



ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΜΕ ΕΦΑΡΜΟΓΕΣ
ΣΤΗ ΒΙΟΙΑΤΡΙΚΗ

**Προσδιορισμός μεγέθους δείγματος στις κλινικές δοκιμές: η
περίπτωση των κλινικών δοκιμών στην Ελλάδα**

Φωταρίδου Ελένη

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ
Υπεύθυνος
Σαχλάς Αθανάσιος
Επίκουρος Καθηγητής

Λαμία, 28/9/2025



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗ
ΒΙΟΙΑΤΡΙΚΗ**

**Προσδιορισμός μεγέθους δείγματος στις κλινικές δοκιμές: η
περίπτωση των κλινικών δοκιμών στην Ελλάδα**

Φωταρίδου Ελένη

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

**Επιβλέπων/σα
Σαχλάς Αθανάσιος
Επίκουρος Καθηγητής**

Λαμία, 28/9/2025

Με ατομική μου ευθύνη και γνωρίζοντας τις κυρώσεις ⁽¹⁾, που προβλέπονται από της διατάξεις της παρ. 6 του άρθρου 22 του Ν. 1599/1986, δηλώνω ότι:

1. Δεν παραθέτω κομμάτια βιβλίων ή άρθρων ή εργασιών άλλων αυτολεξεί **χωρίς να τα περικλείω σε εισαγωγικά** και χωρίς να αναφέρω το συγγραφέα, τη χρονολογία, τη σελίδα. Η αυτολεξεί παράθεση χωρίς εισαγωγικά χωρίς αναφορά στην πηγή, είναι λογοκλοπή. Πέραν της αυτολεξεί παράθεσης, λογοκλοπή θεωρείται και η παράφραση εδαφίων από έργα άλλων, συμπεριλαμβανομένων και έργων συμφοιτητών μου, καθώς και η παράθεση στοιχείων που άλλοι συνέλεξαν ή επεξεργάστηκαν, χωρίς αναφορά στην πηγή. Αναφέρω πάντοτε με πληρότητα την πηγή κάτω από τον πίνακα ή σχέδιο, όπως στα παραθέματα.
2. Δέχομαι ότι η αυτολεξεί **παράθεση χωρίς εισαγωγικά**, ακόμα κι αν συνοδεύεται από αναφορά στην πηγή σε κάποιο άλλο σημείο του κειμένου ή στο τέλος του, είναι αντιγραφή. Η αναφορά στην πηγή στο τέλος π.χ. μιας παραγράφου ή μιας σελίδας, δεν δικαιολογεί συρραφή εδαφίων έργου άλλου συγγραφέα, έστω και παραφρασμένων, και παρουσίασή τους ως δική μου εργασία.
3. Δέχομαι ότι υπάρχει επίσης περιορισμός στο μέγεθος και στη συχνότητα των παραθεμάτων που μπορώ να εντάξω στην εργασία μου εντός εισαγωγικών. Κάθε μεγάλο παράθεμα (π.χ. σε πίνακα ή πλαίσιο, κλπ), προϋποθέτει ειδικές ρυθμίσεις, και όταν δημοσιεύεται προϋποθέτει την άδεια του συγγραφέα ή του εκδότη. Το ίδιο και οι πίνακες και τα σχέδια
4. Δέχομαι όλες τις συνέπειες σε περίπτωση λογοκλοπής ή αντιγραφής.

Ημερομηνία:/...../20.....

Ο – Η Δηλ.

(Υπογραφή)

(1) «Όποιος εν γνώσει του δηλώνει ψευδή γεγονότα ή αρνείται ή αποκρύπτει τα αληθινά με έγγραφη υπεύθυνη δήλωση του άρθρου 8 παρ. 4 Ν. 1599/1986 τιμωρείται με φυλάκιση τουλάχιστον τριών μηνών. Εάν ο υπαίτιος αυτών των πράξεων σκόπευε να προσπορίσει στον εαυτόν του ή σε άλλον περιουσιακό όφελος βλάπτοντας τρίτον ή σκόπευε να βλάψει άλλον, τιμωρείται με κάθειρξη μέχρι 10 ετών.

**Προσδιορισμός μεγέθους δείγματος στις κλινικές δοκιμές: η
περίπτωση των κλινικών δοκιμών στην Ελλάδα**
Ελένη Φωταρίδου

Τριμελής Επιτροπή:

Σαχλάς Αθανάσιος, Επίκουρος Καθηγητής

Κακαρούντας Αθανάσιος, Καθηγητής

Τσαρούχας Παναγιώτης, Καθηγητής

Περιεχόμενα

ΠΕΡΙΛΗΨΗ	vii
ΕΙΣΑΓΩΓΗ	1
ΜΕΘΟΔΟΛΟΓΙΑ	4
ΠΡΩΤΕΥΩΝ ΣΚΟΠΟΣ ΤΩΝ ΚΛΙΝΙΚΩΝ ΔΟΚΙΜΩΝ	8
ΦΑΣΕΙΣ	11
ΜΟΝΤΕΛΟ ΠΑΡΕΜΒΑΣΗΣ	13
ΤΥΦΛΟΠΟΙΗΣΗ, ΤΥΧΑΙΟΠΟΙΗΣΗ ΚΑΙ ΔΗΜΟΓΡΑΦΙΚΑ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ	15
ΜΕΓΕΘΟΣ ΔΕΙΓΜΑΤΟΣ	20
ΤΥΠΟΙ ΥΠΟΛΟΓΙΣΜΟΥ ΜΕΓΕΘΟΥΣ ΔΕΙΓΜΑΤΟΣ	21
ΣΥΓΚΡΙΣΗ ENROLLMENT ΚΑΙ SAMPLE SIZE (Protocol)	34
ΣΤΑΤΙΣΤΙΚΗ ΠΡΟΣΕΓΓΙΣΗ ΔΕΔΟΜΕΝΩΝ	36
ΣΥΖΗΤΗΣΗ	46
ΣΥΜΠΕΡΑΣΜΑΤΑ	49
ΒΙΒΛΙΟΓΡΑΦΙΑ	50

ΠΕΡΙΛΗΨΗ

Ο προσδιορισμός του μεγέθους του δείγματος είναι ένα σημαντικό μέρος του σχεδιασμού των κλινικών δοκιμών. Σκοπός της παρούσας εργασίας είναι να αξιολογηθεί σε ποιον βαθμό οι κλινικές δοκιμές που έχουν διεξαχθεί στην Ελλάδα μέχρι σήμερα συμμορφώνονται με όσα έχουν δηλωθεί στα επίσημα πρωτόκολλά τους, εστιάζοντας συγκεκριμένα στο κομμάτι του μεγέθους δείγματος. Πιο συγκεκριμένα, εξετάζεται κατά πόσο ο αριθμός των συμμετεχόντων που τελικά εντάχθηκαν στη μελέτη αντιστοιχεί στον αριθμό που είχε προβλεφθεί και καταγραφεί εξ αρχής στο εγκεκριμένο πρωτόκολλο της κάθε δοκιμής. Στη συνέχεια, πραγματοποιείται αξιολόγηση του τρόπου με τον οποίο προσδιορίζεται το μέγεθος δείγματος των κλινικών δοκιμών. Η διερεύνηση πραγματοποιήθηκε αντλώντας πληροφορίες από τη δημόσια διαθέσιμη βάση δεδομένων των ClinicalTrials.gov. Τα αποτελέσματα της ανάλυσης έδειξαν ότι οι κλινικές δοκιμές στην Ελλάδα συμβαδίζουν σε μεγάλο βαθμό με τα διεθνή πρότυπα, αλλά εξακολουθούν να υπάρχουν αδυναμίες κυρίως στον τομέα που αφορά το τελικό μέγεθος δείγματος. Οι αποκλίσεις ανάμεσα σε όσα προγραμματίζονται και σε όσα τελικά υλοποιούνται μειώνουν την αξιοπιστία των αποτελεσμάτων, ενώ η ποικιλία μεθόδων υπολογισμού αν και εκ πρώτης δείχνει έλλειψη ομοιομορφίας δικαιολογείται από το αντικείμενο μελέτης της κάθε κλινικής δοκιμής. Το μέγεθος δείγματος φαίνεται να εξαρτάται από πολλούς παράγοντες, όπως η φάση, η διάρκεια και η τυφλοποίηση, κάτι που σημαίνει ότι δεν μπορούμε να το βλέπουμε απομονωμένα. Με πιο ρεαλιστικό σχεδιασμό, καλύτερη τήρηση των πρωτοκόλλων και πιο στοχευμένη στρατολόγηση, οι δοκιμές στη χώρα μας μπορούν να γίνουν πιο αξιόπιστες και να αποκτήσουν μεγαλύτερη διεθνή αξία.

Προσδιορισμός μεγέθους δείγματος στις κλινικές δοκιμές: η περίπτωση των κλινικών δοκιμών στην Ελλάδα

ΕΙΣΑΓΩΓΗ

Οι κλινικές δοκιμές είναι ερευνητικές μελέτες για φαρμακευτικά σκευάσματα, θεραπείες, εμβόλια ή ιατρικές παρεμβάσεις που διεξάγονται σε εθελοντές για να απαντηθούν επιστημονικά ερωτήματα και να βρεθούν καλύτερες μέθοδοι ώστε να διαγνωστούν, να θεραπευθούν ή να προληφθούν παθήσεις. Παράλληλα, οι μελέτες αυτές συχνά βοηθούν τους επιστήμονες να κατανοήσουν σε βάθος μηχανισμούς ασθενειών, να εντοπίσουν λάθη και παρενέργειες. Μια κλινική μελέτη βασίζεται σε άτομα και ασθενείς που εθελοντικά συμμετέχουν και λαμβάνουν ένα νέο φάρμακο ή θεραπεία ή ιατροτεχνολογικό βοήθημα ή συσκευή.

Η έρευνα ξεκινά στο εργαστήριο, όπου το προϊόν που μας ενδιαφέρει μελετάται σε κύτταρα και ζώα προκειμένου να εκτιμηθεί η επίδρασή του, η αποτελεσματική του δόση και η ασφάλεια. Μόνο εφόσον ολοκληρωθεί με επιτυχία η συγκεκριμένη φάση, οι πειραματικές θεραπείες έχουν την δυνατότητα να προχωρήσουν στο δεύτερο βήμα που αποτελείται από κλινικές δοκιμές πλέον σε ανθρώπους. Η κλινική έρευνα είναι το δεύτερο στάδιο ανάπτυξης που οδηγεί στην επικύρωση συνήθως μίας νέας θεραπείας. Κατά τη διάρκεια της κλινικής ανάπτυξης, συλλέγονται ποικίλες πληροφορίες, όπως οι κίνδυνοι και ο βαθμός της αποτελεσματικότητας, μέσω διάφορων εξετάσεων ή ερωτηματολογίων, προκειμένου να εκτιμηθεί η αποτελεσματικότητα του φαρμάκου.

Μέσα από αυτές τις μελέτες, θεραπείες που κάποτε φάνταζαν αδύνατες σήμερα αποτελούν κοινή πρακτική στην ιατρική καθημερινότητα. Για να είναι όμως έγκυρα και αξιόπιστα τα αποτελέσματα μιας κλινικής δοκιμής, απαιτείται ο προσεκτικός και επιστημονικά τεκμηριωμένος σχεδιασμός τους. Οι κλινικές δοκιμές υπόκεινται σε αυστηρούς κανονισμούς και πραγματοποιούνται μόνο μετά από έγκριση αρμόδιων ρυθμιστικών αρχών (π.χ. ΕΟΦ στην Ελλάδα, EMA

στην Ευρώπη, FDA στις ΗΠΑ) και Επιτροπών Ηθικής και Δεοντολογίας. Επίσης, όλες οι μελέτες οφείλουν να ακολουθούν τις αρχές της Καλής Κλινικής Πρακτικής (Good Clinical Practice – GCP), μέσω των οποίων διασφαλίζεται η προστασία των συμμετεχόντων και η διαφάνεια των επιστημονικών αποτελεσμάτων.

Μια κλινική δοκιμή χαρακτηρίζεται από θεμελιώδη στοιχεία που εξασφαλίζουν τη μεθοδολογική πληρότητα και την επιστημονική της αξιοπιστία. Μεταξύ αυτών περιλαμβάνονται το μέγεθος δείγματος, η τυχαιοποίηση, η τυφλοποίηση, οι φάσεις της μελέτης, καθώς και τα κριτήρια ένταξης και αποκλεισμού. Σημαντικό ρόλο παίζουν επίσης τα χαρακτηριστικά των συμμετεχόντων (π.χ. ηλικία, φύλο, δημογραφικά δεδομένα), η ιατρική πάθηση ή κατάσταση στην οποία στοχεύει η παρέμβαση, καθώς και τα αποτελέσματα που προκύπτουν και δημοσιεύονται. Επιπλέον, αναφέρονται στοιχεία σχετικά με τον χορηγό ή φορέα χρηματοδότησης, τους συνεργαζόμενους οργανισμούς, τον τύπο της μελέτης, τον πραγματικό αριθμό συμμετεχόντων, τον τρόπο κατανομής (τυχαιοποιημένη ή όχι) και τις ημερομηνίες έναρξης και ολοκλήρωσης. Όλα αυτά τα δεδομένα συνθέτουν την αναλυτική περιγραφή μιας κλινικής δοκιμής, διασφαλίζοντας την εγκυρότητα και τη διαφάνειά της.

Το έγγραφο που διέπει τη διεξαγωγή μιας κλινικής δοκιμής και καθορίζει όλες τις δραστηριότητες που πραγματοποιούνται εντός του ερευνητικού κέντρου είναι γνωστό ως πρωτόκολλο μελέτης. Το πρωτόκολλο παρέχει αναλυτικές πληροφορίες για τον ερευνητικό σκοπό, τα κριτήρια για τη συμπερίληψη και αποκλεισμό εθελοντών στη μελέτη, το πρόγραμμα των δραστηριοτήτων, τις χρησιμοποιούμενες μεθόδους και θεραπευτικές παρεμβάσεις, τις ακριβείς δοσολογίες, τη χρονική διάρκεια της μελέτης, καθώς και τις διαδικασίες παρακολούθησης και αναφοράς των δεδομένων. Με αυτόν τον τρόπο, εξασφαλίζει τη συνέπεια, την εγκυρότητα και την αξιοπιστία της έρευνας σε όλα τα στάδια της κλινικής δοκιμής.

Σκοπός της παρούσας εργασίας είναι να αξιολογηθεί σε ποιον βαθμό οι κλινικές δοκιμές που έχουν διεξαχθεί στην Ελλάδα μέχρι σήμερα συμμορφώνονται με όσα έχουν δηλωθεί στα επίσημα πρωτόκολλά τους, εστιάζοντας συγκεκριμένα στο κομμάτι του μεγέθους δείγματος. Πιο συγκεκριμένα, εξετάζεται κατά πόσο ο αριθμός των συμμετεχόντων που τελικά εντάχθηκαν στη μελέτη αντιστοιχεί στον αριθμό που είχε προβλεφθεί και

καταγραφεί εξαρχής στο εγκεκριμένο πρωτόκολλο της κάθε δοκιμής. Στη συνέχεια, πραγματοποιείται αξιολόγηση του τρόπου με τον οποίο προσδιορίζεται το μέγεθος δείγματος των κλινικών δοκιμών. Η διερεύνηση πραγματοποιήθηκε αντλώντας πληροφορίες από τη δημόσια διαθέσιμη βάση δεδομένων των ClinicalTrials.gov. Η ανάλυση αυτή έχει ιδιαίτερη σημασία, καθώς η τήρηση του προκαθορισμένου μεγέθους δείγματος συνδέεται άμεσα με τη στατιστική ισχύ της μελέτης, την αξιοπιστία των αποτελεσμάτων και τη συνολική επιστημονική εγκυρότητα των συμπερασμάτων.

Δύο χιλιάδες και ενενήντα εννιά (2099) κλινικές δοκιμές είχαν ολοκληρωθεί στην Ελλάδα τη στιγμή που διεξήχθη αυτή η μελέτη. Από αυτές, μόνο οι 1021 έχουν δημοσιεύσει αποτελέσματα, ενώ από αυτές οι 454 είχαν ελεύθερο στο κοινό το πρωτόκολλό τους.

ΜΕΘΟΔΟΛΟΓΙΑ

Η μεθοδολογική προσέγγιση της παρούσας εργασίας βασίστηκε στη συστηματική συλλογή και δευτερογενή ανάλυση δεδομένων από τη βάση δεδομένων ClinicalTrials.gov. Το ClinicalTrials.gov είναι ένας διαδικτυακός ιστότοπος και ηλεκτρονική βάση δεδομένων που παρέχει αξιόπιστες πληροφορίες σχετικά με κλινικές ερευνητικές μελέτες και τα αποτελέσματά τους, με σκοπό την ενημέρωση του κοινού, των ερευνητών και των επαγγελματιών υγείας. Οι πληροφορίες που δημοσιεύονται βασίζονται στις καταχωρίσεις και τις ενημερώσεις των χορηγών και των ερευνητών. Η πλατφόρμα ενημερώνεται σχεδόν καθημερινά με νέες μελέτες και καλύπτει ερευνητικά προγράμματα που πραγματοποιούνται και στις 50 πολιτείες των ΗΠΑ καθώς και σε περισσότερες από 200 χώρες. Επιπλέον, υποστηρίζει το θεσμικό πλαίσιο που απαιτεί από τους χορηγούς και τους ερευνητές να δημοσιοποιούν τα στοιχεία και τα αποτελέσματα των κλινικών δοκιμών, προάγοντας έτσι τη διαφάνεια και την πρόσβαση στην επιστημονική γνώση.

Αρχικά, πραγματοποιήθηκε αναζήτηση στη βάση ClinicalTrials.gov με κριτήριο τον τόπο διεξαγωγής (Ελλάδα) και το στάδιο ολοκλήρωσης της δοκιμής. Ο συνολικός αριθμός των μελετών που εντοπίστηκαν ήταν 2099. Για κάθε κλινική δοκιμή συλλέχθηκε ένα αρχείο με τα εξής πεδία: NCT Number, Study Title, Study URL, Acronym, Study Status, Study Results, Conditions, Sponsor, Collaborators, Sex, Age, Phases, Enrollment, Sample Size (Protocol), Funder Type, Study Type, Allocation, Intervention Model, Masking, Primary Purpose, Start Date, Completion Date, Duration (days), First Posted, Results First Posted, Last Update Posted, Locations. Μία σύντομη περιγραφή των παραπάνω πεδίων παρουσιάζεται στον Πίνακα 1. Η αρχική αυτή συλλογή δεδομένων αποτέλεσε τη βάση για το επόμενο στάδιο φιλτραρίσματος και ανάλυσης.

Πεδίο	Περιγραφή
-------	-----------

NCT Number	Μοναδικός κωδικός ταυτοποίησης μελέτης στο ClinicalTrials.gov
Study Title	Τίτλος της κλινικής μελέτης
Study URL	Σύνδεσμος προς τη μελέτη στο ClinicalTrials.gov
Acronym	Συντομογραφία/όνομα της μελέτης
Study Status	Τρέχουσα κατάσταση μελέτης (π.χ. ολοκληρωμένη)
Study Results	Διαθεσιμότητα και δημοσίευση αποτελεσμάτων
Conditions	Παθήσεις ή καταστάσεις που μελετώνται
Sponsor	Οργανισμός που διεξάγει ή χρηματοδοτεί τη μελέτη
Collaborators	Συνεργαζόμενοι φορείς ή ιδρύματα
Sex	Φύλο συμμετεχόντων
Age	Ηλικιακό εύρος συμμετεχόντων
Phases	Φάση κλινικής δοκιμής
Enrollment	Τελικός αριθμός συμμετεχόντων
Sample Size (Protocol)	Προβλεπόμενο μέγεθος δείγματος στο πρωτόκολλο
Funder Type	Είδος χρηματοδότη (π.χ. βιομηχανία, πανεπιστήμιο)
Study Type	Είδος μελέτης (π.χ. παρεμβατική, παρατηρητική)
Allocation	Τρόπος κατανομής συμμετεχόντων (τυχαιοποίηση)
Intervention Model	Σχέδιο παρέμβασης (π.χ. παράλληλη, crossover)
Masking	Βαθμός τυφλοποίησης
Primary Purpose	Κύριος στόχος της μελέτης (π.χ. θεραπευτικός)
Start Date	Ημερομηνία έναρξης μελέτης

Completion Date	Ημερομηνία ολοκλήρωσης μελέτης
Duration (days)	Συνολική διάρκεια σε ημέρες
First Posted	Ημερομηνία πρώτης ανάρτησης στο ClinicalTrials.gov
Results First Posted	Ημερομηνία πρώτης ανάρτησης αποτελεσμάτων
Last Update Posted	Τελευταία ενημέρωση της μελέτης
Locations	Γεωγραφικές τοποθεσίες διεξαγωγής μελέτης

Πίνακας 1

Στη συνέχεια, εφαρμόστηκαν κριτήρια ένταξης ώστε να διασφαλιστεί η εγκυρότητα και η πληρότητα της ανάλυσης. Συγκεκριμένα, στο δείγμα συμπεριλήφθηκαν μόνο οι μελέτες για τις οποίες υπήρχε διαθέσιμο πρωτόκολλο. Ο αριθμός των μελετών που πληρούσαν το συγκεκριμένο κριτήριο ήταν 454. Τα πρωτόκολλα αυτών των μελετών εξετάστηκαν προσεκτικά και καταγράφηκε το απαιτούμενο μέγεθος δείγματος, όπου αυτό αναφερόταν ρητά. Οι πληροφορίες αυτές καταχωρήθηκαν και οργανώθηκαν σε αρχείο Excel, ώστε να καταστεί δυνατή η περαιτέρω στατιστική επεξεργασία και ανάλυση.

