

**Ερμηνεύοντας τα Αποτελέσματα των  
Τυχαιοποιημένων Κλινικών Μελετών:  
Προσέγγιση βασιζόμενη στην Κριτική Αξιολόγηση**

**Ηλίας Ζιντζαράς, MSc, PhD, PhD**  
Καθηγητής Βιομαθηματικών-Βιομετρίας

**Εργαστήριο Βιομαθηματικών, Τμήμα Ιατρικής,  
Πανεπιστήμιο Θεσσαλίας**

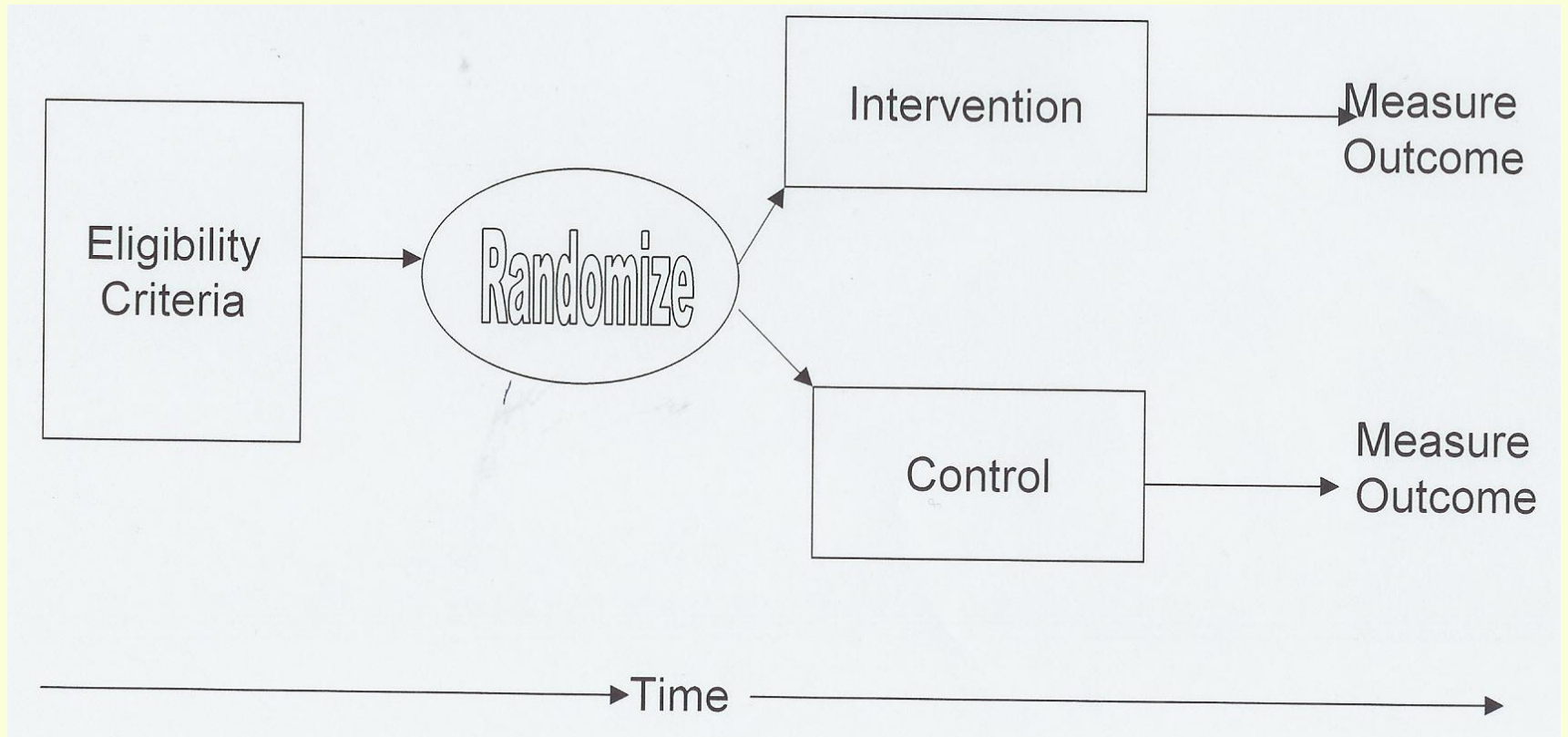
**Institute for Clinical Research and Health Policy Studies,  
Tufts Medical Center**

