκδόσεις Παρισιάνος, Αθήνα, 2002
26. ΜΠΑΜΠΙΝΙΩΤΗΣ Γ. *Λεξικό της νέας ελληνικής γλώσσας*. Κέντρο Λεξικολογίας, Αθήνα, 1998:1520
27. ΣΠΑΡΟΣ ΛΔ. Η έννοια της νοσηρότητας. *Αρχ Ελλ Ιατρ* 2001, 18:303–311
28. INSKIP H. Standardization methods. In: Armitage P, Colton T (eds) *Encyclopedia of biostatistics*. John Wiley & Sons, New York, 1998:4237–4250
29. AHLBOM A, NORELL S. *Εισαγωγή στη σύγχρονη επιδημιολογία* (ελληνική μετάφραση). Ιατρικές Εκδόσεις Λίτσας, Αθήνα, 1992
30. DECOUFLE P, THOMAS TL, PICKLE LW. Comparison of the proportionate mortality ratio and standardized mortality ratio risk measures. *Am J Epidemiol* 1980, 111:263–269
31. ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Διαστρωματική ανάλυση δεδομένων. *Αρχ Ελλ Ιατρ* 2005, 21:378–384
32. ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Βιολογική και στατιστική αλληλεπίδραση. *Αρχ Ελλ Ιατρ* 2004, 21:123–136

Corresponding author:

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

.....