Επιπλέον, υπολογίστηκε η διάρκεια κάθε κλινικής δοκιμής. Ο υπολογισμός βασίστηκε στη διαφορά ανάμεσα στην ημερομηνία ολοκλήρωσης και την ημερομηνία έναρξης της μελέτης. Σε περιπτώσεις όπου οι ημερομηνίες δεν ήταν πλήρως διαθέσιμες (δηλαδή αναφερόταν μόνο μήνας και έτος), θεωρήθηκε ως ημέρα η πρώτη του αντίστοιχου μήνα. Με τον τρόπο αυτό εξασφαλίστηκε η ομοιομορφία των δεδομένων και η συγκρισιμότητα των αποτελεσμάτων.

Η ανάλυση των δεδομένων πραγματοποιήθηκε με το στατιστικό λογισμικό IBM SPSS Statistics (version 29) και την γλώσσα R, τα οποία παρέχουν τη δυνατότητα ορθής επεξεργασίας και συνοπτικής παρουσίασης των μεταβλητών των κλινικών δοκιμών. Οι μεταβλητές οργανώθηκαν και

ομαδοποιήθηκαν κατάλληλα, ώστε να καταστεί εφικτή η εφαρμογή τόσο περιγραφικών όσο και συγκριτικών στατιστικών μεθόδων.

Για την περιγραφική παρουσίαση των δεδομένων χρησιμοποιήθηκαν πίνακες συχνοτήτων και μέσων όρων, που αποτυπώνουν τα βασικά χαρακτηριστικά των δεδομένων. Στο επόμενο στάδιο, για τη διερεύνηση συσχετίσεων και διαφορών μεταξύ μεταβλητών εφαρμόστηκαν τα κατάλληλα στατιστικά εργαλεία, όπως μη παραμετρικοί έλεγχοι, αναλύσεις συσχέτισης και πολυμεταβλητές τεχνικές.

Τα αποτελέσματα παρουσιάζονται σε συνοπτικούς πίνακες, οι οποίοι διευκολύνουν την κατανόηση και τη σύγκριση ανάμεσα στις διαφορετικές φάσεις των κλινικών δοκιμών, προσφέροντας μια σαφή και οργανωμένη εικόνα των ευρημάτων. Για την παραγωγή των πινάκων εφαρμόστηκαν μέθοδοι γραμμικής παλινδρόμησης και ανάλυσης σε συστάδες.

Η γραμμική παλινδρόμηση αποτελεί μια θεμελιώδη στατιστική μέθοδο, η οποία χρησιμοποιείται για τη διερεύνηση και ποσοτικοποίηση της σχέσης μεταξύ μιας εξαρτημένης μεταβλητής (μεταβλητή απόκρισης) και μίας ή περισσότερων ανεξάρτητων μεταβλητών (ερμηνευτικών παραγόντων). Σκοπός της είναι η περιγραφή αυτής της σχέσης μέσω μιας γραμμικής εξίσωσης, η οποία επιτρέπει την εκτίμηση και πρόβλεψη της τιμής της εξαρτημένης μεταβλητής βάσει των τιμών των ανεξάρτητων μεταβλητών.

Η ανάλυση κατά συστάδες (cluster analysis) είναι μια τεχνική μηχανικής μάθησης χωρίς επίβλεψη, η οποία χρησιμοποιείται για την επεξεργασία και οργάνωση δεδομένων. Σε αντίθεση με τις επιβλεπόμενες μεθόδους, δεν είναι εκ των προτέρων γνωστό πόσες και ποιες ομάδες δεδομένων υπάρχουν. Ο αλγόριθμος εντοπίζει αυτόματα στοιχεία με κοινά χαρακτηριστικά και τα οργανώνει σε συστάδες. Η τεχνική αυτή εφαρμόζεται σε σύνολα δεδομένων που δεν είναι αρχικά ταξινομημένα με βάση την ομοιότητά τους, διευκολύνοντας έτσι τον εντοπισμό προτύπων.

Η μεθοδολογική αυτή διαδικασία επέτρεψε την οργάνωση ενός αξιόπιστου συνόλου δεδομένων, το οποίο αποτελεί τη βάση για την περαιτέρω ανάλυση και συζήτηση ενώ η χρήση δευτερογενούς ανάλυσης δεδομένων από

διεθνώς αναγνωρισμένη βάση καταγραφής εγγυάται την αντικειμενικότητα της συλλογής .

ΠΡΩΤΕΥΩΝ ΣΚΟΠΟΣ ΤΩΝ ΚΛΙΝΙΚΩΝ ΔΟΚΙΜΩΝ

Ο πρωτεύων σκοπός (primary purpose) μιας κλινικής δοκιμής αναφέρεται στον κύριο λόγο για τον οποίο γίνεται η μελέτη. Καθορίζει πώς θα οργανωθεί η έρευνα, ποιοι θα συμμετάσχουν και με ποιον τρόπο θα συγκεντρωθούν και θα αναλυθούν τα δεδομένα. Με απλά λόγια, περιγράφει τι θέλει να πετύχει η μελέτη και δίνει ξεκάθαρη κατεύθυνση σε όλα τα επόμενα βήματα της ερευνητικής διαδικασίας.

Η ταξινόμηση του πρωτεύοντα σκοπού σε συγκεκριμένες κατηγορίες συμβάλλει στη συστηματική οργάνωση της κλινικής έρευνας. Η διαδικασία ξεκινά συχνά με μελέτες που στοχεύουν στον έλεγχο (screening), δηλαδή στον εντοπισμό νόσων σε ασυμπτωματικά άτομα, προκειμένου να διευκολυνθεί η έγκαιρη διάγνωση και η άμεση παρέμβαση. Στενά συνδεδεμένες με αυτές είναι οι μελέτες διάγνωσης (diagnostic), οι οποίες αξιολογούν νέες μεθόδους και τεχνολογίες με σκοπό την επιβεβαίωση της παρουσίας μιας ασθένειας με μεγαλύτερη ακρίβεια. Όταν ο εντοπισμός μιας νόσου έχει επιτευχθεί, έρχεται στο προσκήνιο ο πρωτεύοντας σκοπός της πρόληψης (prevention), που αφορά παρεμβάσεις για την αποφυγή εμφάνισης νόσων σε υγιή πληθυσμό ή υποτροπών σε ασθενείς με προηγούμενο ιστορικό.

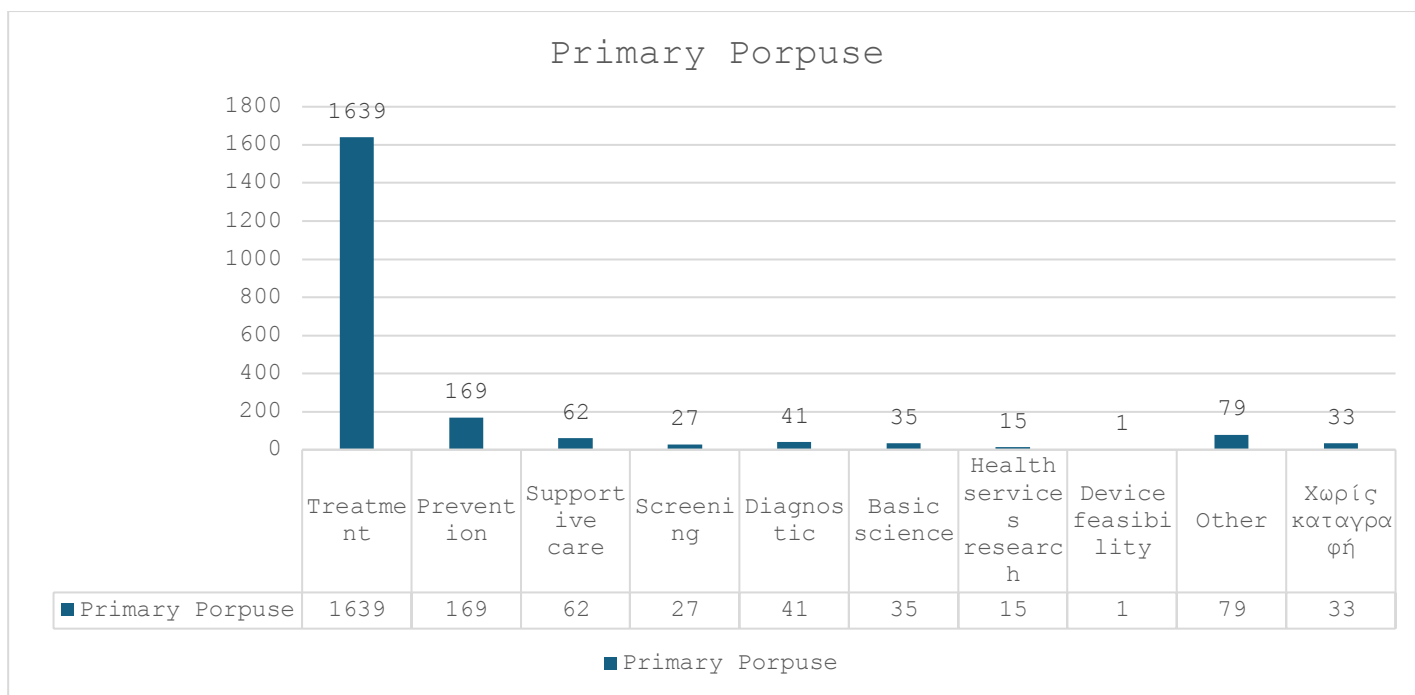
Σε επόμενο στάδιο, οι περισσότερες κλινικές δοκιμές επικεντρώνονται στη θεραπεία (treatment), η οποία αποτελεί τον πιο συχνό σκοπό. Εδώ εξετάζονται η ασφάλεια και η αποτελεσματικότητα νέων φαρμάκων, και θεραπειών με στόχο τη βελτίωση της κλινικής εικόνας και την παράταση της επιβίωσης των ασθενών. Παράλληλα, σημαντική είναι και η κατηγορία της υποστηρικτικής φροντίδας (supportive care), η οποία δεν αποσκοπεί στην ίαση, αλλά στην ανακούφιση των συμπτωμάτων και στη βελτίωση της ποιότητας ζωής.

Εκτός από τις φαρμακευτικές ή βιολογικές παρεμβάσεις, η κλινική έρευνα συχνά στρέφεται και σε τεχνολογικές λύσεις. Οι μελέτες σκοπιμότητας

ιατροτεχνολογικών προϊόντων (device feasibility) εστιάζουν στον αρχικό έλεγχο της ασφάλειας και της λειτουργικότητας νέων ιατρικών συσκευών, ώστε να διαπιστωθεί εάν μπορούν να προχωρήσουν σε πιο εκτεταμένες κλινικές δοκιμές. Συμπληρώνοντας, οι μελέτες βασικής επιστήμης (basic science) αναφέρονται σε μελέτες που αποσκοπούν στη σε βάθος κατανόηση βιολογικών, φυσιολογικών ή παθολογικών μηχανισμών, διερευνώντας, για παράδειγμα, γενετικούς, μοριακούς ή ανοσολογικούς παράγοντες, χωρίς να αξιολογούν άμεσα θεραπευτικές ή διαγνωστικές παρεμβάσεις, αλλά θέτοντας τα θεμέλια για μελλοντική κλινική εφαρμογή.

Τέλος, ιδιαίτερη θέση κατέχουν οι μελέτες που εστιάζουν στο σύστημα υγείας συνολικά. Ο σκοπός της έρευνας υπηρεσιών υγείας (health services research) είναι να αξιολογήσει την ποιότητα, την αποδοτικότητα και την ισότητα στην παροχή φροντίδας, προσφέροντας πολύτιμα στοιχεία για τη χάραξη πολιτικής υγείας. Σε περιπτώσεις που ο ερευνητικός στόχος δεν εντάσσεται σε κάποια από τις παραπάνω κατηγορίες, χρησιμοποιείται η επιλογή «άλλο» (other), η οποία αφορά εξειδικευμένα ή πολυπαραγοντικά ερευνητικά ερωτήματα που δεν μπορούν να καλυφθούν από τις παραπάνω ταξινομήσεις.

Από το σύνολο των 2099 κλινικών δοκιμών, οι περισσότερες, δηλαδή 1639 (ποσοστό 78,1%) είχαν ως πρωτεύοντα σκοπό τη θεραπεία, γεγονός που επιβεβαιώνει πως το κύριο ενδιαφέρον της κλινικής έρευνας εστιάζει στη δοκιμή φαρμάκων και θεραπευτικών παρεμβάσεων. Η πρόληψη αποτέλεσε τον δεύτερο συχνότερο σκοπό με 167 δοκιμές (8,0%), ενώ ακολούθησαν η υποστηρικτική φροντίδα με 62 δοκιμές (3,0%) και η κατηγορία «άλλο» με 79 δοκιμές (3,8%). Λιγότερο συχνές ήταν οι μελέτες screening, που ήταν 27 συνολικά (1,3%), 41 διαγνωστικές (2,0%) και 35 βασικής επιστήμης (1,7%). Ελάχιστες ήταν οι δοκιμές με επίκεντρο την έρευνα υπηρεσιών υγείας (15 δοκιμές, 0,7%) και τη σκοπιμότητα ιατροτεχνολογικών προϊόντων (1 δοκιμή, 0,05%). Σε 33 δοκιμές (1,6%) δεν αναφερόταν ο πρωτεύοντας σκοπός. Στο Σχήμα 1 παρουσιάζεται η κατανομή και η συχνότητα του πρωτεύοντα σκοπού.



Σχήμα 1

Τα δεδομένα αναδεικνύουν ότι το τοπίο της κλινικής έρευνας στην Ελλάδα κυριαρχείται από θεραπευτικές μελέτες, γεγονός αναμενόμενο, καθώς αυτές αποτελούν τη βάση για την έγκριση νέων φαρμάκων και θεραπειών. Παράλληλα, το σχετικά χαμηλό ποσοστό προληπτικών δοκιμών αναδεικνύει την ανάγκη μεγαλύτερης επένδυσης σε παρεμβάσεις που θα μπορούσαν να μειώσουν την εμφάνιση ή την υποτροπή ασθενειών. Η μικρή συμμετοχή ερευνητικών προσπαθειών σε τομείς, όπως οι υπηρεσίες υγείας ή η αξιολόγηση ιατροτεχνολογικών προϊόντων υποδηλώνει περιορισμένη έμφαση σε αυτούς τους τομείς, αν και η ανάπτυξή τους θα μπορούσε να προσφέρει σημαντικά οφέλη τόσο για την αποτελεσματικότητα των συστημάτων υγείας όσο και για την καινοτομία στην ιατρική τεχνολογία.

ΦΑΣΕΙΣ

Στην κλινική έρευνα, η ανάπτυξη ενός νέου φαρμάκου ή θεραπείας ακολουθεί μια δομημένη και σταδιακή διαδικασία που χωρίζεται σε επιμέρους φάσεις. Κάθε φάση έχει σαφείς στόχους και συγκεκριμένο σχεδιασμό. Η σωστή κατανόηση των φάσεων είναι κρίσιμη, καθώς διασφαλίζει ότι οι αποφάσεις που λαμβάνονται κατά την εξέλιξη μιας κλινικής δοκιμής βασίζονται σε επιστημονικά τεκμηριωμένα δεδομένα και ότι η υγεία των συμμετεχόντων προστατεύεται πλήρως. Οι δοκιμές συνήθως κατατάσσονται σε τέσσερις Φάσεις.

Στις μελέτες Φάσης I (PHASE 1), οι ερευνητές μελετούν μία διερευνητική θεραπεία για πρώτη φορά σε μία μικρή ομάδα ατόμων (20-100), συνήθως υγιών εθελοντών, προκειμένου να αξιολογηθεί ο τρόπος απορρόφησης, κατανομής, μεταβολισμού και απομάκρυνσης του εκάστοτε φαρμάκου από το σώμα, προκειμένου να εντοπιστεί η βέλτιστη οδός χορήγησης και δόσης. Σε μικρή συχνότητα, ορισμένες δοκιμές έχουν ένα προγενέστερο στάδιο που ονομάζεται Φάση 0 (EARLY PHASE 1). Αυτές είναι γνωστές ως μελέτες μικροδοσολογίας σε ανθρώπους και έχουν σχεδιαστεί για να επιταχύνουν την ανάπτυξη συνήθως πολλά υποσχόμενων φαρμάκων.

Στις κλινικές μελέτες Φάσης II (PHASE 2), η διερευνητική θεραπεία χορηγείται σε μία στοχευμένη ομάδα ασθενών (100-300) προκειμένου να ελεγχθεί η βέλτιστη δοσολογία και να αξιολογηθεί περαιτέρω η ασφάλεια της χρήσης της. Οι συμμετέχοντες στη Φάση II είναι ασθενείς που πάσχουν από την νόσο για την οποία αναπτύσσεται η φαρμακευτική αγωγή.

Στις κλινικές μελέτες Φάσης III (PHASES 3), η διερευνητική θεραπεία χορηγείται σε ομάδες ασθενών (300-3000) προκειμένου να επιβεβαιωθεί η αποτελεσματικότητα συγκριτικά με τις καθιερωμένες θεραπείες ώστε να συλλεχθούν πληροφορίες που θα επιτρέψουν στο φάρμακο ή τη θεραπεία να χρησιμοποιηθεί με τον πιο ωφέλιμο τρόπο για τον ασθενή. Οι θεραπείες που ολοκληρώνουν τις κλινικές δοκιμές Φάσης III, εφόσον τα αποτελέσματά τους αξιολογηθούν θετικά από τις εγκριτικές αρχές, λαμβάνουν στη συνέχεια την απαραίτητη έγκριση κυκλοφορίας.

Οι μελέτες Φάσης IV (PHASES 4) που αφορούν σκευάσματα τα οποία έχουν ήδη κυκλοφορήσει και ενταχθεί στην κλινική πράξη, καταγράφονται

πρόσθετες πληροφορίες από την καθημερινή κλινική εφαρμογή τους, ειδικότερα τα οφέλη, οι ανεπιθύμητες ενέργειες (παρενέργειες) προκειμένου να επιβεβαιωθεί η καλύτερη δυνατή χρήση τους. Επιπλέον με τις κλινικές μελέτες Φάσης IV αξιολογούνται τα μακροπρόθεσμα αποτελέσματα των φαρμακευτικών προϊόντων και αξιολογούνται σπάνιες μεν αλλά δυνητικά σοβαρές ανεπιθύμητες ενέργειες.

Φάση	Πρωταρχικός στόχος	Δόση	Παρακολούθηση ασθενούς	Τυπικός αριθμός συμμετεχόντων
Φάση 0	Φαρμακοκινητική, ιδιαίτερα βιοδιαθεσιμότητα από το στόμα και χρόνος ημιζωής του φαρμάκου	Μικρό, υποθεραπευτικό	Κλινικός ερευνητής	10 άτομα
Φάση I	Δοσολογία σε υγιείς εθελοντές για λόγους ασφάλειας	Συχνά υποθεραπευτικό, με αυξανόμενες δόσεις	Κλινικός ερευνητής	20–100 φυσιολογικοί υγιείς εθελοντές (ή ασθενείς με καρκίνο)
Φάση II	Δοκιμή του φαρμάκου στους συμμετέχοντες για αξιολόγηση αποτελεσματικότητας και παρενεργειών	Θεραπευτική δόση	Κλινικός ερευνητής	100–300 συμμετέχοντες με συγκεκριμένη ασθένεια
Φάση III	Δοκιμή του φαρμάκου σε συμμετέχοντες για αξιολόγηση αποτελεσματικότητας, ασφάλειας και αποτελεσματικότητας	Θεραπευτική δόση	Κλινικός ερευνητής και προσωπικός ιατρός	300–3.000 άτομα με συγκεκριμένη ασθένεια
Φάση IV	Δημόσια επιτήρηση μετά την κυκλοφορία	Θεραπευτική δόση	Προσωπικός γιατρός	Όποιος αναζητά θεραπεία

Πίνακας 2

ΜΟΝΤΕΛΟ ΠΑΡΕΜΒΑΣΗΣ

Στις κλινικές δοκιμές, το Intervention Model (ή αλλιώς «Μοντέλο Παρέμβασης») αποτελεί θεμελιώδες στοιχείο του σχεδιασμού τους, καθώς προσδιορίζει με ποιον τρόπο κατανέμονται οι συμμετέχοντες σε ομάδες και πώς εκτίθενται στην υπό μελέτη παρέμβαση. Η επιλογή του κατάλληλου μοντέλου εξαρτάται από το ερευνητικό ερώτημα, τη φύση της παρέμβασης και το στάδιο ανάπτυξης του φαρμάκου ή της θεραπευτικής μεθόδου.