# Τι είναι Κριτική Αξιολόγηση (Critical Appraisal)?

Μία τεχνική:

- α) για να ελέγχουμε την **εγκυρότητα** και **ποιότητα** των δημοσιευμένων Τυχαιοποιημένων Κλινικών Μελετών (**RCTs**) και
- β) να αξιολογούμε συστηματικά τις RCTs ώστε τα αποτελέσματά τους να μπορούν να **εφαρμοστούν** στην καθημερινή **κλινική πράξη** (external validity).

# Τυχαιοποιημένη κλινική μελέτη



# Ιεράρχηση των ενδείξεων



## Κλινικό σενάριο (διδακτικό παράδειγμα)

- 75-χρονη με Διαβήτη τύπου II διεγνώσθη με **παροξυσμική κοιλιακή μαρμαρυγή (AF)**.
- Λόγω αυξημένου κινδύνου **εγκεφαλικού** επεισοδίου, ο θεράπων γιατρός της πρότεινε να ξεκινήσει **αντιπηκτική αγωγή με warfarin**.

## Κλινικό σενάριο

- Ωστόσο, η ασθενής δε φάνηκε πρόθυμη να ελέγχει το **INR** συχνά και ως εκ τούτου, ο ιατρός εκτίμησε ότι υπάρχει αυξημένη πιθανότητα μη συμμόρφωσης.
- Στη συνέχεια, ο ιατρός σκέφτηκε να χορηγήσει ένα νέο από του στόματος **αντιπηκτικό** παράγοντα, όπως το **dabigatran**, το οποίο δεν απαιτεί συχνό εργαστηριακό έλεγχο.
- Αποφάσισε όμως να συζητήσει το θέμα με τους συναδέλφους του στην κλινική.

## Κλινικό σενάριο

- Ο ιατρός παρουσίασε στους συναδέλφους του μια σειρά **κλινικών μελετών** οι οποίες υποστηρίζουν την **καταλληλότητα** του **dabigatran**.
- Όταν ρωτήθηκε αν οι μελέτες ήταν **τυχαιοποιημένες**, απαντά ότι δεν είναι σίγουρος.
- Τότε του προτείνουν να γίνει ανασκόπηση της σχετικής **βιβλιογραφίας** πριν ληφθεί οποιαδήποτε απόφαση.

## Βιβλιογραφική έρευνα

Ένα άρθρο θεωρήθηκε χρήσιμο αν:

- περιλαμβάνει ασθενείς με παροξυσμική **κολπική μαρμαρυγή**
- οι ασθενείς έχουν τυχαιοποιηθεί σε δύο ομάδες: **warfarin vs. dabigatran**
- η κύρια μεταβλητή (outcome) είναι η εμφάνιση **εγκεφαλικού** επεισοδίου



Η αναζήτηση έγινε στο **PubMed**  
χρησιμοποιώντας τις ακόλουθες λέξεις κλειδιά:

***“atrial fibrillation” and***

***“dabigatran” and***

***“randomized controlled trial”***

**41** άρθρα ανακτήθηκαν και μελετήθηκαν οι περιλήψεις

**1** άρθρο θεωρήθηκε κατάλληλο

**Dabigatran versus warfarin in patients with atrial fibrillation. Connolly SJ, Ezekowitz MD, Yusuf S, Eikelboom J, Oldgren J, Parekh A, Pogue J, Reilly PA, Themeles E, Varrone J, Wang S, Alings M, Xavier D, Zhu J, Diaz R, Lewis BS, Darius H, Diener HC, Joyner CD, Wallentin L; RE-LY Steering Committee and Investigators. N Engl J Med. 2009 Sep 17;361(12):1139-51. doi: 10.1056/NEJMoa0905561.**

## Το κατάλληλο άρθρο

Το άρθρο περιγράφει μία RCT όπου στην μία ομάδα χορηγείται το **dabigatran** και στην άλλη ομάδα η **warfarin**.

Η κύρια μεταβλητή για να αξιολογηθεί η αποτελεσματικότητά του **dabigatran** είναι η πρόληψη του εγκεφαλικού επεισοδίου.

## Το κατάλληλο άρθρο

Η RCTs αποτελείται από:

6075 ασθενείς που έλαβαν **dabigatran** 150mg

6021 ασθενείς που έλαβαν **warfarin**

## Αποτελέσματα

**dabigatran 150 mg: 134 ασθενείς (2%) με εγκεφαλικό**

**vs.**

**Warfarin: 199 ασθενείς (3%) με εγκεφαλικό**

**P-value <0.05**

# Προβληματισμός για τη αξία του άρθρου

**Είναι τα αποτελέσματα της μελέτης  
έγκυρα, ακριβή, και εφαρμόσιμα?**

**Τρεις ερωτήσεις χρειάζεται να απαντηθούν όταν αξιολογείται μία RCT:**

- 1. Είναι τα αποτελέσματα της μελέτης έγκυρα (valid)?**
- 2. Τι μέγεθος και ακρίβεια έχουν τα αποτελέσματα?**
- 3. Θα με βοηθήσουν τα αποτελέσματα στη φροντίδα των ασθενών μου?**

# 1. Είναι τα αποτελέσματα της μελέτης έγκυρα?

## Πρωτεύοντα Κριτήρια

- Ήταν το **μέγεθος** του δείγματος αρκετά μεγάλο (power analysis)?
- Έχουν κατανεμηθεί οι ασθενείς στις θεραπείες με **τυχαιοποίηση**?
- Ήταν κατάλληλη η **παρακολούθηση** (follow up) των ασθενών?
- Έχουν **αναλυθεί** οι ασθενείς με βάση τις ομάδες στις οποίες είχαν τυχαιοποιηθεί?



# 1. Είναι τα αποτελέσματα της μελέτης έγκυρα?

## Δευτερεύοντα Κριτήρια

- Ήταν «**τυφλή**» η μελέτη?
- Ήταν οι δύο συγκρινόμενες ομάδες (treatment & control) **όμοιες** στην αρχή της μελέτης?
- **Αντιμετωπίστηκαν** και οι δύο ομάδες με τον ίδιο τρόπο.

# 1. Είναι τα αποτελέσματα της μελέτης έγκυρα?

## Ήταν το μέγεθος του δείγματος αρκετά μεγάλο?

Το μέγεθος του δείγματος εξαρτάται από:

- την επιθυμητή ισχύ (**power**) (δηλ. την πιθανότητα να ανιχνεύσουμε μία διαφορά μεταξύ των ομάδων σε πιθανότητα λάθους  $P < 0.05$ , όταν αυτή υπάρχει πραγματικά).

Η ισχύς πρέπει να είναι περισσότερο από **80%**.

- την **διαφορά** μεταξύ των δύο ομάδων (treatment vs control) που μας ενδιαφέρει να ανιχνεύσουμε.
- την **μεταβλητότητα** της κύριας μεταβλητής.