Το πιο απλό μοντέλο είναι το Single Group Assignment (κατανομή μονής ομάδας), στο οποίο υπάρχει μόνο ένα θεραπευτικό σκέλος διαθέσιμο σε όλους τους συμμετέχοντες στη μελέτη. Με άλλα λόγια, όλοι οι συμμετέχοντες στη δοκιμή λαμβάνουν την ίδια θεραπεία (συμπεριλαμβανομένης της δοσολογίας, της συχνότητας και της μεθόδου χορήγησης) που μελετάται καθ' όλη τη διάρκεια της δοκιμής. Δεν υπάρχει σύγκριση μεταξύ διαφορετικών ομάδων, και η αξιολόγηση επικεντρώνεται στην παρακολούθηση της επίδρασης της παρέμβασης μέσα στην ίδια ομάδα. Το μοντέλο αυτό χρησιμοποιείται συχνά σε πρώιμα στάδια κλινικών δοκιμών (π.χ. Φάση I ή II), όπου ο στόχος είναι η αξιολόγηση της ασφάλειας και αποτελεσματικότητας μιας παρέμβασης. 327 δοκιμές χρησιμοποίησαν κατανομή μονής ομάδας.

Το πιο διαδεδομένο μοντέλο είναι το Parallel Assignment (παράλληλη κατανομή), όπου οι συμμετέχοντες κατανέμονται σε δύο ή περισσότερες ομάδες και κάθε ομάδα λαμβάνει διαφορετική παρέμβαση (π.χ. φάρμακο έναντι placebo (αδρανές φάρμακο) ή σύγκριση δύο φαρμάκων) καθ' όλη τη διάρκεια της μελέτης. Κάθε συμμετέχοντας παραμένει στην ίδια ομάδα σε όλη τη διάρκεια της δοκιμής, γεγονός που απλοποιεί την ανάλυση και μειώνει τον κίνδυνο σύγχυσης. Για παράδειγμα, μια δοκιμή με δύο βραχίονες παράλληλης κατανομής περιλαμβάνει δύο ομάδες συμμετεχόντων. Η μία ομάδα λαμβάνει το φάρμακο Α και η άλλη ομάδα λαμβάνει το φάρμακο Β. Έτσι, κατά τη διάρκεια της δοκιμής, οι συμμετέχοντες στη μία ομάδα λαμβάνουν το φάρμακο Α «παράλληλα» με τους συμμετέχοντες στην άλλη ομάδα, οι οποίοι λαμβάνουν το φάρμακο Β. Όπως είναι αναμενόμενο στην πλειοψηφία των κλινικών δοκιμών της μελέτης, δηλαδή στις 1564, επιλέχθηκε το μοντέλο παράλληλης κατανομής και η ποσοστιαία ανομοιομορφία τους με τα άλλα μοντέλα παρουσιάζεται στον Σχήμα 2.

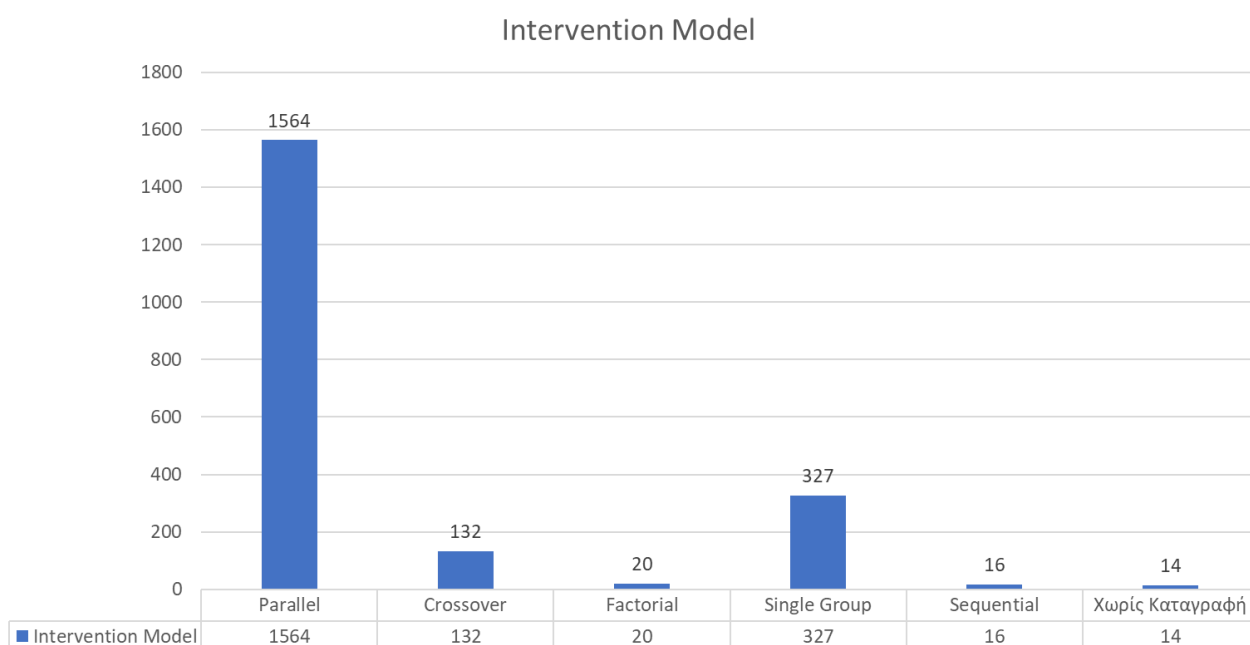
Ένα εξίσου σημαντικό μοντέλο είναι το Crossover Assignment (διασταυρούμενη κατανομή), στο οποίο οι συμμετέχοντες λαμβάνουν διαδοχικά περισσότερες από μία παρεμβάσεις σε συγκεκριμένη σειρά, με ενδιάμεσες περιόδους (washout periods) ώστε να αποτραπεί η αλληλεπίδραση μεταξύ τους. Σε αντίθεση με την παράλληλη κατανομή, μια δοκιμή διασταυρούμενης κατανομής δύο επί δύο περιλαμβάνει δύο ομάδες συμμετεχόντων. Η μία ομάδα λαμβάνει το φάρμακο Α κατά την αρχική φάση της δοκιμής, ακολουθούμενο από το φάρμακο Β σε μεταγενέστερη φάση. Η άλλη ομάδα λαμβάνει το φάρμακο Β κατά την αρχική φάση, ακολουθούμενο από το φάρμακο Α. Όλοι οι συμμετέχοντες λαμβάνουν το φάρμακο Α και το φάρμακο Β σε κάποιο σημείο της δοκιμής, αλλά με διαφορετική σειρά, ανάλογα με την ομάδα στην οποία έχουν κατανεμηθεί. Ένα σημαντικό πλεονέκτημα αυτού του μοντέλου είναι ότι η σύγκριση πραγματοποιείται ενδοατομικά, με τον ίδιο τον συμμετέχοντα να λειτουργεί ως εσωτερικό σημείο αναφοράς, μειώνοντας τη διακύμανση και επιτρέποντας την ανίχνευση διαφορών με μικρότερο δείγμα. Το Crossover μοντέλο ακολούθησαν 132 από τις συνολικές δοκιμές της χώρας.

Ένα τρίτο μοντέλο, με βάση το οποίο πραγματοποιήθηκαν 20 κλινικές δοκιμές, είναι το Factorial Assignment (παραγοντική κατανομή), όπου οι συμμετέχοντες κατανέμονται σε ομάδες που λαμβάνουν διαφορετικούς συνδυασμούς παρεμβάσεων. Σε μια παραγοντική δύο επί δύο δοκιμή περιλαμβάνονται τέσσερις ομάδες συμμετεχόντων. Κάθε ομάδα λαμβάνει έναν από τους ακόλουθους συνδυασμούς παρεμβάσεων: (1) φάρμακο Α και φάρμακο Β, (2) φάρμακο Α και εικονικό φάρμακο (placebo), (3) εικονικό φάρμακο και φάρμακο Β, ή (4) εικονικό φάρμακο και εικονικό φάρμακο. Έτσι, κατά τη διάρκεια της δοκιμής, όλοι οι δυνατοί συνδυασμοί των δύο φαρμάκων (Α και Β) και των placebos χορηγούνται σε διαφορετικές ομάδες συμμετεχόντων, με αποτέλεσμα να μελετάται ταυτόχρονα η επίδραση δύο ή περισσότερων θεραπειών, καθώς και οι πιθανές αλληλεπιδράσεις μεταξύ τους.

Τέλος, μόνο 16 κλινικές δοκιμές πραγματοποιήθηκαν βασισμένες στο Sequential Assignment (σειριακή κατανομή). Είναι ένα μοντέλο παρέμβασης, όπου οι συμμετέχοντες λαμβάνουν δύο ή περισσότερες παρεμβάσεις σε προκαθορισμένη σειρά. Σε αντίθεση με τα παράλληλη ή παραγοντικά μοντέλα, κάθε συμμετέχοντας εκτίθεται σε όλες τις παρεμβάσεις, αλλά όχι απαραίτητα με

ενδιάμεσες περιόδους, όπως γίνεται στον διασταυρούμενο σχεδιασμό. Η αξιολόγηση γίνεται ενδοατομικά, δηλαδή κάθε συμμετέχοντας συγκρίνεται με τον εαυτό του σε διαφορετικές φάσεις της δοκιμής. Κατά το συγκεκριμένο μοντέλο, μειώνεται η διακύμανση των αποτελεσμάτων και απαιτείται μικρότερο δείγμα για την ανίχνευση διαφορών.

Κάθε μοντέλο προσφέρει διαφορετικά πλεονεκτήματα ανάλογα με τους στόχους της μελέτης, τον τύπο της παρέμβασης και τις ανάγκες σε μέγεθος δείγματος. Η επιλογή του κατάλληλου μοντέλου επιτρέπει την ακριβέστερη εκτίμηση της αποτελεσματικότητας και της ασφάλειας των παρεμβάσεων.



Σχήμα 2

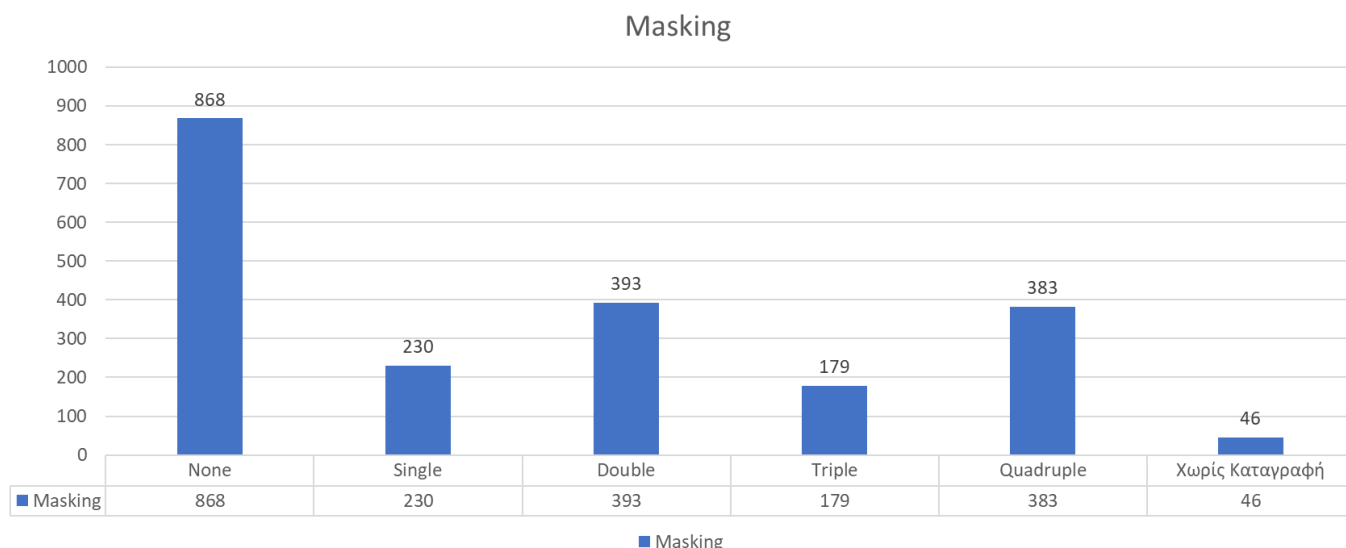
ΤΥΦΛΟΠΟΙΗΣΗ, ΤΥΧΑΙΟΠΟΙΗΣΗ ΚΑΙ ΔΗΜΟΓΡΑΦΙΚΑ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ

Η τυφλοποίηση (Masking) είναι μια διαδικασία κατά την οποία ένα ή περισσότερα μέλη σε μια δοκιμή δεν γνωρίζουν σε ποια θεραπευτικά σκέλη έχουν κατανεμηθεί οι συμμετέχοντες. Με άλλα λόγια, δεν γνωρίζουν ποια θεραπεία έλαβαν. Η τυφλοποίηση είναι μια σημαντική πτυχή κάθε κλινικής δοκιμής που πραγματοποιείται, προκειμένου να αποφευχθεί και να προληφθεί

η μεροληψία κατά τον σχεδιασμό και την εκτέλεση μιας κλινικής δοκιμής. Η τυφλοποίηση διακρίνεται σε πέντε βασικούς τύπους: καθόλου τυφλοποίηση (None), μονή τυφλοποίηση (Single-blind), διπλή τυφλοποίηση (Double-blind), τριπλή τυφλοποίηση (Triple-blind) και τετραπλή τυφλοποίηση (Quadruple-blind).

Στην καθόλου τυφλοποίηση όλοι οι εμπλεκόμενοι γνωρίζουν ποια θεραπεία λαμβάνουν οι συμμετέχοντες, ενώ στη μονή τυφλοποίηση μόνο ο ασθενής αγνοεί τη θεραπεία που του χορηγείται. Στη διπλή τυφλοποίηση, ούτε ο ασθενής ούτε ο γιατρός γνωρίζουν τη θεραπεία, μειώνοντας περαιτέρω τις πιθανότητες μεροληψίας. Στην τριπλή τυφλοποίηση, προστίθεται η τυφλοποίηση του αναλυτή των δεδομένων, ενώ στην τετραπλή τυφλοποίηση και ο χορηγός της μελέτης παραμένει εν αγνοία της κατανομής των θεραπειών. Η επιλογή του τύπου τυφλοποίησης εξαρτάται συχνά από τη φύση της θεραπείας, την πρακτική δυνατότητα εφαρμογής και τους στόχους της δοκιμής.

Στην ανάλυση των 2099 κλινικών δοκιμών που εξετάστηκαν, διαπιστώθηκε ότι σε 868 δοκιμές δεν εφαρμόζεται καμία μορφή τυφλοποίησης, γεγονός που υποδηλώνει ότι μεγάλο ποσοστό δοκιμών βασίζεται σε ανοιχτή κατανομή θεραπειών. Σε 230 δοκιμές εφαρμόστηκε μονή τυφλοποίηση, σε 393 διπλή τυφλοποίηση, σε 179 τριπλή τυφλοποίηση και σε 383 τετραπλή τυφλοποίηση, ενώ σε 46 δοκιμές δεν αναφέρεται καμία μορφή τυφλοποίησης. Τα δεδομένα παρουσιάζονται στο Σχήμα 3.



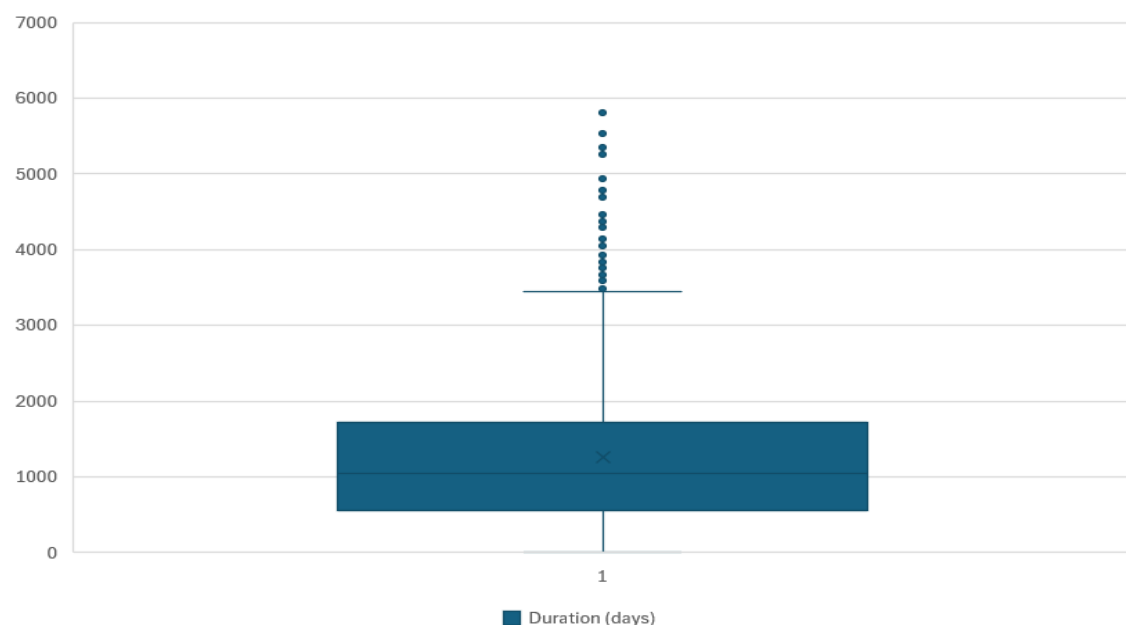
Σχήμα 3

Μία επιπλέον βασική μεθοδολογική πρακτική στις κλινικές δοκιμές είναι η τυχαιοποίηση. Η τυχαιοποίηση είναι μια διαδικασία κατά την οποία οι συμμετέχοντες κατανέμονται τυχαία σε διαφορετικές ομάδες, συνήθως σε μια ομάδα που λαμβάνει το πειραματικό φάρμακο και μια ομάδα που λαμβάνει εικονικό φάρμακο (placebo) ή κάποια άλλη θεραπεία. Αυτό γίνεται για να εξασφαλιστεί ότι οι ομάδες είναι όσο το δυνατόν πιο ισοδύναμες στην αρχή της δοκιμής, μειώνοντας έτσι για ακόμη μια φορά τις πιθανότητες μεροληψίας και διασφαλίζοντας ότι οποιεσδήποτε διαφορές στα αποτελέσματα οφείλονται κυρίως στη θεραπεία και όχι σε άλλους παράγοντες. Έτσι εξασφαλίζεται ότι οι παράγοντες που δεν σχετίζονται με τη θεραπεία, όπως η ηλικία, το φύλο, η σοβαρότητα της νόσου κ.λπ., κατανέμονται εξίσου στις ομάδες. Από τις ολοκληρωμένες 2099 κλινικές δοκιμές, οι 1676 ήταν τυχαιοποιημένες, 155 μη τυχαιοποιημένες και για 268 δεν είχε καταγραφεί ο τρόπος κατανομής.