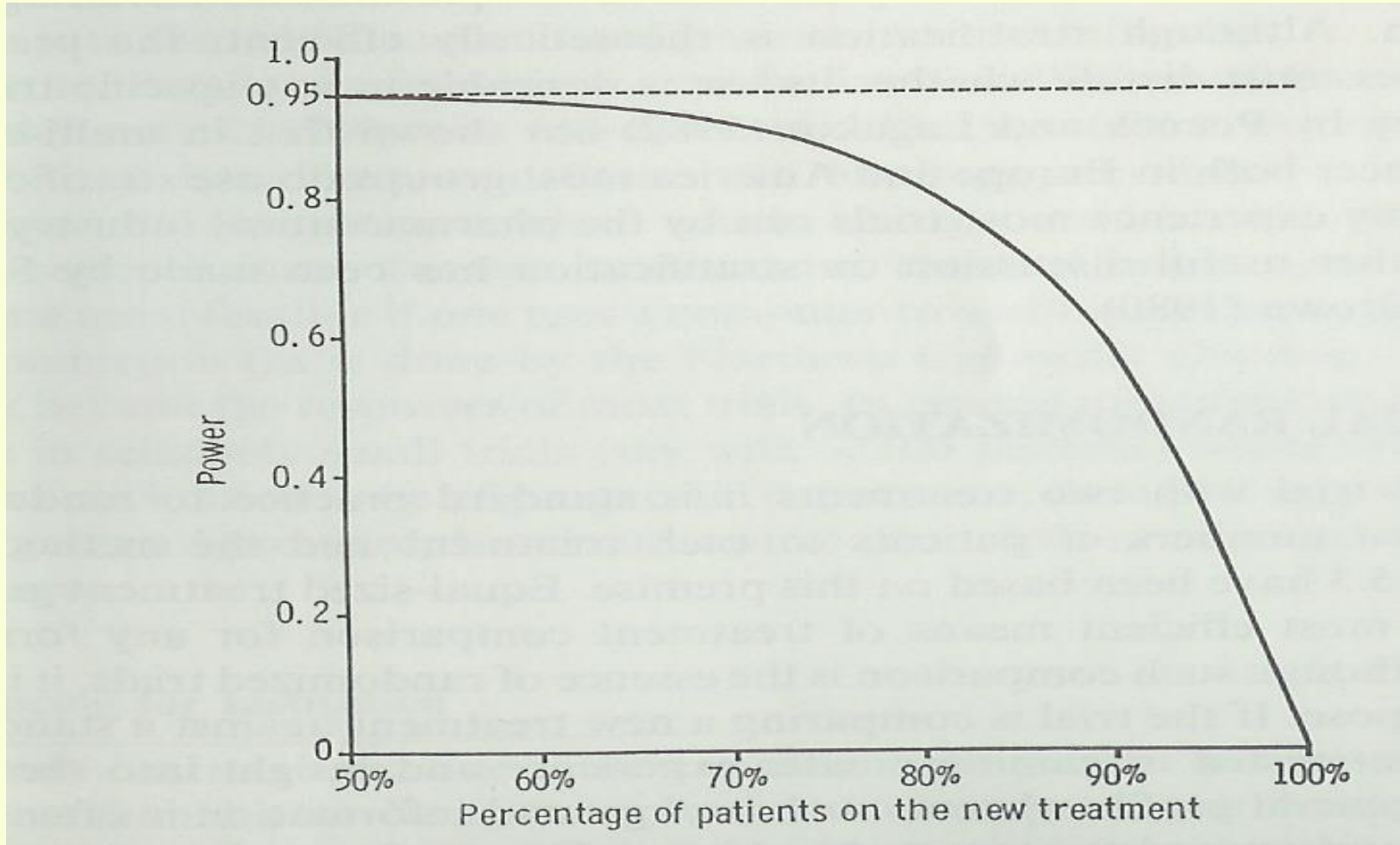
# 1. Είναι τα αποτελέσματα της μελέτης έγκυρα?

Έχουν κατανεμηθεί οι ασθενείς στις  
θεραπείες με τυχαιοποίηση?

- Η τυχαιοποίηση εξασφαλίζει ότι γνωστοί και άγνωστοι **συγχυτικοί** (confounders) παράγοντες της κύριας μεταβλητής (εμφάνιση εγκεφαλικού) **κατανέμονται ομοιόμορφα** μεταξύ των δύο ομάδων.
- Οι μη-τυχαιοποιημένες μελέτες μπορεί να παράγουν ψευδή θετικά αποτελέσματα (false-positive results) όσον αφορά την αποτελεσματικότητα της θεραπείας

## Ανισοκατανεμημένη τυχαιοποίηση

Όσο μεγαλώνει η διαφορά μεγέθους μεταξύ των ομάδων τόσο ελαττώνεται η ισχύς (power) της μελέτης



# 1. Είναι τα αποτελέσματα της μελέτης έγκυρα?

Ήταν κατάλληλη η παρακολούθηση (follow-up) των ασθενών?

Αν ένας σημαντικός αριθμός ασθενών έχει χαθεί από την παρακολούθηση (**lost-to-follow-up**), η εγκυρότητα της μελέτης αμφισβητείται.

Οι ασθενείς είναι δυνατόν να χαθούν επειδή:

- αποσύρθηκαν λόγω δυσμενούς κλινικής έκβασης (**withdrawal**)
- δεν ήθελαν να συμμορφωθούν με τους κανόνες της μελέτης (**non-compliance**)

# 1. Είναι τα αποτελέσματα της μελέτης έγκυρα?

Έχουν αναλυθεί οι ασθενείς με βάση τις ομάδες στις οποίες έχουν τυχαιοποιηθεί?

- **Intent-to-treat (ITT) ανάλυση:** περιλαμβάνει την ανάλυση των δεδομένων από όλους τους ασθενείς ασχέτως αν έχουν χαθεί κατά τη διάρκεια της μελέτης ή άλλαξαν θεραπεία.
- **Per protocol (PP) analysis:** περιλαμβάνει μόνο τα δεδομένα από τους ασθενείς που ολοκλήρωσαν την μελέτη σύμφωνα με το πρωτόκολλο.

**ITT και PP (per protocol) αναλύσεις πρέπει να δίνουν παρόμοια αποτέλεσμα, διαφορετικά η ITT είναι πιο αποδεκτή.**

# 1. Είναι τα αποτελέσματα της μελέτης έγκυρα?

Ήταν οι ασθενείς, γιατροί, νοσηλευτές και εργαζόμενοι στη μελέτη «τυφλοί» ως προς τις χορηγούμενες θεραπείες?

Η έλλειψη τυφλότητας, ελαττώνει την αξιοπιστία των αποτελέσμων της μελέτης και δημιουργεί συστηματικό σφάλμα (**bias**)

# 1. Είναι τα αποτελέσματα της μελέτης έγκυρα?

## Ήταν οι ομάδες όμοιες στην αρχή της μελέτης?

- Οι **προγνωστικοί** παράγοντες πρέπει να είναι **παρόμοιοι** για τις δύο ομάδες (π.χ. φύλο, ηλικία).
- Η **τυχαιοποίηση** δεν επιτυγχάνει πάντα την ισοκατανομή των γνωστών προγνωστικών παραγόντων στις δύο ομάδες.
- Υπάρχουν στατιστιστικές τεχνικές που λαμβάνουν υπόψη (**adjustment**) διαφορές στην αρχή της μελέτης.



# 1. Είναι τα αποτελέσματα της μελέτης έγκυρα?

**Αντιμετωπίστηκαν και οι δύο ομάδες με τον ίδιο τρόπο?**

Αν η μία ομάδα (πχ με **dabigatran**) έλαβε **εντατικότερη παρακολούθηση** τότε

- είναι πιο πιθανότερο να καταγραφούν τυχόν **μεταβολές** στην ανταπόκριση αυτών των ασθενών
- οι ασθενείς αυτοί μπορεί να λάβουν θεραπείες που δεν περιλαμβάνονταν στο πρωτόκολλο (**co-intervention**).