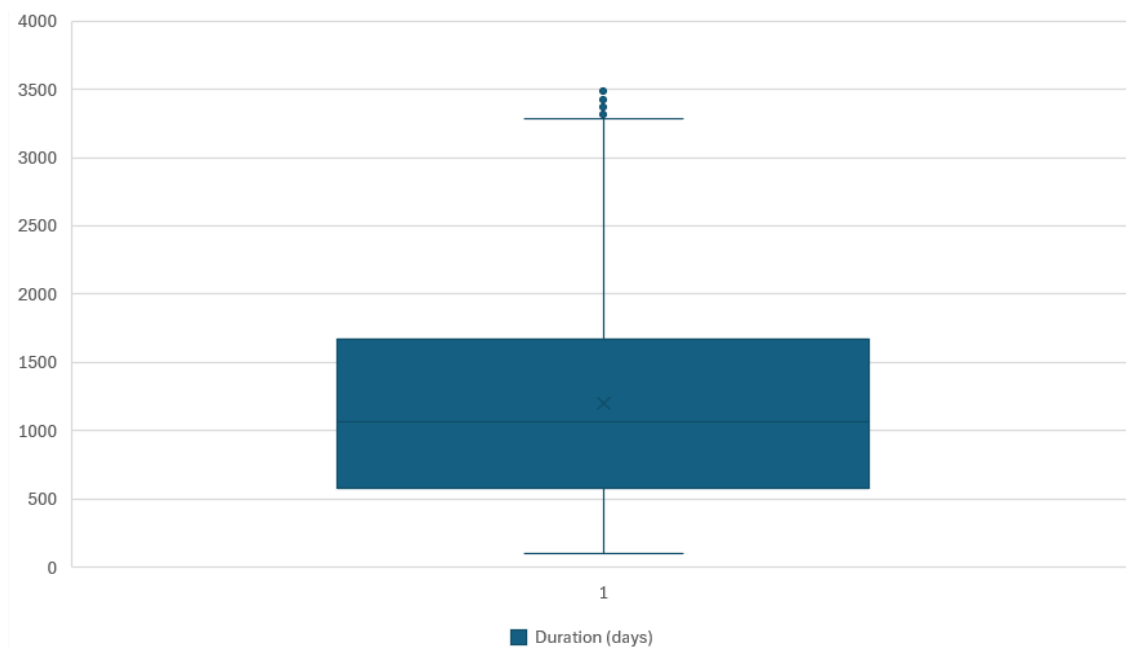
Όσον αφορά τα δημογραφικά χαρακτηριστικά των συμμετεχόντων, στις περισσότερες κλινικές δοκιμές (1831) συμμετείχαν άτομα και των δύο φύλων, σε 149 δοκιμές συμμετείχαν μόνο γυναίκες, ενώ σε 109 συμμετείχαν μόνο άνδρες. Σε 10 δοκιμές δεν αναφερόταν το φύλο των συμμετεχόντων. Ο μέσος αριθμός συμμετεχόντων σε όλες τις δοκιμές ήταν 833 άτομα, με εύρος από 1 έως 49001 συμμετέχοντες και διάμεση τιμή 230 άτομα, υποδηλώνοντας μεγάλη

διακύμανση στον αριθμό των συμμετεχόντων και κατ' επέκταση στη στατιστική ισχύ των μελετών.

Η διάρκεια των κλινικών δοκιμών, παρουσίασε μέση τιμή 1460 ημέρες, δηλαδή περίπου τέσσερα έτη, με εύρος από 91 έως 4883 ημέρες. Η μεγάλη αυτή διακύμανση αντικατοπτρίζει τις διαφορετικές απαιτήσεις παρακολούθησης. Η διακύμανση της διάρκειας όλων συνολικά των κλινικών παρουσιάζεται σε μορφή Box plot στο Σχήμα 4 ενώ η ίδια διακύμανση αφαιρώντας τις ακραίες τιμές (μεγαλύτερες του 3500 και μικρότερες του 100) παρουσιάζονται στο Σχήμα 5.



Σχήμα 4



Σχήμα 5

ΜΕΓΕΘΟΣ ΔΕΙΓΜΑΤΟΣ

Το μέγεθος δείγματος ορίζεται ως ο αριθμός των ατόμων που συμμετέχουν σε μια κλινική δοκιμή, με σκοπό τη συλλογή επαρκών δεδομένων ώστε να εξαχθούν αξιόπιστα και γενικεύσιμα επιστημονικά συμπεράσματα. Αποτελεί βασική παράμετρο στον σχεδιασμό κάθε μελέτης, καθώς σχετίζεται άμεσα με την ποιότητα των αποτελεσμάτων.

Το μέγεθος δείγματος αποτελεί έναν από τους σημαντικότερους παράγοντες στον σχεδιασμό μιας κλινικής δοκιμής, καθώς επηρεάζει άμεσα την αξιοπιστία και την εγκυρότητα των αποτελεσμάτων. Η επιλογή του δεν γίνεται αυθαίρετα, αλλά στηρίζεται σε θεωρητικά και στατιστικά κριτήρια, με σκοπό να εξασφαλιστεί ότι η μελέτη θα έχει επαρκή ισχύ ώστε να εντοπίσει κλινικά σημαντικές διαφορές μεταξύ των υπό εξέταση ομάδων. Ένα πολύ μικρό δείγμα ενέχει τον κίνδυνο λανθασμένων συμπερασμάτων, καθώς αυξάνει την πιθανότητα σφάλματος τύπου II (να μη βρεθεί διαφορά ενώ υπάρχει). Από την άλλη, ένα υπερβολικά μεγάλο δείγμα μπορεί να οδηγήσει σε ανίχνευση μη ουσιαστικών διαφορών, επιβαρύνοντας άσκοπα τόσο τους συμμετέχοντες όσο και τους διαθέσιμους πόρους.

Ο καθορισμός του μεγέθους δείγματος εξαρτάται από πολλούς παράγοντες. Μεταξύ αυτών συγκαταλέγονται το αναμενόμενο μέγεθος του αποτελέσματος, η μεταβλητότητα των δεδομένων, το επιλεγμένο επίπεδο στατιστικής σημαντικότητας (α) και η επιθυμητή στατιστική ισχύς (power), που συνήθως ορίζεται στο 80% ή 90%. Επίσης, σημαντικό ρόλο παίζει η ετερογένεια του πληθυσμού και ο βαθμός ακρίβειας που απαιτείται για τα αποτελέσματα. Σε πρακτικό επίπεδο, λαμβάνονται υπόψη και οι πιθανές απώλειες ή αποχωρήσεις συμμετεχόντων (dropouts), ώστε να υπάρχει επάρκεια δεδομένων έως το τέλος της μελέτης. Έτσι, το μέγεθος δείγματος αποτελεί προϊόν ισορροπίας μεταξύ επιστημονικών αναγκών και ρεαλιστικών περιορισμών.

ΤΥΠΟΙ ΥΠΟΛΟΓΙΣΜΟΥ ΜΕΓΕΘΟΥΣ ΔΕΙΓΜΑΤΟΣ

Η ορθή εκτίμηση του μεγέθους δείγματος αποτελεί θεμέλιο λίθο για τον σχεδιασμό και την αξιολόγηση των κλινικών δοκιμών. Ο υπολογισμός του δείγματος πρέπει να γίνεται με ακρίβεια, λαμβάνοντας υπόψη στατιστικές παραμέτρους, κλινικές υποθέσεις και ηθικούς περιορισμούς, ώστε να εξασφαλίζεται η εγκυρότητα των αποτελεσμάτων χωρίς να τίθεται σε κίνδυνο περισσότερος πληθυσμός από όσο είναι απαραίτητο. Υπό αυτό το πρίσμα, η επιστημονική κοινότητα έχει αναπτύξει ποικίλες στατιστικές μεθόδους και μοντέλα, τα οποία χρησιμοποιούνται για την τεκμηριωμένη εκτίμηση του κατάλληλου αριθμού συμμετεχόντων ανάλογα με τον τύπο και τα χαρακτηριστικά της μελέτης.

Μεταξύ αυτών των μεθόδων περιλαμβάνονται έλεγχοι, όπως το Log-Rank Test, το Wilcoxon-Mann-Whitney Test, το Chi-square Test, το Fisher's Exact Test, το Independent Samples t-test, καθώς και πιο σύνθετες στατιστικές μεθόδους, όπως το Cox Proportional Hazards Regression Model, το Negative Binomial Regression Model, και προσεγγίσεις βασισμένες στη Bayesian στατιστική. Επιπλέον, λογισμικά όπως το nQuery Advisor και κανόνες, όπως αυτός του Peduzzi et al. (1996) προσφέρουν πρακτικά εργαλεία στον υπολογισμό του δείγματος.

Στο πλαίσιο της παρούσας εργασίας, πραγματοποιήθηκε συστηματική έρευνα σε όλα τα διαθέσιμα πρωτόκολλα κλινικών δοκιμών που διεξήχθησαν στην Ελλάδα και για τα οποία υπήρχε δημοσίως προσβάσιμο υλικό. Σε κάθε πρωτόκολλο εξετάστηκε εάν υπήρχε σαφής αναφορά στον τρόπο υπολογισμού του μεγέθους δείγματος. Όπου αυτή η πληροφορία ήταν διαθέσιμη, καταγράφηκε ο συγκεκριμένος στατιστικός τύπος ή μεθοδολογία που χρησιμοποιήθηκε. Το υλικό αυτό αναλύεται και παρουσιάζεται αναλυτικότερα παρακάτω. Πιο συγκεκριμένα, ονομαστικά όλα τα εργαλεία που χρησιμοποιήθηκαν για τον υπολογισμό του απαραίτητου μεγέθους δείγματος στις Ελληνικές κλινικές δοκιμές είναι τα εξής:

1. ANOVA – ANCOVA
2. Gray's test
3. Two-sample t-test

4. Wilcoxon-Mann-Whitney test
5. Chi-square test
6. Bayesian Hierarchical Model
7. MCP-Mod – Multiple Comparison Procedure and Modeling
8. Cox proportional hazards regression model
9. Log-Rank Test
10. Negative Binomial Regression Model
11. CMH test (Cochran-Mantel-Haenszel test)
12. Fisher's Exact Test
13. Peduzzi et al. (1996) rule
14. nQuery Advisor

1. ANOVA-ANCOVA

Η ANOVA (Ανάλυση Διακύμανσης ή Ανάλυση Διασποράς) είναι μια στατιστική τεχνική που χρησιμοποιείται για τον έλεγχο της ισότητας των μέσων τιμών σε πολλαπλές ομάδες ή επίπεδα. Χρησιμοποιείται για να προσδιοριστεί εάν υπάρχουν σημαντικές διαφορές μεταξύ των μέσων τιμών περισσότερων των δύο ομάδων. Η ANCOVA (Ανάλυση Συνδιακύμανσης) είναι μια στατιστική τεχνική που χρησιμοποιείται για την ανάλυση της σχέσης μεταξύ μιας εξαρτημένης μεταβλητής και μίας ή περισσότερων ανεξάρτητων μεταβλητών, προσαρμόζοντας παράλληλα τις επιδράσεις μίας ή περισσότερων συμμεταβλητών. Μπορεί να θεωρηθεί ως ένας συνδυασμός ANOVA και ανάλυσης παλινδρόμησης, καθώς επιτρέπει τον έλεγχο των επιδράσεων άλλων μεταβλητών. Με λίγα λόγια, η ANOVA είναι μια απλούστερη μέθοδος και χρησιμοποιείται όταν υπάρχουν μόνο δύο ή περισσότερες ομάδες για σύγκριση, ενώ η ANCOVA είναι πιο σύνθετη και χρησιμοποιείται όταν υπάρχουν περισσότερες από δύο ομάδες για σύγκριση.

Η χρήση της ANOVA και της ANCOVA δεν περιορίζεται μόνο στην ανάλυση των δεδομένων αλλά μπορεί να προηγηθεί της μελέτης, βοηθώντας στον υπολογισμό του κατάλληλου μεγέθους δείγματος που απαιτείται ώστε η μελέτη να έχει επαρκή στατιστική ισχύ.

Για να υπολογιστεί το μέγεθος δείγματος σε ένα σχέδιο που θα αναλυθεί με ANOVA, απαιτούνται ορισμένες βασικές παράμετροι: ο αριθμός των ομάδων που θα συγκριθούν, το επίπεδο σημαντικότητας (συνήθως 0,05), η επιθυμητή στατιστική ισχύς (συνήθως 80% ή 90%), καθώς και το μέγεθος επίδρασης (effect size), που αντανάκλα τη δύναμη της διαφοράς μεταξύ των ομάδων). Εάν υπάρχει διαθέσιμη πληροφορία για τη διακύμανση, αυτή ενσωματώνεται στον υπολογισμό για μεγαλύτερη ακρίβεια. Στην περίπτωση της ANCOVA, ο υπολογισμός του μεγέθους δείγματος απαιτεί τις ίδιες βασικές παραμέτρους, όπως και η ANOVA, αλλά επιπλέον λαμβάνει υπόψη την ύπαρξη των συγκατανεμημένων μεταβλητών. Η παρουσία συμμεταβλητών μειώνει την ανεκμήνευτη διακύμανση, γεγονός που αυξάνει την ακρίβεια των συγκρίσεων και συνεπώς μειώνει τον απαιτούμενο αριθμό συμμετεχόντων. Η επιλογή μεταξύ των δύο εξαρτάται από το εάν υπάρχουν ή όχι συμμεταβλητές που επηρεάζουν την εξαρτημένη μεταβλητή, καθώς και από τη διαθεσιμότητα των σχετικών πληροφοριών για την παραμετροποίηση του μοντέλου.

2.Gray's test

Το τεστ του Gray, το οποίο εισήχθη από τον Gray το 1988, είναι ένα στατιστικό τεστ που χρησιμοποιείται όταν θέλουμε να συγκρίνουμε περισσότερες από δύο ομάδες (π.χ. ομάδες ασθενών) σε μελέτες όπου έχουμε ανταγωνιστικά συμβάντα (competing risks) (Gray, 1988). Δηλαδή, όταν μπορεί να συμβεί το συμβάν που μας ενδιαφέρει, αλλά και άλλα συμβάντα που εμποδίζουν την εμφάνιση του πρώτου. Η βασική ιδέα είναι ότι το τεστ υπολογίζει, για κάθε ομάδα, έναν «σταθμισμένο μέσο όρο» του ρυθμού με τον οποίο εμφανίζεται το συμβάν ενδιαφέροντος. Στη συνέχεια, συγκρίνει τις ομάδες χρησιμοποιώντας μια μαθηματική συνάρτηση (η στατιστική QGray), η οποία, εάν δεν υπάρχουν πραγματικές διαφορές μεταξύ των ομάδων, ακολουθεί περίπου την κατανομή χ^2 . Έτσι, μπορούμε να υπολογίσουμε την πιθανότητα οι διαφορές που παρατηρούμε να οφείλονται στην τύχη.

Η χρήση του Gray's test στον υπολογισμό του απαιτούμενου μεγέθους δείγματος βασίζεται στη στατιστική του ισοδυναμίας, υπό ασυμπτωτικές συνθήκες, στο πλαίσιο του μοντέλου αναλογικών υποκινδύνων Fine–Gray. Το

μοντέλο Fine–Gray (Fine and Gray, 1999) είναι παραμετρικό/ημιπαραμετρικό μοντέλο αναλογικών υποκινδύνων (proportional subdistribution hazards model) και αποτελεί ένα Gray’s test χωρίς συνδιακυμάνσεις αλλά με ειδική επιλογή βαρών. Για τον υπολογισμό ποσοτήτων όπως αυτό του μεγέθους δείγματος, οι ερευνητές συνήθως βασίζονται στους τύπους ισχύος του Fine–Gray, καθώς αυτοί περιγράφουν πλήρως τη στατιστική συμπεριφορά του τεστ Gray σε μεγάλα δείγματα. Δεδομένου, λοιπόν, ότι το τεστ του Gray αποτελεί ειδική περίπτωση αυτού του ελέγχου, η μεθοδολογία υπολογισμού ισχύος και μεγέθους δείγματος μπορεί να προσαρμοστεί απευθείας από τις γνωστές προσεγγίσεις για το μοντέλο Fine–Gray.

Στη διαδικασία σχεδιασμού, ο ερευνητής καθορίζει εκ των προτέρων τις τιμές των βασικών παραμέτρων, όπως το επίπεδο σημαντικότητας, την επιθυμητή ισχύ ($1-\beta$), τον στόχο μεγέθους επίδρασης, την αναλογία κατανομής μεταξύ των ομάδων και τον χρονικό ορίζοντα, στον οποίο θα εκτιμηθεί η επίπτωση του συμβάντος ενδιαφέροντος. Κατά το μοντέλο Fine–Gray, ο απαιτούμενος αριθμός συμβάντων προκύπτει από την εξίσωση ισχύος του ελέγχου τύπου Wald ή Score, όπου η μη κεντρική παράμετρος σχετίζεται γραμμικά με τον αριθμό των συμβάντων και τετραγωνικά με τον δείκτη SHR (Sub-Hazard Ratio), ο οποίος εκφράζει πόσο αυξάνεται ή μειώνεται η πιθανότητα να συμβεί το γεγονός που μελετάμε, όταν υπάρχει και η πιθανότητα να συμβούν άλλα ανταγωνιστικά γεγονότα. Ο συνολικός απαιτούμενος αριθμός συμμετεχόντων προκύπτει κατόπιν διαίρεσης του αριθμού συμβάντων με το αναμενόμενο ποσοστό εμφάνισης του συμβάντος ενδιαφέροντος.

3. Two-sample t-test

Το Two-sample t-test, γνωστό και ως independent samples t-test, αποτελεί μια από τις βασικότερες παραμετρικές μεθόδους για τη σύγκριση των μέσων τιμών δύο ανεξάρτητων ομάδων. Ουσιαστικά ελέγχει εάν οι μέσοι όροι δύο πληθυσμών είναι ίσοι ή όχι. Το τεστ αυτό χρησιμοποιείται ευρέως σε ερευνητικά πρωτόκολλα όπου το ζητούμενο είναι να διαπιστωθεί εάν μία παρέμβαση έχει στατιστικώς σημαντική επίδραση συγκριτικά με μια ομάδα ελέγχου, μετρώντας ένα ποσοτικό τελικό σημείο (όπως π.χ. αρτηριακή πίεση,

επίπεδα σακχάρου, σκορ λειτουργικότητας). Για να είναι έγκυρα τα αποτελέσματα ενός t-test δύο δειγμάτων προϋποθέτει ότι οι παρατηρήσεις στο ένα δείγμα θα είναι ανεξάρτητες από τις παρατηρήσεις στο άλλο δείγμα, τα δεδομένα θα είναι περίπου κανονικά κατανεμημένα και τα δύο δείγματα θα πρέπει να έχουν περίπου την ίδια διακύμανση..

Ο έλεγχος Two-sample t-test, ως παραμετρική μέθοδος σύγκρισης των μέσων τιμών δύο ανεξάρτητων ομάδων, παρέχει το στατιστικό πλαίσιο για τον υπολογισμό του απαιτούμενου μεγέθους δείγματος σε μελέτες με ποσοτικό τελικό σημείο. Η μαθηματική του διατύπωση συνδέει την αναμενόμενη διαφορά των μέσων τιμών, τη διακύμανση των δεδομένων και τον αριθμό παρατηρήσεων σε κάθε ομάδα με το προκαθορισμένο επίπεδο σημαντικότητας και την επιθυμητή στατιστική ισχύ, επιτρέποντας την εκτίμηση του ελάχιστου αριθμού συμμετεχόντων που απαιτείται ώστε να ανιχνευθεί μια διαφορά με την καθορισμένη πιθανότητα. Η διαδικασία βασίζεται στην ποσοτικοποίηση τεσσάρων βασικών παραμέτρων όπως της στατιστικής σημαντικότητας (συνήθως $\alpha=0,05$), της επιθυμητής ισχύος (π.χ. 80% ή 90%), της εκτιμώμενης διαφοράς μεταξύ των μέσων όρων των δύο ομάδων (effect size) και της τυπικής απόκλισης της μετρούμενης μεταβλητής.

4.Wilcoxon-Mann-Whitney test

Το Wilcoxon-Mann-Whitney test, επίσης γνωστό ως Mann-Whitney U Test, είναι ένας μη παραμετρικός στατιστικός έλεγχος που χρησιμοποιείται για τη σύγκριση δύο δειγμάτων ή ομάδων. Αξιολογεί εάν δύο ομάδες δείγματος είναι πιθανό να προέρχονται από τον ίδιο πληθυσμό και ουσιαστικά απαντά στο ερώτημα: έχουν αυτοί οι δύο πληθυσμοί το ίδιο σχήμα (την ίδια μορφή) ως προς τα δεδομένα τους; Με άλλα λόγια, θέλουμε στοιχεία για το εάν οι ομάδες προέρχονται από πληθυσμούς με διαφορετικά επίπεδα μιας μεταβλητής ενδιαφέροντος. Συνεπώς, οι υποθέσεις σε ένα Mann-Whitney U Test είναι: Η μηδενική υπόθεση (H_0) είναι ότι οι δύο πληθυσμοί είναι ίσοι και η εναλλακτική υπόθεση (H_1) είναι ότι οι δύο πληθυσμοί δεν είναι ίσοι. Το Mann-Whitney U Test είναι ένας κοινός στατιστικός έλεγχος που χρησιμοποιείται σε πολλούς τομείς, όπως οι βιολογικές επιστήμες και η επιδημιολογία. Είναι ιδιαίτερα χρήσιμος

όταν αξιολογούμε τη διαφορά μεταξύ δύο ανεξάρτητων ομάδων με χαμηλό αριθμό ατόμων σε κάθε ομάδα (συνήθως λιγότερα από 30), οι οποίες δεν κατανέμονται κανονικά και όπου τα δεδομένα είναι συνεχή.