Πιθανή διαστρέβλωση αποτελεσμάτων (**bias**).

# Κλινικό σενάριο

## 1. Είναι τα αποτελέσματα της μελέτης έγκυρα?

Πόσο καλά η μελέτη με το **dabigatran** σε ασθενείς με παροξυσμική κοιλιακή μαρμαρυγή εκπληρώνει τα κριτήρια εγκυρότητας?

# Κλινικό σενάριο

## 1. Είναι τα αποτελέσματα της μελέτης έγκυρα?

### Πρωτεύοντα κριτήρια:

- Η μελέτη είχε μεγάλη ισχύ >80%.
- Η τυχαιοποίηση έχει διεξαχθεί με αυστηρότητα.
- Η παρακολούθηση (follow-up) των ασθενών ήταν κατάλληλη (<15% των ασθενών χάθηκε).
- Η ανάλυση ήταν ITT.

### Δευτερεύοντα κριτήρια:

- Υπήρχε τυφλότητα.
- Οι δύο ομάδες ήταν παρόμοιες στην αρχή της μελέτης.
- Δεν υπάρχουν πληροφορίες σχετικά με τον αν αντιμετωπίστηκαν και οι δύο ομάδες με τον ίδιο τρόπο. **X**

**Η μεθοδολογία σε αυτή τη μελέτη ήταν  
γενικά έγκυρη**

Μπορούμε με ασφάλεια να συμπεράνουμε  
ότι η χορήγηση **dabigatran** στον ασθενή  
με παροξυσμική κοιλιακή μαρμαρυγή είναι  
κατάλληλη?

**Όχι ακόμη!**

**Πρέπει να απαντήσουμε σε δύο ακόμη ερωτήσεις**

**2. Τι μέγεθος και ακρίβεια έχουν τα αποτελέσματα?**

**3. Θα με βοηθήσουν τα αποτελέσματα στη φροντίδα των ασθενών μου?**

## 2. Τι μέγεθος και ακρίβεια έχουν τα αποτελέσματα?

**Πόσο μεγάλο ήταν το θεραπευτικό αποτέλεσμα?**

Στις RCTs, συνήθως η κύρια μεταβλητή είναι δυαδική (ναι εγκεφαλικό / όχι εγκεφαλικό).

Τότε η αποτελεσματικότητα της θεραπείας μετράται με το odds ratio (OR).

Στην μελέτη της κοιλιακής μαρμαρυγής, 2% των ασθενών με **dabigatran** εμφάνισε εγκεφαλικό και με το **warfarin** το 3% των ασθενών εμφάνισε εγκεφαλικό.

- Κίνδυνος (odds) με **dabigatran**:  $Y=0.02$
- Κίνδυνος (odds) με **warfarin**:  $X=0.03$

Οπότε, το αποτέλεσμα της θεραπείας μπορεί να εκφρασθεί ως:

- Odds Ratio (**OR**)= $(134/(6075-134))/(199/(6021-199))=0.65$ 
  - Η «πιθανότητα» εγκεφαλικού με **dabigatran** είναι 35% μικρότερη σχέση με **warfarin**.



## 2. Τι μέγεθος και ακρίβεια έχουν τα αποτελέσματα?

### Πόσο ακριβής ήταν η εκτίμηση του θεραπευτικού αποτελέσματος?

- Η ακρίβεια του OR εκτιμάται από το 95% διάστημα εμπιστοσύνης (δε), το οποίο προσδιορίζει το διάστημα που μπορεί να κυμανθεί το OR με 95% βεβαιότητα.
- Το 95% δε είναι ισοδύναμο με τη στατιστική σημαντικότητα  $p < 0.05$  (error probability).