Ο υπολογισμός του μεγέθους δείγματος με τη χρήση του Wilcoxon-Mann-Whitney test βασίζεται στην εκτίμηση της κατανομής της μεταβλητής ενδιαφέροντος και στη διαφορά που αναμένεται μεταξύ δύο ανεξάρτητων ομάδων. Το τεστ αυτό χρησιμοποιείται για να συγκρίνει την κατανομή ενός χαρακτηριστικού μεταξύ των ομάδων, χωρίς να απαιτείται η υπόθεση κανονικότητας. Για τον υπολογισμό του μεγέθους δείγματος, λαμβάνονται υπόψη οι εκτιμώμενες κατανομές της μεταβλητής στις δύο ομάδες, το επίπεδο σημαντικότητας (α) και η επιθυμητή ισχύς (power) της μελέτης.

5. Chi-square test

Ένα Chi-square test (έλεγχος χ^2) είναι ένα στατιστικό τεστ υπόθεσης που χρησιμοποιείται στην ανάλυση πινάκων συνάφειας όταν τα μεγέθη του δείγματος είναι μεγάλα. Με απλά λόγια, αυτό το τεστ χρησιμοποιείται κυρίως για να εξετάσει αν δύο κατηγορικές μεταβλητές είναι ανεξάρτητες, με βάση την τιμή ενός στατιστικού ελέγχου. Ο έλεγχος αυτός είναι αξιόπιστος όταν η τιμή του ακολουθεί την αναμενόμενη κατανομή σύμφωνα με τη μηδενική υπόθεση, όπως συμβαίνει στο τεστ χ^2 του Pearson και στις παραλλαγές του. Το τεστ χ^2 του Pearson χρησιμοποιείται για να προσδιοριστεί εάν υπάρχει στατιστικά σημαντική διαφορά μεταξύ των αναμενόμενων συχνοτήτων και των παρατηρούμενων συχνοτήτων σε μία ή περισσότερες κατηγορίες ενός πίνακα συνάφειας. Εάν η μηδενική υπόθεση δείχνει ότι δεν υπάρχουν διαφορές μεταξύ των κατηγοριών στον πληθυσμό είναι αληθής, το στατιστικό στοιχείο του τεστ που υπολογίζεται από τις παρατηρήσεις ακολουθεί μια χ^2 κατανομή.

Ο υπολογισμός του μεγέθους δείγματος σε κλινικές δοκιμές συχνά στηρίζεται στη χρήση του ελέγχου χ^2 για τη σύγκριση ποσοστών ή αναλογιών μεταξύ δύο ή περισσότερων ανεξάρτητων ομάδων. Για τον σκοπό αυτό, εκτιμώνται οι αναμενόμενες αναλογίες ανταπόκρισης στις διάφορες ομάδες. Στη συνέχεια, καθορίζεται το μέγεθος δείγματος που απαιτείται ώστε να επιτευχθεί προκαθορισμένη στατιστική ισχύς (συνήθως 80% ή 90%) για την ανίχνευση

μιας ελάχιστης σημαντικής διαφοράς μεταξύ των ομάδων, σε συγκεκριμένο επίπεδο σημαντικότητας. Η ορθή χρήση του ελέγχου χ^2 για τον καθορισμό του μεγέθους δείγματος είναι καθοριστική όταν οι βασικές μεταβλητές της μελέτης είναι δίτιμες ή κατηγορικές, καθώς διασφαλίζει την επάρκεια του δείγματος για τη στατιστικά έγκυρη ανίχνευση διαφορών μεταξύ των υπό σύγκριση ομάδων.

6. Bayesian Hierarchical Model

Το μοντέλο Bayesian Hierarchical (Μπεϋζιανή ιεραρχική μοντελοποίηση) αποτελεί ένα στατιστικό πλαίσιο που οργανώνεται σε πολλαπλά επίπεδα και βασίζεται στο θεώρημα του Bayes. Ουσιαστικά, συνδυάζει επιμέρους υπομοντέλα τα οποία ενοποιούνται με τα παρατηρούμενα δεδομένα, επιτρέποντας την εκτίμηση τόσο των παραμέτρων όσο και των υπερπαραμέτρων. Η βασική διαφορά με την κλασσική στατιστική είναι ότι στο Μπεϋζιανό πλαίσιο οι παράμετροι θεωρούνται τυχαίες μεταβλητές, ενώ η κλασσική προσέγγιση τις αντιμετωπίζει ως σταθερές αλλά άγνωστες ποσότητες. Αυτό σημαίνει ότι οι Μπεϋζιανοί αναλυτές μπορούν να ενσωματώσουν υποκειμενικές πληροφορίες ή προϋπάρχουσα γνώση στη διαδικασία. Έτσι, ενώ τα αποτελέσματα των δύο προσεγγίσεων μπορεί να φαίνονται διαφορετικά, στην πραγματικότητα αντανακλούν διαφορετικά ερευνητικά ερωτήματα.

Η χρήση του Bayesian Hierarchical Model στον υπολογισμό μεγέθους δείγματος είναι ιδιαίτερα χρήσιμη σε πολυκεντρικές κλινικές δοκιμές, συχνά σε μελέτες με πολλαπλές χρονικές μετρήσεις. Η δυνατότητα ενσωμάτωσης προγενέστερων πληροφοριών μειώνει την ανάγκη για υπερβολικά μεγάλα δείγματα, καθιστώντας πιο αποδοτική τη σχεδίαση της μελέτης. Παράλληλα, η προσομοιωτική φύση της ανάλυσης επιτρέπει την αξιολόγηση διαφόρων σεναρίων και υποθέσεων, παρέχοντας μια πιο ολοκληρωμένη βάση για τη λήψη αποφάσεων σχετικά με το μέγεθος του δείγματος.

7. MCP-Mod – Multiple Comparison Procedure and Modeling

To MCP-Mod (Multiple Comparison Procedure and Modeling) αποτελεί μια εξελιγμένη στατιστική μεθοδολογία που εφαρμόζεται κυρίως σε πρώιμες φάσεις κλινικών δοκιμών, ιδιαίτερα στη Φάση II, με σκοπό την αξιολόγηση της σχέσης δόσης-απόκρισης. Η μέθοδος αυτή συνδυάζει την ισχύ των διαδικασιών πολλαπλών συγκρίσεων με την ευελιξία της μοντελοποίησης, επιτρέποντας την ταυτόχρονη αξιολόγηση διαφορετικών υποθετικών μορφών της καμπύλης δόσης-απόκρισης. Αντί να προϋποθέτει την επιλογή ενός συγκεκριμένου μοντέλου, το MCP-Mod δίνει τη δυνατότητα εξέτασης πολλαπλών εναλλακτικών, όπως γραμμικά, λογαριθμικά ή λογιστικά μοντέλα, προκειμένου να διαπιστωθεί ποιο μοντέλο περιγράφει καλύτερα τη σχέση μεταξύ της χορηγούμενης δόσης και της θεραπευτικής απόκρισης. Το πλεονέκτημα αυτής της προσέγγισης είναι ότι αυξάνει τη στατιστική ισχύ και την αξιοπιστία των αποτελεσμάτων, μειώνοντας ταυτόχρονα τον κίνδυνο αποτυχίας λόγω λανθασμένων αρχικών υποθέσεων για τη μορφή της σχέσης δόσης-απόκρισης.

Ο υπολογισμός του μεγέθους δείγματος με τη χρήση του MCP-Mod δεν βασίζεται σε έναν απλό αναλυτικό τύπο, όπως συμβαίνει σε πιο κλασικές στατιστικές δοκιμές, αλλά απαιτεί τη διενέργεια εκτεταμένων προσομοιώσεων. Οι προσομοιώσεις αυτές περιλαμβάνουν την καθορισμένη επιλογή υποψήφιων μοντέλων δόσης-απόκρισης, την εκτίμηση του προσδοκώμενου μεγέθους επίδρασης, τη διασπορά των μετρήσεων, τον αριθμό των βραχιόνων δόσης που θα συγκριθούν, το επιθυμητό επίπεδο στατιστικής σημαντικότητας (συνήθως 0,05) και την επιδιωκόμενη στατιστική ισχύ. Κατά τη διαδικασία, δημιουργούνται χιλιάδες εικονικά σενάρια βάσει των παραπάνω παραμέτρων, προκειμένου να εκτιμηθεί η πιθανότητα να ανιχνευθεί με ακρίβεια μία στατιστικά σημαντική σχέση μεταξύ δόσης και απόκρισης. Το μέγεθος του δείγματος προσαρμόζεται έως ότου επιτευχθεί ένα αποδεκτό ποσοστό επιτυχίας στην ταυτοποίηση αυτής της σχέσης και στον εντοπισμό της βέλτιστης δόσης.

8.Cox proportional hazards regression model

To Cox proportional hazards regression model, γνωστό και ως μοντέλο αναλογικού κινδύνου του Cox, αποτελεί μία από τις σημαντικότερες μεθόδους

ανάλυσης επιβίωσης στη βιοστατιστική και την ιατρική έρευνα. Αναπτύχθηκε από τον David Cox το 1972 και χρησιμοποιείται για τη μελέτη της επίδρασης διαφόρων παραγόντων (μεταβλητών) στον χρόνο έως την εμφάνιση ενός γεγονότος, όπως ο θάνατος, η υποτροπή της νόσου ή οποιοδήποτε άλλο καθορισμένο τελικό σημείο (endpoint). Το κύριο χαρακτηριστικό του μοντέλου είναι ότι δεν απαιτεί την υπόθεση συγκεκριμένης κατανομής για τον χρόνο επιβίωσης (είναι δηλαδή ημιπαραμετρικό), ενώ βασίζεται στην υπόθεση της αναλογικότητας των κινδύνων, δηλαδή ο λόγος των κινδύνων (hazard ratios) μεταξύ ομάδων παραμένει σταθερός με την πάροδο του χρόνου.

Η χρήση του Cox proportional hazards model για τον υπολογισμό του απαιτούμενου μεγέθους δείγματος σε μια κλινική δοκιμή που μελετά δεδομένα επιβίωσης βασίζεται στην επιθυμία να εντοπιστεί μια στατιστικά σημαντική διαφορά στους κινδύνους μεταξύ των υπό σύγκριση ομάδων. Ο σχεδιασμός της μελέτης απαιτεί αρχικά τον καθορισμό βασικών παραμέτρων όπως η αναμενόμενη αναλογία κινδύνων (hazard ratio), το ποσοστό γεγονότων (δηλαδή ο αριθμός των συμμετεχόντων που αναμένεται να βιώσουν το συμβάν εντός της περιόδου παρακολούθησης), ο συνολικός χρόνος παρακολούθησης, το επίπεδο σημαντικότητας και η επιδιωκόμενη στατιστική ισχύς.

9.Log-Rank test

Το Log-Rank test είναι ένα στατιστικό τεστ που χρησιμοποιείται για τη σύγκριση καμπυλών επιβίωσης μεταξύ δύο ή περισσότερων ομάδων. Βασίζεται στην ανάλυση επιβίωσης και ειδικότερα στη μέθοδο Kaplan–Meier, η οποία εκτιμά την πιθανότητα επιβίωσης σε διαφορετικά χρονικά σημεία. Το Log-Rank test εξετάζει εάν οι διαφορές που παρατηρούνται στις καμπύλες επιβίωσης μεταξύ των ομάδων είναι στατιστικά σημαντικές, δηλαδή εάν ο χρόνος μέχρι το γεγονός (π.χ. θάνατος, υποτροπή νόσου) διαφέρει ουσιαστικά μεταξύ τους. Το Log-Rank test χρησιμοποιείται ευρέως σε κλινικές μελέτες, όπου το ενδιαφέρον επικεντρώνεται στον χρόνο μέχρι την εμφάνιση ενός γεγονότος. Για παράδειγμα, μπορεί να αξιολογήσει εάν μία νέα θεραπεία παρατείνει την επιβίωση σε σύγκριση με την τυπική θεραπεία. Ωστόσο, κάνει την υπόθεση ότι οι λόγοι κινδύνου (hazard ratios) παραμένουν σταθεροί με την πάροδο του

χρόνου, κάτι που σημαίνει ότι εάν οι καμπύλες επιβίωσης διασταυρώνονται έντονα, το τεστ μπορεί να δώσει παραπλανητικά αποτελέσματα. Παρ' όλα αυτά, παραμένει το πιο συχνά χρησιμοποιούμενο τεστ για τη σύγκριση επιβίωσης σε κλινικές δοκιμές.

Ο υπολογισμός του μεγέθους δείγματος με τη βοήθεια του Log-Rank test βασίζεται στη σύγκριση των καμπυλών επιβίωσης ή χρόνου έως το γεγονός μεταξύ δύο ή περισσότερων ομάδων. Αρχικά καθορίζεται η υπόθεση που θα ελεγχθεί, ο δείκτης κινδύνου (hazard ratio) και τα επίπεδα σφάλματος τύπου I και ισχύος που απαιτούνται. Στη συνέχεια προσδιορίζεται ο αριθμός των συμβάντων που χρειάζονται για να επιτευχθεί η επιθυμητή ισχύς υπό τις παραδοχές αυτές. Τέλος, λαμβάνοντας υπόψη την αναμενόμενη συχνότητα των συμβάντων, τον ρυθμό στρατολόγησης των συμμετεχόντων, τη διάρκεια της μελέτης και τυχόν απώλειες, υπολογίζεται ο συνολικός αριθμός συμμετεχόντων που πρέπει να τυχαιοποιηθούν, ώστε να αναμένονται τα απαραίτητα γεγονότα και το Log-Rank test να έχει την προβλεπόμενη ισχύ.

10.Negative Binomial Regression Model

Το μοντέλο αρνητικής διωνυμικής παλινδρόμησης (Negative Binomial Regression) είναι μια μέθοδος στατιστικής ανάλυσης που χρησιμοποιείται για να μοντελοποιήσει δεδομένα, όπου η εξαρτημένη μεταβλητή είναι αριθμός γεγονότων ή μετρήσεων που μπορούν να λάβουν μη αρνητικές ακέραιες τιμές (0, 1, 2, 3, κ.λπ.). Το μοντέλο αυτό επιτρέπει μεγαλύτερη διακύμανση από τη μέση τιμή, αντιμετωπίζοντας έτσι το φαινόμενο της υπερδιασποράς (overdispersion). Η παραδοσιακή εκδοχή του μοντέλου, γνωστή ως NB2, βασίζεται σε ένα μείγμα κατανομών Poisson-Γάμμα, που σημαίνει ότι η αρχική Poisson κατανομή «μικραίνει» ή «μεγαλώνει» ανάλογα με μια τυχαία μεταβλητή που ακολουθεί Γάμμα κατανομή. Αυτό επιτρέπει στο μοντέλο να μοντελοποιεί την ετερογένεια των δεδομένων, δηλαδή τις διαφορές στις πιθανότητες εμφάνισης των γεγονότων μεταξύ των παρατηρήσεων.

Ο υπολογισμός του μεγέθους δείγματος σε κλινικές δοκιμές βασίζεται σε μια προσέγγιση που λαμβάνει υπόψη τον τύπο των γεγονότων που μετρώνται ανά συμμετέχοντα, όπως υποτροπές ή ασθένεια. Αρχικά, ορίζονται βασικές

παράμετροι, όπως ο αναμενόμενος ρυθμός γεγονότων στην ομάδα ελέγχου, η στοχευόμενη σχετική διαφορά στη θεραπευτική ομάδα και η σχέση τυχαιοποίησης μεταξύ των ομάδων. Στη συνέχεια καθορίζονται η επιθυμητή στατιστική ισχύς και το επίπεδο σημαντικότητας για να εξασφαλιστεί η πιθανότητα ανίχνευσης της αναμενόμενης θεραπευτικής επίδρασης. Αυτή η προσέγγιση επιτρέπει στους ερευνητές να σχεδιάζουν κλινικές δοκιμές με επαρκή αριθμό συμμετεχόντων, λαμβάνοντας υπόψη τόσο τη μεταβλητότητα των δεδομένων όσο και τις πιθανές απώλειες συμμετοχής.

11.CMH test (Cochran-Mantel-Haenszel test)

Στη στατιστική, το τεστ Cochran-Mantel-Haenszel (CMH) είναι ένα τεστ που χρησιμοποιείται στην ανάλυση στρωματοποιημένων ή αντιστοιχισμένων κατηγορικών δεδομένων. Ονομάστηκε έτσι από τους William G. Cochran, Nathan Mantel και William Haenszel. Επιτρέπει σε έναν ερευνητή να ελέγξει τη συσχέτιση μεταξύ ενός δίτιμου προγνωστικού παράγοντα ή θεραπείας και ενός δίτιμου αποτελέσματος, όπως η κατάσταση περίπτωσης ή ελέγχου, λαμβάνοντας παράλληλα υπόψη τη στρωματοποίηση. Το τεστ CMH χειρίζεται αυθαίρετα μεγέθη στρωμάτων. Οι επεκτάσεις αυτού του τεστ σε μια κατηγορική απόκριση ή/και σε πολλές ομάδες ονομάζονται συνήθως στατιστικές Cochran-Mantel-Haenszel. Χρησιμοποιείται συχνά σε παρατηρησιακές μελέτες, στις οποίες η τυχαία ανάθεση υποκειμένων σε διαφορετικές θεραπείες δεν μπορεί να ελεγχθεί, αλλά μπορούν να μετρηθούν συγχυτικές συμμεταβλητές.

Ο υπολογισμός του μεγέθους δείγματος με βάση το τεστ CMH γίνεται με την αντίστροφη χρήση των εξισώσεων ισχύος του τεστ, εισάγοντας τις επιθυμητές τιμές ισχύος, σημαντικότητας, αναμενόμενων αναλογιών και στρωματοποίησης. Με αυτόν τον τρόπο εξασφαλίζεται ότι η μελέτη έχει επαρκές δείγμα για να ανιχνεύσει στατιστικά σημαντική συσχέτιση μεταξύ θεραπείας και αποτελέσματος, λαμβάνοντας υπόψη πιθανές συγχύσεις.

12.Fisher's Exact Test

Το τεστ του Fisher είναι ένα στατιστικό εργαλείο που χρησιμοποιείται για την ανάλυση πινάκων συνάφειας, κυρίως 2×2 , με στόχο να εξεταστεί εάν υπάρχει στατιστικά σημαντική σχέση μεταξύ δύο κατηγοριών. Σε αντίθεση με άλλους στατιστικούς ελέγχους, όπως ο έλεγχος χ^2 , το τεστ του Fisher δεν βασίζεται σε προσεγγίσεις που γίνονται ακριβείς μόνο σε μεγάλα δείγματα. Αντίθετα, υπολογίζει την πιθανότητα των παρατηρούμενων δεδομένων με ακρίβεια, δεδομένης της μηδενικής υπόθεσης ανεξαρτησίας των δύο μεταβλητών. Η μέθοδος αυτή χρησιμοποιεί την υπεργεωμετρική κατανομή για να καθορίσει την ακριβή πιθανότητα των παρατηρούμενων συνδυασμών, με βάση τα σταθερά περιθώρια του πίνακα συνάφειας. Αυτό επιτρέπει την εκτίμηση της στατιστικής σημαντικότητας χωρίς να απαιτείται μεγάλος αριθμός παρατηρήσεων ή συγκεκριμένες παραδοχές για τη μορφή κατανομής των δεδομένων. Ουσιαστικά, το τεστ χρησιμοποιείται κυρίως όταν τα δείγματα είναι μικρά, καθώς άλλες προσεγγιστικές μέθοδοι μπορεί να μην είναι αξιόπιστες σε αυτές τις περιπτώσεις. Ωστόσο, η ισχύς του δεν περιορίζεται από το μέγεθος του δείγματος και μπορεί να εφαρμοστεί σε οποιοδήποτε μέγεθος, κάτι που το καθιστά πολύ ευέλικτο σε διάφορα πειραματικά ή κλινικά πλαίσια.

Ο υπολογισμός ξεκινά με τον καθορισμό της μηδενικής υπόθεσης, κατά την οποία οι πιθανότητες εμφάνισης του γεγονότος ενδιαφέροντος είναι ίδιες στις δύο ομάδες, και της εναλλακτικής υπόθεσης, κατά την οποία οι πιθανότητες διαφέρουν. Η διαδικασία υπολογισμού του μεγέθους δείγματος για το Fisher's Exact Test βασίζεται σε έναν 2×2 πίνακα συνάφειας, όπου οι συμμετέχοντες κατατάσσονται ανάλογα με την ομάδα και την επίτευξη του καταληκτικού σημείου. Με βάση τα εκτιμώμενα ποσοστά, υπολογίζεται ο αριθμός των συμμετεχόντων που απαιτείται ώστε η συνολική πιθανότητα απόρριψης της μηδενικής υπόθεσης να φτάνει ή να υπερβαίνει την επιθυμητή ισχύ.