Στη μελέτη του **dabigatran** το 95% δε είναι

**(0.53, 0.82)**

Το OR είναι σημαντικό: το 1 δεν συμπεριλαμβάνεται στο 95% δε

Σε μία υποθετική μελέτη ήταν:

Ομάδα με **dabigatran**: 600 ασθενείς και 20 με εγκεφαλικό

Ομάδα με **warfarin**: 600 ασθενείς και 15 με εγκεφαλικό

Κίνδυνος (odds) με **dabigatran**:  $Y=0.02$

Κίνδυνος (odds) με **warfarin**:  $X=0.03$

Όπως και στην πραγματική μελέτη!

**Τότε στην υποθετική μελέτη:**

**OR= 0.65**

**95% δε (0.29, 1.40)**

**Το OR δεν είναι σημαντικό επειδή το 1 συμπεριλαμβάνεται στο 95% δε**

**Η μελέτη δεν μπορεί να μας βοηθήσει για να αποφασίσουμε αν πρέπει να χορηγήσουμε τη θεραπεία.**

**Σε μία μελέτη με μεγάλο αριθμό ασθενών υπάρχει μεγαλύτερη πιθανότητα υπάρχει να παρατηρήσουμε ένα στατιστικά σημαντικό αποτέλεσμα.**

## Επιφυλακτικότητα στα στατιστικά tests και στα P-values

- Προσοχή στα **αόριστα test** (π.χ. anova) ή στη χρήση παραμετρικών tests όταν τα δεδομένα δεν τα δικαιολογούν.
- Προσοχή σε **στατιστικά σημαντικά** αποτελέσματα που **δεν** είναι **κλινικά σημαντικά**.

## Επιφυλακτικότητα στα στατιστικά tests και στα P-values

- Το δε είναι πάντα καλύτερο από μία τιμή P-value.
- Το μέγεθος των P-values δεν είναι ανάλογο του μεγέθους της θεραπευτικής αποτελεσματικότητας:
  - το  **$P < 0.001$** , συγκρινόμενο με το  **$P < 0.05$** ,  
**δεν σημαίνει μεγαλύτερο θεραπευτικό αποτέλεσμα.**
- Σημαίνει ότι είναι **πιο πιθανό να υπάρχει διαφορά** μεταξύ των δύο ομάδων.

**Η μελέτη έχει αποδεκτό μέγεθος και ακρίβεια όσον αφορά την αποτελεσματικότητά του**

Μπορούμε με ασφάλεια να συμπεράνουμε  
ότι η χορήγηση **dabigatran** στον ασθενή  
με παροξυσμική κοιλιακή μαρμαρυγή είναι  
κατάλληλη?

**Όχι ακόμη!**

**Πρέπει να απαντήσουμε και στην τελευταία ερώτηση**

**3. Θα με βοηθήσουν τα αποτελέσματα στη φροντίδα των ασθενών μου?**



**3. Θα με βοηθήσουν τα αποτελέσματα στη φροντίδα των ασθενών μου?**

**Μπορούν τα αποτελέσματα να εφαρμοσθούν για την θεραπεία των ασθενών μου?**

**Πόσο διαφορετικοί είναι οι ασθενείς στη μελέτη μα αυτούς στο ιατρείο μου?**

**Γενικά, δεν μπορούμε να γενικεύσουμε τα αποτελέσματα των RCTs στον γενικό πληθυσμό των αρρώστων.**

Οι RCTs διεξάγονται με **συγκεκριμένο χρόνο παρακολούθησης**, οι ασθενείς πληρούν καθορισμένα **κριτήρια ένταξης- αποκλεισμού**, ενώ υπάρχει και συνεχής **επιτήρηση** της διεξαγωγής τους.

**Ιδανική περίπτωση!**

Για την εξαγωγή ασφαλέστερων συμπερασμάτων όσον αφορά στη σχετική αποτελεσματικότητα και ανεκτικότητα του **dabigatran** στην κολπική μαρμαρυγή, χρειάζεται να ληφθούν υπόψη και τα αποτελέσματα μακροπρόθεσμων **μελετών παρατήρησης**.