13. Peduzzi et al. (1996) rule

Ο κανόνας των Peduzzi et al. (1996) αποτελεί μια καθιερωμένη κατευθυντήρια γραμμή στον σχεδιασμό μελετών επιβίωσης και ανάλυσης Cox proportional hazards. Περιγράφει τη σχέση ανάμεσα στον αριθμό των ανεξάρτητων μεταβλητών σε ένα πολυπαραγοντικό μοντέλο και την

απαιτούμενη επάρκεια των δεδομένων για αξιόπιστες εκτιμήσεις. Στόχος του είναι να εξασφαλίσει σταθερότητα στις εκτιμήσεις των παραμέτρων, να μειώσει τον κίνδυνο υπερπροσαρμογής και να διατηρήσει την ακρίβεια των διαστημάτων εμπιστοσύνης. Ο κανόνας επισημαίνει ότι η αναλογία μεταξύ αριθμού γεγονότων και αριθμού μεταβλητών είναι κρίσιμη για την εγκυρότητα των αποτελεσμάτων και την αξιοπιστία των συμπερασμάτων που προκύπτουν από το μοντέλο. Χρησιμοποιείται ευρέως ως πρακτικό εργαλείο σχεδιασμού κλινικών μελετών και αναλύσεων χρόνου έως το υπό μελέτη γεγονός.

Ο κανόνας του Peduzzi et al. (1996) αποτελεί μια καθιερωμένη οδηγία στον σχεδιασμό μελετών που χρησιμοποιούν πολυπαραγοντικά μοντέλα επιβίωσης. Ο κανόνας παρέχει μια πρακτική αναλογία που συνδέει τον αριθμό των γεγονότων με τον αριθμό των μεταβλητών που περιλαμβάνονται στο μοντέλο, υποδεικνύοντας ουσιαστικά ότι κάθε μεταβλητή χρειάζεται έναν επαρκή αριθμό γεγονότων για να μπορεί να εκτιμηθεί με αξιοπιστία. Παρότι αποτελεί έναν συντηρητικό εμπειρικό κανόνα, η μέθοδος του Peduzzi χρησιμοποιείται ιδιαίτερα όταν η εφαρμογή πιο ολοκληρωμένων προσεγγίσεων υπολογισμού του μεγέθους δείγματος, βασισμένων στην ισχύ ή στην κατανομή του καταληκτικού σημείου, δεν είναι εφικτή ή πρακτική. Η μέθοδος παρέχει μια αρχική, ασφαλή εκτίμηση, η οποία μπορεί να καθοδηγήσει αποτελεσματικά τη λήψη αποφάσεων κατά την αρχική φάση σχεδιασμού μιας κλινικής δοκιμής.

14.nQuery Advisor

Το nQuery Advisor είναι ένα εξειδικευμένο στατιστικό λογισμικό που χρησιμοποιείται εκτενώς στον τομέα της βιοστατιστικής και ειδικά στον σχεδιασμό κλινικών δοκιμών. Αναπτύχθηκε για να διευκολύνει την ακριβή και αποτελεσματική εκτίμηση του απαιτούμενου μεγέθους δείγματος και της στατιστικής ισχύος (power) σε αρκετούς τύπους μελετών. Το πρόγραμμα καλύπτει ευρύ φάσμα στατιστικών μοντέλων και μεθόδων, όπως παραμετρικά και μη παραμετρικά τεστ, αναλύσεις επιβίωσης, μοντέλα παλινδρόμησης, αναλύσεις ισοδυναμίας, και πολλά άλλα. Η έκδοση nQuery Advisor 7.0, συγκεκριμένα, αποτέλεσε σημαντική αναβάθμιση του λογισμικού,

ενσωματώνοντας επιπλέον επιλογές και διεπαφές χρήστη για πιο φιλική εμπειρία.

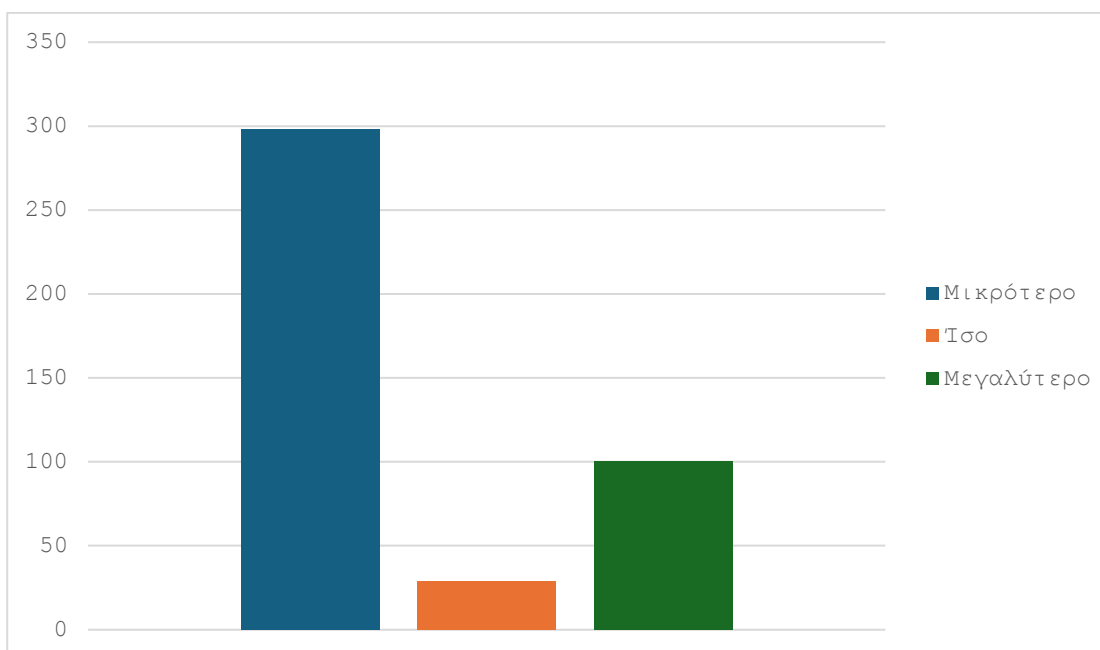
Το nQuery Advisor μπορεί να χρησιμοποιηθεί άμεσα για τον υπολογισμό του μεγέθους δείγματος σε μια κλινική δοκιμή, προσφέροντας προκαθορισμένα πρότυπα (templates) που βασίζονται σε στατιστικά μοντέλα ανάλογα με τον τύπο της υπόθεσης και του σχεδίου της μελέτης. Εισάγονται κρίσιμες παράμετροι της μελέτης, όπως το επίπεδο σημαντικότητας, η στατιστική ισχύς, το προσδοκώμενο μέγεθος επίδρασης και οι εκτιμώμενες διακυμάνσεις. Το πρόγραμμα υπολογίζει το απαιτούμενο μέγεθος δείγματος ανά ομάδα, λαμβάνοντας υπόψη και πρόσθετες μεταβλητές, όπως τον ρυθμό αποχωρήσεων (dropout rate) ή τον τύπο κατανομής των δεδομένων. Επιπλέον, μπορεί να παρέχει γραφήματα και πίνακες ευαισθησίας που βοηθούν στην κατανόηση του πώς οι αλλαγές στις παραμέτρους επηρεάζουν το τελικό δείγμα.

ΣΥΓΚΡΙΣΗ ENROLLMENT ΚΑΙ SAMPLE SIZE (Protocol)

Η ανάλυση της σχέσης ανάμεσα στον δηλωμένο τελικό αριθμό συμμετεχόντων (Enrollment) και το μέγεθος δείγματος που αναφερόταν στο πρωτόκολλο ανέδειξε σημαντικές διαφοροποιήσεις. Από το σύνολο των 454 πρωτοκόλλων, σε 27 (5,9%) δεν υπήρχε διαθέσιμη αναφορά για το απαιτούμενο μέγεθος δείγματος, γεγονός που υποδηλώνει περιορισμούς στην πληρότητα των πρωτοκόλλων.

Στις υπόλοιπες 427 περιπτώσεις, καταγράφηκαν τρία βασικά μοτίβα. Συγκεκριμένα, σε 298 πρωτόκολλα (65,6%) το μέγεθος δείγματος που αναφερόταν στο πρωτόκολλο ήταν μικρότερο από τον τελικό αριθμό συμμετεχόντων, σε 100 πρωτόκολλα (22,1%) το μέγεθος δείγματος που αναφερόταν στο πρωτόκολλο ήταν μεγαλύτερο, ενώ σε 29 πρωτόκολλα (6,4%) τα δύο μεγέθη ταυτίζονταν. Στο σημείο αυτό να επισημανθεί ότι σε περιπτώσεις όπου μέγεθος δείγματος που αναφερόταν στο πρωτόκολλο παρουσιαζόταν ως διάστημα τιμών, ο τελικός αριθμός συμμετεχόντων θεωρήθηκε ισοδύναμος όταν βρισκόταν εντός αυτού του διαστήματος. Το ακόλουθο ραβδόγραμμα του Σχήματος 6 παρουσιάζει με αναλυτικό τρόπο τη μέγιστη διαφοροποίηση που καταγράφηκε μεταξύ των αποτελεσμάτων, προσφέροντας μια σαφή απεικόνιση

της έντασης της απόκλισης. Συγκεκριμένα, παρουσιάζεται σε τι αριθμό των κλινικών δοκιμών, το προβλεπόμενο μέγεθος δείγματος από το πρωτόκολλο τους είναι μικρότερο, ίσο ή μεγαλύτερο από τον πραγματικό αριθμό συμμετεχόντων της μελέτης.



Σχήμα 6

Τα αποτελέσματα δείχνουν ότι σε ένα μεγάλο ποσοστό των μελετών ο τελικός αριθμός συμμετεχόντων ξεπερνά το μέγεθος δείγματος που αναφερόταν στο πρωτόκολλο, πιθανόν ως στρατηγική υπερστρατολόγησης για να αντισταθμιστούν πιθανές αποχωρήσεις συμμετεχόντων ή μη ολοκληρωμένες συμμετοχές. Αντίθετα, σχεδόν στο 1/5 των πρωτοκόλλων όπου το μέγεθος δείγματος που αναφερόταν στο πρωτόκολλο υπερέβαινε τον τελικό αριθμό συμμετεχόντων πιθανόν να αντανάκλα πρακτικές δυσκολίες στρατολόγησης ή περιορισμούς κατά τη διεξαγωγή. Οι περιπτώσεις ισότητας υποδεικνύουν ακριβή τήρηση του αρχικού σχεδιασμού, επιβεβαιώνοντας την αξιοπιστία των πρωτοκόλλων. Τα αποτελέσματα δείχνουν ότι οι αποκλίσεις μεταξύ προγραμματισμένου και πραγματικού δείγματος είναι συχνό φαινόμενο, το οποίο δύναται να επηρεάσει την ερμηνεία των ευρημάτων και, συνεπώς, πρέπει να λαμβάνεται υπόψη κατά την κρίση της μεθοδολογικής αρτιότητας των κλινικών δοκιμών.

ΣΤΑΤΙΣΤΙΚΗ ΠΡΟΣΕΓΓΙΣΗ ΔΕΔΟΜΕΝΩΝ

Για την μελέτη και ανάλυση των δεδομένων χρησιμοποιήθηκε το αρχικό δείγμα αποτελούμενο από 2099 κλινικές δοκιμές, από τις οποίες εξαιρέθηκαν 15, καθώς είχαν κενά όλα τα κύρια χαρακτηριστικά. Έτσι το τελικό δείγμα της μελέτης αποτελείται από 2084 κλινικές δοκιμές. Επιπλέον, η Φάση 0 συμπεριλήφθηκε στην Φάση Ι καθώς μόνο 3 δοκιμές βρίσκονταν στην συγκεκριμένη Φάση, χωρίς να υπάρχει κάποιο στατιστικό ενδιαφέρον.

1.Κλινική Φάση – Φύλο και Ηλικιακή ομάδα

Συγκεντρωτικός πίνακας συχνότητας μεταβλητών Φύλου και Ηλικίας ανά κλινική δοκιμή								
	PHASE 1 (N=46)	PHASE1 PHASE2 (N=31)	PHASE2 (N=342)	PHASE2 PHASE3 (N=40)	PHASE3 (N=898)	PHASE4 (N=219)	NA (N=508)	TOTAL (N=2084)
Sex								
Male	1 (2,2%)	2 (6,5%)	11 (3,2%)	1 (2,5%)	35 (3,9%)	10 (4,6%)	46 (9,0%)	106(5,0%)
Female	4 (8,7%)	6 (19,4%)	20 (5,8%)	4 (10%)	49 (5,5%)	16 (7,3%)	50 (9,8%)	149(7,1%)
All	41 (89,1%)	23 (74,2%)	311 (90,9%)	35 (87,5%)	814 (90,6%)	193 (88,1%)	412 (81,1%)	1829(87,7%)
Age								
CHILD	6 (13%)	-	14 (4,1%)	2 (5%)	44 (4,9%)	6 (2,7%)	37 (7,3%)	109(5,2%)
CHILD, ADULT	3 (6,5%)	1 (3,2%)	-	1 (2,5%)	13 (1,4%)	2 (0,9%)	11 (2,2%)	31(1,4%)
CHILD, ADULT, OLDER_ADULT	1 (2,2%)	-	24 (7%)	5 (12,5%)	75 (8,4%)	9 (4,1%)	32 (6,3%)	146(7,0%)
ADULT	14 (30,4%)	3 (9,7%)	9 (2,6%)	5 (12,5%)	29 (3,2%)	10 (4,6%)	119 (23,4%)	189(9,0%)
ADULT, OLDER_ADULT	22 (47,8%)	26 (83,9%)	294 (86%)	27 (67,5%)	731 (81,4%)	190 (86,8%)	298 (58,7%)	1588(76,1%)
OLDER_ADULT	-	1 (3,2%)	1 (0,3%)	-	6 (0,7%)	2 (0,9%)	11 (2,2%)	21(1,0%)

Πίνακας 3

Ο Πίνακας 3 συνοψίζει την κατανομή των συμμετεχόντων ανά φύλο και ηλικιακή ομάδα στις διάφορες φάσεις των κλινικών δοκιμών. Σχετικά με το φύλο, η συντριπτική πλειοψηφία των μελετών περιλαμβάνει και άντρες και γυναίκες μαζί (All), με ποσοστά που ξεπερνούν το 80% σε όλες τις φάσεις και φτάνουν έως το 90,6% στην PHASE 3. Οι μελέτες που αφορούν αποκλειστικά άντρες ή γυναίκες είναι πολύ περιορισμένες, κυμαίνονται σε χαμηλά ποσοστά (κάτω από 10%) και δεν αποτελούν τον κανόνα. Αυτό δείχνει ότι οι περισσότερες κλινικές δοκιμές έχουν σχεδιαστεί ώστε να είναι αντιπροσωπευτικές και για τα δύο φύλα.

Όσον αφορά την ηλικία, παρατηρείται ότι οι περισσότερες δοκιμές εστιάζουν σε ενήλικες και ηλικιωμένους (ADULT, OLDER ADULT), οι οποίοι αντιπροσωπεύουν την πλειοψηφία σε όλες τις φάσεις, με εντυπωσιακά υψηλά ποσοστά, όπως 86% στην PHASE 2 και 81,4% στην PHASE 3. Οι μελέτες που περιλαμβάνουν παιδιά είναι πολύ περιορισμένες και σε χαμηλά ποσοστά, γεγονός που φανερώνει τη δυσκολία ή τον περιορισμό στη συμμετοχή παιδικών πληθυσμών σε κλινικές δοκιμές. Σε ορισμένες περιπτώσεις καταγράφονται συνδυασμοί ηλικιακών ομάδων (π.χ. CHILD, ADULT ή CHILD, ADULT, OLDER ADULT), όμως αυτοί αποτελούν τη μειοψηφία. Συνολικά, τα αποτελέσματα αναδεικνύουν ότι το τυπικό προφίλ συμμετεχόντων στις περισσότερες φάσεις είναι άνδρες και γυναίκες μαζί, κυρίως ενήλικες και ηλικιωμένοι, με τα παιδιά να μην προτιμώνται.

2.Κλινική Φάση - Βασικά Χαρακτηριστικά

Στον Πίνακα 4 παρουσιάζονται τα βασικά χαρακτηριστικά των κλινικών δοκιμών ανά φάση, με μεταβλητές τον τελικό αριθμό συμμετεχόντων (Enrollment), τη διάρκεια σε ημέρες (Duration), τον τρόπο κατανομής (Allocation), το μοντέλο παρέμβασης (Intervention model) και τη μέθοδο τυφλοποίησης (Masking). Σχετικά με τον τελικό αριθμό συμμετεχόντων, παρατηρούμε ότι οι μέσοι όροι αυξάνονται καθώς προχωρούν οι φάσεις. Από μόλις 50 συμμετέχοντες στην PHASE 1 σε 1384 συμμετέχοντες στην PHASE 3, γεγονός που αντανάκλα την κλασική λογική των κλινικών δοκιμών όπου οι πρώιμες φάσεις έχουν μικρό αριθμό για λόγους ασφάλειας, ενώ οι μεταγενέστερες φάσεις χρειάζονται μεγάλα δείγματα για αξιοπιστία. Όσον αφορά τη διάρκεια, παρατηρούμε μεγάλες μέσες διάρκειες, με υψηλή τυπική απόκλιση, κάτι που δείχνει μεγάλη διακύμανση μεταξύ των μελετών. Οι πιο μεγάλες διάρκειες εμφανίζονται στην PHASE 1/2 (περίπου 1859 μέρες) και στην PHASE 3 (περίπου 1603 μέρες).

	PHASE 1 (N=46)	PHASE1 PHASE2 (N=31)	PHASE2 (N=342)	PHASE2 PHASE3 (N=40)	PHASE3 (N=898)	PHASE4 (N=219)	NA (N=508)	TOTAL (N=2084)
Enrollment	50,17 (S.D.58,11) 833,02	112,65 (S.D.109,75) 1858,19	249,11 (S.D.356,80) 1259,57	416,63 (S.D.584,57) 1172,05	1384,35 (S.D.2819,73) 1603,46	894,57 (S.D.2794,47) 1046,90	376,77 (S.D.2921,21) 741,1	834,03 (S.D.2570,69) 1256,83
Duration (days)	(S.D.749,79)	(S.D.1181,97)	(S.D.863,66)	(S.D.634,26)	(S.D.986,36)	(S.D.752,48)	(S.D.703,45)	(S.D.944,12)
Allocation								
NON_RANDOMIZED	13 (30,2%)	10 (32,3%)	36 (10,6%)	3 (7,5%)	33 (3,7%)	22 (10,1%)	38 (7,5%)	155(7,4%)
RANDOMIZED	19 (44,2%)	14 (45,2%)	238 (70,2%)	34 (85%)	785 (87,8%)	170 (78,3%)	403 (79,3%)	1663(79,7%)
NA	11 (25,6%)	7 (22,6%)	65 (19,2%)	3 (7,5%)	76 (8,5%)	25 (11,5%)	67 (13,2%)	254(12,1%)
Intervention model								
SINGLE_GROUP	13 (29,5%)	13 (41,9%)	82 (24,4%)	4 (10%)	98 (11,1%)	40 (18,3%)	77 (15,2%)	327(15,6%)
PARALLEL	16 (36,4%)	17 (54,8%)	247 (73,5%)	35 (87,5%)	762 (86,3%)	157 (72%)	330 (65,1%)	1564(75,0%)
SEQUENTIAL	5 (11,4%)	1 (3,2%)	3 (0,9%)	-	3 (0,3%)	1 (0,5%)	3 (6%)	16(0,7%)
CROSSOVER	10 (22,7%)	-	1 (0,3%)	-	14 (1,6%)	18 (8,3%)	89 (17,6%)	132(6,3%)
FACTORIAL	-	-	3 (0,9%)	1 (2,5%)	6 (0,7%)	2 (0,9%)	8 (1,6%)	20(0,9%)
Masking								
NONE	24 (54,5%)	22 (71%)	155 (46,1%)	11 (27,5%)	329 (37,4%)	115 (53%)	212 (41,8%)	868(41,6%)
SINGLE	15 (34,1%)	2 (6,5%)	15 (4,5%)	5 (12,5%)	21 (2,4%)	37 (17,1%)	135 (26,6%)	230(11,0%)
DOUBLE	4 (9,1%)	2 (6,5%)	60 (17,9%)	7 (17,5%)	204 (23,2%)	34 (15,7%)	83 (16,4%)	394(18,9%)
TRIPLE	-	1 (3,2%)	32 (9,5%)	8 (20%)	92 (10,5%)	9 (4,1%)	37 (7,3%)	179(8,5%)
QUADRUPLE	1 (2,3%)	4 (12,9%)	74 (22%)	9 (22,5%)	233 (26,5%)	22 (10,1%)	40 (7,9%)	383(18,3%)

Πίνακας 4

Σχετικά με τη μεθοδολογία, παρατηρούμε ότι οι περισσότερες μελέτες από την Φάση 2 και μετά είναι τυχαιοποιημένες (70,2% στην Φάση 2 και 87,8% στην Φάση 3), ενώ οι μη τυχαιοποιημένες επικρατούν μόνο στην Φάση 1. Όσον αφορά το μοντέλο παρέμβασης, το παράλληλο μοντέλο κυριαρχεί στις πιο ώριμες φάσεις (86,3% στην Φάση 3), ενώ το μοντέλο μιας ομάδας και το διασταυρούμενο μοντέλο είναι πιο συχνά στις πρώιμες φάσεις. Όσον αφορά το είδος τυφλοποίησης, στην Φάση 1 κυριαρχεί η απουσία τυφλοποίησης (54,5%), ενώ στις Φάση 2 και Φάση 3 βλέπουμε περισσότερη χρήση διπλής (Double), τριπλής (Triple) και τετραπλής (Quadruple) τυφλοποίησης. Χαρακτηριστικά, στην Φάση 3 η τετραπλή τυφλοποίηση φτάνει το 26,5%. Συνολικά, ο πίνακας δείχνει μια ξεκάθαρη τάση: όσο προχωρούν οι φάσεις, τόσο αυξάνεται το μέγεθος δείγματος, η διάρκεια, αλλά και η αυστηρότητα στη μεθοδολογία (τυχαιοποίηση, ερευνητικός σχεδιασμός, τυφλοποίηση), κάτι που ενισχύει σημαντικά την αξιοπιστία των αποτελεσμάτων.

Στην συνέχεια, πραγματοποιήθηκε έλεγχος κανονικότητας με το τεστ Kolmogorov–Smirnov, ο οποίος έδειξε ότι τα δεδομένα δεν ακολουθούν την κανονική κατανομή ($p < 0,05$). Για τον λόγο αυτό επιλέχθηκε η χρήση μη παραμετρικών ελέγχων. Συγκεκριμένα, εφαρμόστηκε ο έλεγχος Kruskal–Wallis H, ο οποίος είναι κατάλληλος για τη σύγκριση ομάδων όταν οι μεταβλητές υπό

μελέτη περιλαμβάνουν περισσότερες από δύο κατηγορίες και οι προϋποθέσεις παραμετρικών ελέγχων, όπως η ANOVA, δεν ικανοποιούνται. Ο έλεγχος αυτός επιτρέπει τη σύγκριση των μέσων όρων κατάταξης (mean ranks) μεταξύ των διαφορετικών φάσεων, αναδεικνύοντας εάν υπάρχουν στατιστικά σημαντικές διαφορές στη μεθοδολογία και τα χαρακτηριστικά των κλινικών δοκιμών.

<i>Σύγκριση μέσων όρων διάταξης των μεταβλητών Allocation, Intervention model, Masking και Enrollment ανά κλινική φάση</i>					
	Phases	N	Mean Rank	Kruskal-Wallis H	p
Allocation	PHASE 1	43	957,38	9,211	0,162
	PHASE1 PHASE2	31	910,21		
	PHASE2	339	1074,25		
	PHASE2 PHASE3	40	990,71		
	PHASE3	894	1034,93		
	PHASE4	217	1005,27		
	NA	508	1045,42		
Intervention Model	PHASE 1	44	1116,28	74,969	<0,001
	PHASE1 PHASE2	31	738,48		
	PHASE2	336	896,77		
	PHASE2 PHASE3	40	1038,45		
	PHASE3	883	1027,33		
	PHASE4	218	1019,60		
	NA	507	1137,08		
Masking	PHASE 1	44	732,40	101,461	<0,001
	PHASE1 PHASE2	31	746,81		
	PHASE2	336	1036,65		
	PHASE2 PHASE3	40	1204,71		
	PHASE3	879	1146,20		
	PHASE4	217	855,43		
	NA	507	918,08		
Enrollment	PHASE 1	46	309,83	855,137	<0,001
	PHASE1 PHASE2	31	578,81		
	PHASE2	341	836,39		
	PHASE2 PHASE3	40	969,73		
	PHASE3	895	1444,96		
	PHASE4	219	1047,71		
	NA	508	561,73		

Πίνακας 5

Στον Πίνακα 5, παρουσιάζονται τα αποτελέσματα του ελέγχου Kruskal–Wallis, ο οποίος χρησιμοποιείται για να ελέγξει εάν υπάρχουν στατιστικά σημαντικές διαφορές ανάμεσα στις φάσεις των κλινικών δοκιμών για τέσσερις μεταβλητές: Τυχαιοποίηση (Allocation), Μοντέλο παρέμβασης (Intervention model), Τυφλοποίηση (Masking) και Τελικός αριθμός συμμετεχόντων (Enrollment). Για την μεταβλητή τυχαιοποίηση παρατηρείται ότι, παρόλο που οι μέσες τιμές (Mean Ranks) διαφέρουν ελαφρώς μεταξύ των φάσεων, το αποτέλεσμα δεν είναι στατιστικά σημαντικό ($H=9,211$, $p=0,162$). Αυτό σημαίνει ότι ο τρόπος κατανομής (τυχαιοποιημένος ή μη) των συμμετεχόντων στις

παρεμβάσεις παραμένει σχετικά σταθερός από τη μία φάση στην άλλη και δεν διαφοροποιείται ουσιαστικά στη μεθοδολογία.

Αντίθετα, για το μοντέλο παρέμβασης, την τυφλοποίηση και τον τελικό αριθμό συμμετεχόντων τα αποτελέσματα είναι ιδιαίτερα εντυπωσιακά, καθώς οι τιμές του Kruskal–Wallis H είναι υψηλές και οι p τιμές πολύ χαμηλές ($p < 0,001$), δείχνοντας στατιστικά σημαντικές διαφορές μεταξύ των φάσεων. Στο μοντέλο παρέμβασης, οι μέσες τάξεις (Mean Ranks) διαφοροποιούνται έντονα, υποδηλώνοντας αλλαγή στον τρόπο σχεδιασμού καθώς οι δοκιμές προχωρούν. Στην τυφλοποίηση, οι υψηλότερες τιμές Mean Rank σε προχωρημένες φάσεις δείχνουν ότι εφαρμόζονται πιο αυστηρές τεχνικές τυφλοποίησης (π.χ. Double, Triple, Quadruple Masking). Τέλος, ο τελικός αριθμός συμμετεχόντων εμφανίζει τη μεγαλύτερη διαφορά, με πολύ χαμηλό Mean Rank στην PHASE 1 και πολύ υψηλό στην PHASE 3, γεγονός που αποτυπώνει την εκρηκτική αύξηση του αριθμού συμμετεχόντων στις πιο ώριμες φάσεις. Συνολικά, από τον πίνακα αντιλαμβανόμαστε ότι, αν και η κατανομή συμμετεχόντων (Allocation) μένει σταθερή, οι υπόλοιπες παράμετροι βελτιώνονται θεαματικά και γίνονται πιο αυστηρές στις προχωρημένες φάσεις, ενισχύοντας την αξιοπιστία των αποτελεσμάτων.

3.Enrollment – Βασικά Χαρακτηριστικά

Στην ενότητα αυτή θα παρουσιάσουμε τα αποτελέσματα της πολυμεταβλητής ανάλυσης με εξαρτημένη μεταβλητή τον τελικό αριθμό συμμετεχόντων (Enrollment) και ανεξάρτητες μεταβλητές το την τυχαιοποίηση (Allocation), την τυφλοποίηση (Masking), τις Φάσεις των κλινικών δοκιμών (Phases), το Φύλο και την Ηλικία των συμμετεχόντων. Το μοντέλο παρέμβασης (Intervention Model) εξαιρέθηκε από την πολυμεταβλητή ανάλυση, δεδομένου ότι προηγουμένως δεν είχε διαπιστωθεί στατιστικά σημαντική συσχέτισή του με τον τελικό αριθμό συμμετεχόντων (Enrollment).

Ο Πίνακας 6 παρουσιάζει τα αποτελέσματα της πολλαπλής παλινδρόμησης με εξαρτημένη μεταβλητή τον τελικό αριθμό συμμετεχόντων (Enrollment). Ο συντελεστής προσδιορισμού $R^2 = 0.05328$, το οποίο υποδεικνύει ότι το μοντέλο εξηγεί μόλις το 5,3% της διακύμανσης που υπάρχει

στα δεδομένα. Παρόλο που το ποσοστό αυτό είναι χαμηλό και υποδηλώνει την ύπαρξη σημαντικών παραγόντων που δεν περιλήφθηκαν (όπως το είδος της ασθένειας, η χρηματοδότηση ή η γεωγραφική περιοχή), το συνολικό μοντέλο είναι στατιστικά σημαντικό ($F=6,628$, $p<0,001$), γεγονός που επιτρέπει την περαιτέρω ερμηνεία του.

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	825.94922	296.74525	2.783	0.005431	**
factor(Sex)FEMALE	-52.95497	217.49884	-0.243	0.807665	
factor(Sex)MALE	-293.66034	250.34049	-1.173	0.240919	
factor(Age)ADULT, OLDER_ADULT	411.70577	207.78500	1.981	0.047684	*
factor(Age)CHILD	-86.93354	300.26284	-0.290	0.772210	
factor(Age)CHILD, ADULT	-10.78448	468.82085	-0.023	0.981650	
factor(Age)CHILD, ADULT, OLDER_ADULT	-139.26215	287.58885	-0.484	0.628269	
factor(Age)OLDER_ADULT	64.66770	556.28246	0.116	0.907466	
factor(Phases)PHASE1	-270.03382	386.61955	-0.698	0.484979	
factor(Phases)PHASE1 PHASE2	-516.56021	452.86290	-1.141	0.254152	
factor(Phases)PHASE2	-453.79665	183.30609	-2.476	0.013384	*
factor(Phases)PHASE2 PHASE3	-192.62965	401.45057	-0.480	0.631398	
factor(Phases)PHASE3	650.14643	158.17395	4.110	4.11e-05	***
factor(Phases)PHASE4	140.43888	204.12997	0.688	0.491541	
factor(Allocation)NON_RANDOMIZED	-229.62703	250.97064	-0.915	0.360326	
factor(Allocation)RANDOMIZED	-349.34638	187.14485	-1.867	0.062089	.
factor(Masking)NONE	-582.01777	166.05673	-3.505	0.000467	***
factor(Masking)QUADRUPLE	103.38704	176.06033	0.587	0.557120	
factor(Masking)SINGLE	-573.68734	213.69284	-2.685	0.007321	**
factor(Masking)TRIPLE	-283.32301	218.11146	-1.299	0.194101	
`Duration (days)`	0.08778	0.06340	1.385	0.166322	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2388 on 1980 degrees of freedom
(98 observations deleted due to missingness)

Multiple R-squared: 0.06275, Adjusted R-squared: 0.05328

F-statistic: 6.628 on 20 and 1980 DF, p-value: < 2.2e-16

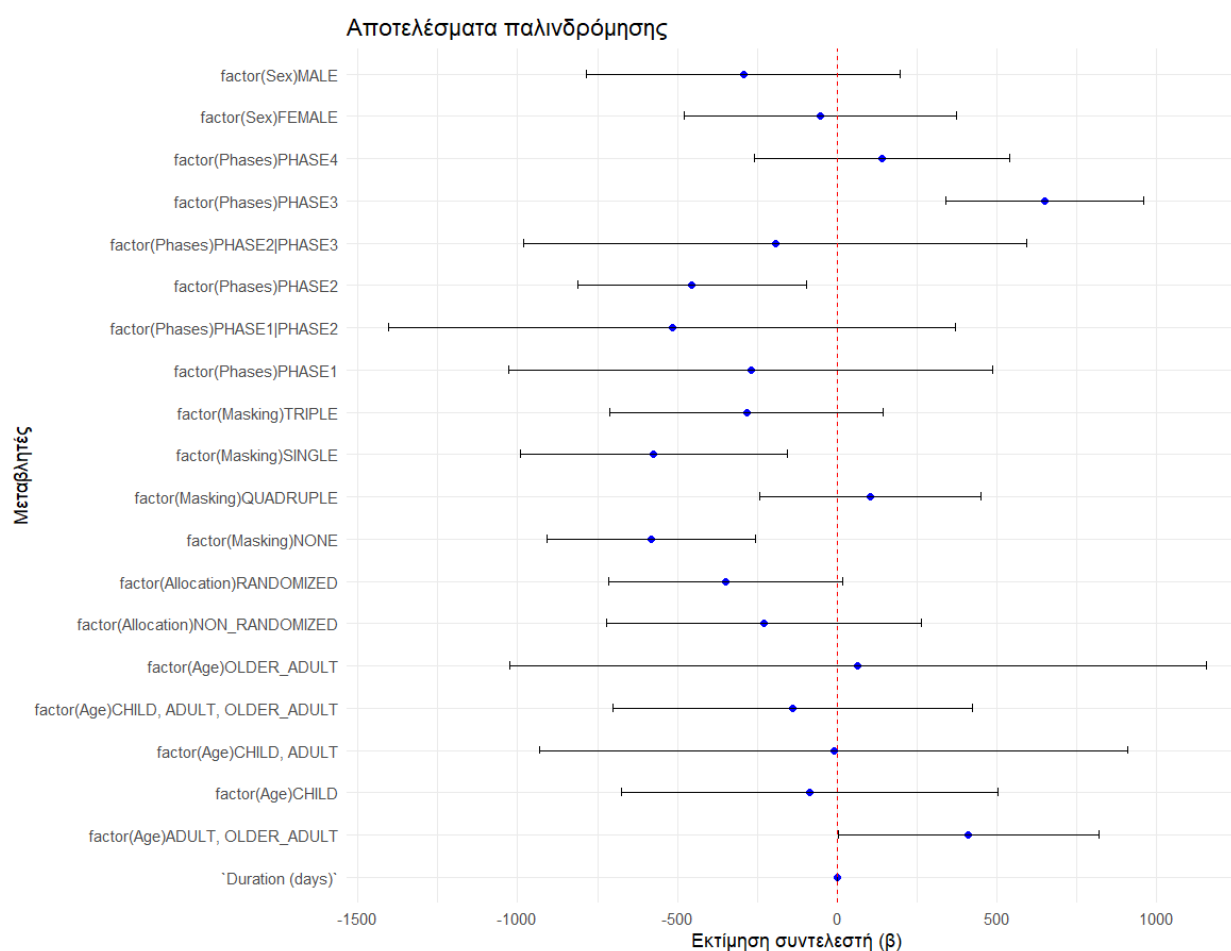
Πίνακας 6

Από τις έξι ανεξάρτητες μεταβλητές που αξιολογήθηκαν προέκυψε ότι οι τρεις από αυτές επηρεάζουν τον τελικό αριθμό των συμμετεχόντων. Συγκεκριμένα, αυτές είναι η ηλικία των συμμετεχόντων (Age), η φάση της κλινικής δοκιμής (Phases) και η τυφλοποίηση (Masking). Οριακά θα μπορούσαμε να πούμε ότι ο τελικός αριθμός συμμετεχόντων επηρεάζεται και από τον τρόπο κατανομής – τυχαιοποίηση (Allocation). Με τον τρόπο αυτό

μπορεί να προκύψει ένα τελικό μοντέλο παλινδρόμησης και να αποτυπωθεί με την εξής εξίσωση:

$$\text{Enrollment} = 825.95 + 411.71 \cdot \text{Age(ADULT, OLDER_ADULT)} - 453.80 \cdot \text{Phase(PHASE2)} + 650.15 \cdot \text{Phase(PHASE3)} - 582.02 \cdot \text{Masking(NONE)} - 573.69 \cdot \text{Masking(SINGLE)} + \varepsilon$$

Διαγραμματικά, τα αποτελέσματα της παλινδρόμησης παρουσιάζονται στο Σχήμα 7. Στις περιπτώσεις των στατιστικά σημαντικών αποτελεσμάτων, η κόκκινη διακεκομμένη γραμμή δεν τέμνει το 95% διάστημα εμπιστοσύνης του συντελεστή παλινδρόμησης.



Σχήμα 7

Το παραπάνω μοντέλο υποδεικνύει ότι η συμμετοχή ενηλίκων και των ηλικιωμένων (ADULT, OLDER_ADULT) έχει θετική επίδραση στον τελικό

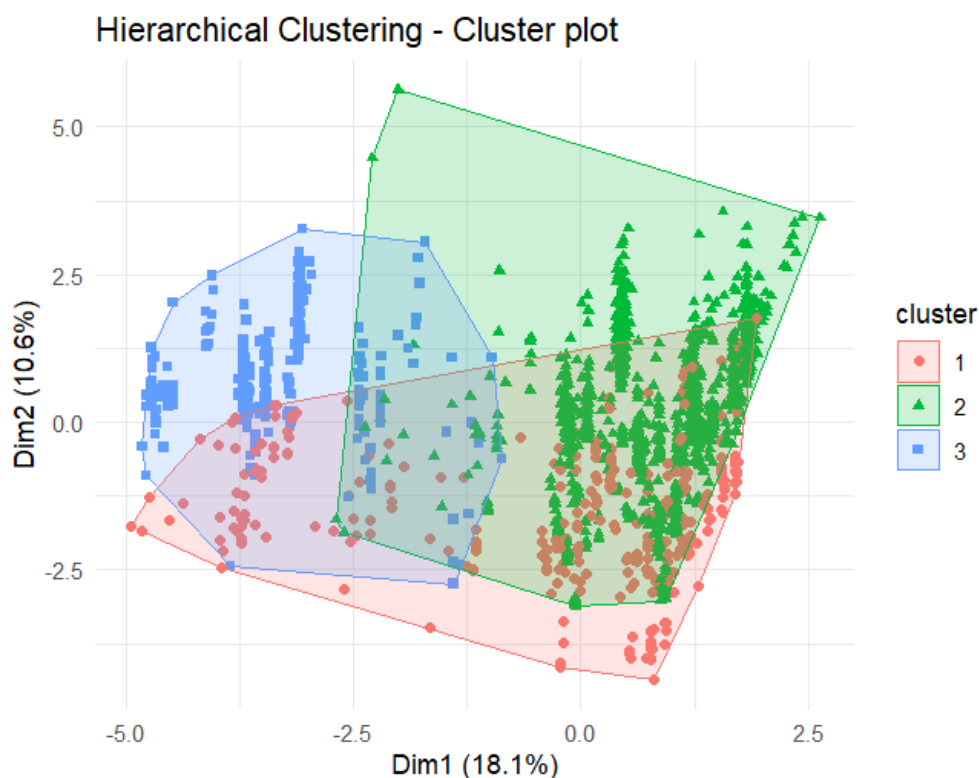
αριθμό των συμμετεχόντων σε σχέση με την baseline κατηγορία ηλικίας (ενήλικες μόνο). Η Φάση II έχει αρνητική επίδραση ενώ η Φάση III έχει ισχυρά θετική επίδραση στον τελικό αριθμό των συμμετεχόντων. Η μη τυφλοποίηση και η μονή τυφλοποίηση έχουν αρνητική επίδραση στον τελικό αριθμό των συμμετεχόντων.

Συνοψίζοντας, το τελικό μοντέλο αναδεικνύει η Φάση II μειώνει σημαντικά την εξαρτημένη μεταβλητή (τον τελικό αριθμό συμμετεχόντων), ενώ η Φάση III την αυξάνει σημαντικά. Η συμμετοχή ενηλίκων και ηλικιωμένων έχει μεγαλύτερη θετική τιμή από την αναφορά. Η απουσία τυφλοποίησης ή η μονή τυφλοποίησης σχετίζονται με σημαντική μείωση σε σχέση με τα υπόλοιπα είδη τυφλοποίησης. Το φύλο και η διάρκεια δεν φαίνεται να επηρεάζουν στατιστικά σημαντικά τον τελικό αριθμό των συμμετεχόντων στην κλινική δοκιμή.

4.Ομαδοποίηση κλινικών δοκιμών

Στη συνέχεια θα παρουσιάσουμε τα αποτελέσματα της ιεραρχικής ομαδοποίησης (cluster analysis) με σκοπό να ομαδοποιήσουμε τις κλινικές δοκιμές με βάση τα χαρακτηριστικά τους.

Το ιεραρχικό μοντέλο οδήγησε σε 3 συστάδες (ομάδες), οι οποίες παρουσιάζονται στο Σχήμα 7. Οι άξονες του γραφήματος (Dim1 και Dim2) είναι οι δύο πρώτες συνιστώσες που προέκυψαν από την ανάλυση κύριων συνιστωσών (PCA), οι οποίες εξηγούν τη μεγαλύτερη μεταβλητότητα των δεδομένων. Μαζί εξηγούν περίπου το 28.7% της συνολικής διακύμανσης.



Σχήμα 8

Κάθε χρωματισμένη περιοχή αντιπροσωπεύει μια ομάδα (cluster). Με πορτοκαλί χρώμα παρουσιάζεται η πρώτη ομάδα (Cluster 1), με πράσινο παρουσιάζεται η δεύτερη ομάδα (Cluster 2) και με γαλάζιο παρουσιάζεται η τρίτη ομάδα (Cluster 3). Οι περιοχές αυτές δείχνουν που βρίσκονται συγκεντρωμένα τα σημεία κάθε ομάδας. Κάθε σημείο είναι μια παρατήρηση (μια κλινική δοκιμή στην περίπτωση μας). Όσο πιο κοντά είναι δύο σημεία, τόσο πιο όμοιες είναι οι δοκιμές ως προς τα χαρακτηριστικά τους. Οι αποστάσεις δείχνουν “διαφορές” μεταξύ των κλινικών δοκιμών. Εάν δύο ομάδες είναι καλά χωρισμένα (με μικρή επικάλυψη), σημαίνει ότι οι ομάδες έχουν σαφώς διαφορετικά χαρακτηριστικά. Στην περίπτωση που επικαλύπτονται, σημαίνει ότι οι ομάδες μοιάζουν σε κάποια χαρακτηριστικά.

Στον Πίνακα 7 παρουσιάζονται τα χαρακτηριστικά της κάθε ομάδας. Στην πρώτη ομάδα ανήκουν 396 κλινικές δοκιμές, στη δεύτερη ομάδα ανήκουν 853 κλινικές δοκιμές, ενώ στην τρίτη ομάδα ανήκουν 748 κλινικές δοκιμές. Στην τρίτη ομάδα ανήκουν οι μεγαλύτερες σε διάρκεια και αριθμό συμμετεχόντων κλινικές

δοκιμές (κατά βάση δοκιμές Φάσης III παράλληλου σχεδιασμού). Στην δεύτερη ομάδα ανήκουν οι μικρότερες σε διάρκεια και αριθμό συμμετεχόντων κλινικές δοκιμές (κατά βάση δοκιμές Φάσης II και IV παράλληλου και διασταυρούμενου σχεδιασμού), ενώ στην πρώτη ομάδα ανήκουν οι μεσαίες σε διάρκεια και αριθμό συμμετεχόντων κλινικές δοκιμές (κατά βάση δοκιμές Φάσης II και III μιας ομάδας). Ως προς την τυφλοποίηση, στην πρώτη ομάδα ανήκουν κλινικές δοκιμές στις οποίες δεν εφαρμόζεται τυφλοποίηση. Στη δεύτερη ομάδα ανήκουν κλινικές δοκιμές στις οποίες είτε δεν εφαρμόζεται τυφλοποίηση είτε εφαρμόζεται μονή ή διπλή τυφλοποίηση. Στην τρίτη ομάδα ανήκουν κλινικές δοκιμές στις οποίες είτε δεν εφαρμόζεται τυφλοποίηση είτε εφαρμόζεται διπλή ή τετραπλή τυφλοποίηση. Ως προς την ηλικία των συμμετεχόντων δεν υπάρχει μεγάλη διαφοροποίηση μεταξύ των 3 ομάδων. Στην πρώτη ομάδα ανήκουν κυρίως μη τυχαιοποιημένες κλινικές δοκιμές και οι κλινικές δοκιμές για την οποία δεν έχει καταγραφεί η πληροφορία σχετικά με το ένα είναι τυχαιοποιημένες ή όχι. Στη δεύτερη και στην τρίτη ομάδα είναι σχεδόν ισομοιρασμένες οι τυχαιοποιημένες κλινικές δοκιμές.

Cluster	mean_participants	mean_duration	n	
1	1	596.	1358.	396
2	2	294.	870.	853
3	3	1499.	1622.	748

Cluster	NA	PHASE1	PHASE1 PHASE2	PHASE2	PHASE2 PHASE3	PHASE3	PHASE4
1	103	21	16	99	5	106	46
2	390	22	12	227	32	11	159
3	6	0	3	0	2	736	1

Cluster	DOUBLE	NONE	QUADRUPLE	SINGLE	TRIPLE
1	8	380	0	7	1
2	173	241	145	208	86
3	189	226	233	8	92

Cluster	ADULT	ADULT, OLDER_ADULT	CHILD	CHILD, ADULT	CHILD, ADULT, OLDER_ADULT	OLDER_ADULT
1	26		298	28	12	27
2	136		595	48	12	52
3	21		632	31	7	51

Cluster	NA	NON_RANDOMIZED	RANDOMIZED
1	249	138	9
2	1	6	846
3	0	0	748

Cluster	CROSSOVER	FACTORIAL	PARALLEL	SEQUENTIAL	SINGLE_GROUP
1	1	1	76	8	310
2	120	13	709	5	6
3	9	6	727	3	3

Πίνακας 7

ΣΥΖΗΤΗΣΗ

Η ανάλυση των κλινικών δοκιμών στην Ελλάδα ανέδειξε μερικά αναμενόμενα αλλά και αρκετά απρόσμενα στοιχεία. Όπως αναμέναμε, το μέγεθος δείγματος αυξάνεται όσο προχωράμε από τις πρώιμες προς τις πιο ώριμες φάσεις, κάτι που συμβαδίζει με τα διεθνή πρότυπα. Εντύπωση όμως προκαλεί ότι στη Φάση I/II οι μελέτες διαρκούν περισσότερο απ' όλες, γεγονός που ίσως συνδέεται με τις δυσκολίες στον καθορισμό δοσολογίας και τον μεταβατικό χαρακτήρα της φάσης.

Ενδιαφέρον παρουσιάζουν τα χαρακτηριστικά των συμμετεχόντων. Οι περισσότερες μελέτες περιλαμβάνουν άνδρες και γυναίκες ενήλικες, κάτι θετικό για την αντιπροσωπευτικότητα, αλλά η συμμετοχή παιδιών ή άλλων ειδικών ομάδων παραμένει περιορισμένη. Αν και αυτό ίσως αφήνει κενά που θα μπορούσαν να καλυφθούν με πιο στοχευμένες μελέτες, εδώ θα πρέπει να ληφθεί υπόψη και η ασθένεια που μελετάται καθώς κάτι τέτοιο ενδέχεται να περιορίζει το δείγμα.

Από τα πιο ανησυχητικά ευρήματα ήταν η μεγάλη απόκλιση ανάμεσα στο αρχικά προγραμματισμένο και στο τελικό δείγμα, που βρέθηκε σε ένα σημαντικό ποσοστό των μελετών. Αυτό δείχνει ότι, παρότι οι στόχοι τίθενται σωστά, στην πράξη υπάρχουν εμπόδια στην ένταξη και διατήρηση των συμμετεχόντων. Πρόκειται για ζήτημα που μειώνει την ισχύ των αποτελεσμάτων και χρειάζεται καλύτερο σχεδιασμό και οργάνωση.

Οι στατιστικές αναλύσεις προσέφεραν χρήσιμες ενδείξεις για το πώς οι μεθοδολογικές επιλογές επηρεάζουν την πορεία των κλινικών δοκιμών. Το Kruskal-Wallis test έδειξε σημαντικές διαφορές στο μοντέλο παρέμβασης, στην τυφλοποίηση και στον τελικό αριθμό συμμετεχόντων μεταξύ των φάσεων, κάτι που δείχνει ότι ο σχεδιασμός διαφοροποιείται ουσιαστικά καθώς οι δοκιμές προχωρούν. Αντίθετα, η τυχαιοποίηση δεν εμφάνισε στατιστικά σημαντικές διαφορές, γεγονός που υποδηλώνει ότι ο τρόπος κατανομής των συμμετεχόντων παραμένει σχετικά σταθερός ανεξαρτήτως φάσης. Η πολλαπλή παλινδρόμηση κατέδειξε πως ο τελικός αριθμός συμμετεχόντων επηρεάζεται περισσότερο από την φάση, την τυφλοποίηση και την ηλικία, ενώ το φύλο και η διάρκεια δεν είχαν στατιστικά σημαντική επίδραση. Ενδιαφέρον είναι το γεγονός

ότι, παρά τα παραπάνω, το μοντέλο εξηγεί μόλις το 5,3% της μεταβλητότητας, κάτι που αφήνει να εννοηθεί ότι άλλοι, μη μετρημένοι και μη καταχωρημένοι στη βάση του ClinicalTrials.gov παράγοντες πιθανόν να παίζουν καθοριστικό ρόλο στον προσδιορισμό του τελικού αριθμού συμμετεχόντων.

Με βάση τα χαρακτηριστικά τους οι κλινικές δοκιμές μπορούν να κατηγοριοποιηθούν σε 3 ομάδες. Η πρώτη ομάδα αφορά κατά βάση μη τυχαιοποιημένες μεσαίου μεγέθους και διάρκειας κλινικές δοκιμές Φάσης II και III που δεν χρησιμοποιούν τυφλοποίηση και στις οποίες συμμετέχουν ηλικιωμένα άτομα. Ο βασικός σχεδιασμός σε αυτή την ομάδα είναι ο σχεδιασμός μιας ομάδας. Η δεύτερη ομάδα αφορά κατά βάση τυχαιοποιημένες μικρού μεγέθους και διάρκειας κλινικές δοκιμές Φάσης II και IV που είτε δεν χρησιμοποιούν τυφλοποίηση ή χρησιμοποιούν μονή τυφλοποίηση και στις οποίες συμμετέχουν ενήλικα και ηλικιωμένα άτομα. Οι βασικοί σχεδιασμοί σε αυτή την ομάδα είναι ο παράλληλος σχεδιασμός δύο ομάδων και ο διασταυρούμενος σχεδιασμός. Η τρίτη ομάδα αφορά κατά βάση τυχαιοποιημένες μεγάλου μεγέθους και διάρκειας κλινικές δοκιμές Φάσης III που χρησιμοποιούν διπλή, τριπλή ή τετραπλή τυφλοποίηση και στις οποίες συμμετέχουν κυρίως ηλικιωμένα άτομα. Ο βασικός σχεδιασμός σε αυτή την ομάδα είναι ο παράλληλος σχεδιασμός δύο ομάδων.

Η μελέτη βασίστηκε αποκλειστικά σε δεδομένα από τη βάση ClinicalTrials.gov, κάτι που έχει ορισμένους περιορισμούς. Τα στοιχεία που δημοσιεύονται δεν είναι πάντα πλήρη ή ενημερωμένα, καθώς εξαρτώνται από το τι αναρτούν οι ερευνητές και οι φορείς των μελετών. Έτσι, μπορεί να υπάρχουν ελλείψεις σε σημαντικές πληροφορίες. Επιπλέον, στη βάση πιθανών να μην καταγράφονται κάποιοι άλλοι παράγοντες που επηρεάζουν τη διεξαγωγή των δοκιμών, όπως για παράδειγμα η χρηματοδότηση ή άλλες ιδιαιτερότητες της κλινικής δοκιμής. Για τον λόγο αυτό, τα αποτελέσματα της ανάλυσης ενδέχεται να μην ανταποκρίνονται πλήρως την πραγματικότητα και να υπάρχουν μικρές αποκλίσεις.

Αν και η βάση δεδομένων ClinicalTrials.gov είναι μια κρατική υπηρεσία, η κυβέρνηση των ΗΠΑ δεν εξετάζει ούτε εγκρίνει την ασφάλεια και την επιστήμη όλων των μελετών που αναφέρονται στη βάση. Το National Institutes of Health

(NIH) και άλλες υπηρεσίες των ΗΠΑ εγκρίνουν μόνο τις μελέτες που χρηματοδοτούν. Οι υπηρεσίες της κυβέρνησης των ΗΠΑ δεν χρηματοδοτούν όλες τις μελέτες που αναφέρονται στο ClinicalTrials.gov. Η National Library of Medicine (NLM) πραγματοποιεί περιορισμένη αναθεώρηση των πληροφοριών της μελέτης που υποβάλλουν οι χορηγοί και οι ερευνητές για να συμπεριληφθούν στο αρχείο της μελέτης στη βάση δεδομένων ClinicalTrials.gov. Ο χορηγός ή ο ερευνητής της μελέτης είναι υπεύθυνος για την ασφάλεια, την επιστήμη και την ακρίβεια οποιασδήποτε μελέτης που αναφέρει στη βάση. Είναι επίσης υπεύθυνος για τη διασφάλιση ότι η μελέτη του ακολουθεί όλους τους σχετικούς νόμους και κανονισμούς.

ΣΥΜΠΕΡΑΣΜΑΤΑ

Συνοψίζοντας, οι κλινικές δοκιμές στην Ελλάδα δείχνουν να συμβαδίζουν σε μεγάλο βαθμό με τα διεθνή πρότυπα, αλλά εξακολουθούν να υπάρχουν αδυναμίες κυρίως στον τομέα που αφορά το τελικό μέγεθος δείγματος. Οι αποκλίσεις ανάμεσα σε όσα προγραμματίζονται και σε όσα τελικά υλοποιούνται μειώνουν την αξιοπιστία των αποτελεσμάτων, ενώ η ποικιλία μεθόδων υπολογισμού αν και εκ πρώτης δείχνει έλλειψη ομοιομορφίας δικαιολογείται από το αντικείμενο μελέτης της κάθε κλινικής δοκιμής. Το μέγεθος δείγματος φαίνεται να εξαρτάται από πολλούς παράγοντες, όπως η φάση, η διάρκεια και η τυφλοποίηση, κάτι που σημαίνει ότι δεν μπορούμε να το βλέπουμε απομονωμένα. Με πιο ρεαλιστικό σχεδιασμό, καλύτερη τήρηση των πρωτοκόλλων και πιο στοχευμένη στρατολόγηση, οι δοκιμές στη χώρα μας μπορούν να γίνουν πιο αξιόπιστες και να αποκτήσουν μεγαλύτερη διεθνή αξία.

Μια ενδιαφέρουσα κατεύθυνση για μελλοντική έρευνα θα ήταν η διεύρυνση της παρούσας μελέτης σε ευρύτερα γεωγραφικά πλαίσια, εφαρμόζοντάς την σε χώρες της Ευρώπης ή ακόμη και σε παγκόσμιο επίπεδο, προκειμένου να διερευνηθούν συγκριτικά οι διαφοροποιήσεις στο μέγεθος και στα χαρακτηριστικά των δειγμάτων μεταξύ διαφορετικών χωρών και να ανιχνευτούν τυχόν ευρύτερες τάσεις ή παράγοντες που τα επηρεάζουν.

BIBΛΙΟΓΡΑΦΙΑ

1. Abd ElHafeez, S., D'Arrigo, G., Leonardis, D., Fusaro, M., Tripepi, G., & Roumeliotis, S. (2021). Methods to analyze time-to-event data: the Cox regression analysis. *Oxidative medicine and cellular longevity*, 2021(1), 1302811.
2. Bhide, A., Shah, P. S., & Acharya, G. (2018). A simplified guide to randomized controlled trials. *Acta obstetricia et gynecologica Scandinavica*, 97(4), 380-387.
3. Bretz, F., Hothorn, T., & Westfall, P. (2008). Multiple comparison procedures in linear models. In *Compstat 2008: Proceedings in Computational Statistics* (pp. 423-431). Heidelberg: Physica-Verlag HD.
4. Chin, R., & Lee, B. Y. (2008). *Principles and practice of clinical trial medicine*. Elsevier.
5. Christian, J. B., Brouwer, E. S., Girman, C. J., Bennett, D., Davis, K. J., & Dreyer, N. A. (2020). Masking in pragmatic trials: who, what, and when to blind. *Therapeutic Innovation & Regulatory Science*, 54(2), 431-436.
6. Cressie, N. A. C., & Whitford, H. J. (1986). How to use the two sample t-test. *Biometrical Journal*, 28(2), 131-148.
7. Dignam, J. J., & Kocherginsky, M. N. (2008). Choice and interpretation of statistical tests used when competing risks are present. *Journal of Clinical Oncology*, 26(24), 4027-4034.
8. Friedman, L. M., Furberg, C. D., DeMets, D. L., Reboussin, D. M., & Granger, C. B. (2015). *Fundamentals of clinical trials*. Springer.
9. Friedman, L. M., Furberg, C. D., DeMets, D. L., Reboussin, D. M., & Granger, C. B. (2015). *Fundamentals of clinical trials*. springer.
10. Gad, S. C., & Rousseaux, C. G. (2002). Use and misuse of statistics in the design and interpretation of studies. In *Handbook of toxicologic pathology* (pp. 327-418). Academic Press.
11. Giannini, E. H. (2005). Chapter 6-design, measurement, and analysis of clinical investigations, Editor (s): Cassidy, JT, Petty, RE, Laxer, RM, Lindsley, CB Textbook of Pediatric Rheumatology.
12. Goldberg, D. E., & Scheiner, S. M. (2020). ANOVA and ANCOVA: field competition experiments. In *Design and analysis of ecological experiments* (pp. 69-93). Chapman and Hall/CRC.

13. Hilbe, J. M. (2011). *Negative binomial regression*. Cambridge University Press.
14. Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*. John Wiley & Sons.
15. Jones, L., Black, S. R., Arnold, L. E., & Fristad, M. A. (2019). Not all masks are created equal: Masking success in clinical trials of children and adolescents. *Journal of Clinical Child & Adolescent Psychology*, 48(sup1), S227-S233.
16. MacFarland, T. W., & Yates, J. M. (2016). Mann-whitney u test. In *Introduction to nonparametric statistics for the biological sciences using R* (pp. 103-132). Cham: Springer International Publishing.
17. Marshall, G., & Jonker, L. (2011). An introduction to inferential statistics: A review and practical guide. *Radiography*, 17(1), e1-e6.
18. McGlothlin, A. E., & Viele, K. (2018). Bayesian hierarchical models. *Jama*, 320(22), 2365-2366.
19. Miyamoto, S. (2012, November). An overview of hierarchical and non-hierarchical algorithms of clustering for semi-supervised classification. In *International Conference on Modeling Decisions for Artificial Intelligence* (pp. 1-10). Berlin, Heidelberg: Springer Berlin Heidelberg.
20. National Library of Medicine. ClinicalTrials.gov Glossary Terms. <https://clinicaltrials.gov/study-basics/glossary>
21. nQuery Sample Size Software.
https://en.wikipedia.org/wiki/NQuery_Sample_Size_Software
22. Pandis, N. (2016). The chi-square test. *American journal of orthodontics and dentofacial orthopedics*, 150(5), 898-899.
23. Piantadosi, S. (2024). *Clinical trials: a methodologic perspective*. John Wiley & Sons.
24. Pizarro, J., Guerrero, E., & Galindo, P. L. (2002). Multiple comparison procedures applied to model selection. *Neurocomputing*, 48(1-4), 155-173.
25. Pocock, S. J. (2013). *Clinical trials: a practical approach*. John Wiley & Sons.
26. Rayner, J. C., & Livingston Jr, G. C. (2023). *An introduction to Cochran-Mantel-Haenszel testing and nonparametric ANOVA*. John Wiley & Sons.

27. Roche (2025). Προϋποθέσεις για τη διεξαγωγή κλινικής δοκιμής. <https://www.roche.gr/innovation/clinical-trials/for-hcps/conditions>
28. Sakpal, T. V. (2010). Sample size estimation in clinical trial. *Perspectives in clinical research*, 1(2), 67-69.
29. Taylor, R. (1990). Interpretation of the correlation coefficient: a basic review. *Journal of diagnostic medical sonography*, 6(1), 35-39.
30. Thorat, S. B., Banarjee, S. K., Gaikwad, D. D., Jadhav, S. L., & Thorat, R. M. (2010). Clinical trial: a review. *Int J Pharm Sci Rev Res*, 1(2), 101-106.
31. Wright, B. (2017). Clinical trial phases. In *A comprehensive and practical guide to clinical trials* (pp. 11-15). Academic Press.
32. Κέντρο Ερευνών Νοσοκομείου Παπαγεωργίου. Διαδικασία Έγκρισης Κλινικών Μελετών. <https://www.rcph.gr/procedure/>
33. Χαλκιαδάκη, Ε., Γαλάνης, Π., & Λάμνισος, Δ. (2020). Ο νέος Ευρωπαϊκός Κανονισμός 536/2014 για τις κλινικές δοκιμές. *Archives of Hellenic Medicine/Arheia Ellenikes Iatrikes*, 37(1).
34. Gray, R. J. (1988). A class of K-sample tests for comparing the cumulative incidence of a competing risk. *The Annals of statistics*, 1141-1154.
35. Fine, J. P., & Gray, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American statistical association*, 94(446), 496-509.