



**ΜΠΣ Βιοστατιστικής**

*Αθήνα, Εαρινό Εξάμηνο 2015*

## **Μάθημα: Κλινικές Δοκιμές**

Βασισμένο στο μάθημα του Καθηγητή Steve Lagakos, και σε σημειώσεις των R. Gelber, K. Stanley & J. Ware.

Ουρανία Δαφνή  
Καθηγήτρια, Διευθύντρια  
Εργαστήριο Βιοστατιστικής  
Τμήμα Νοσηλευτικής  
Πανεπιστήμιο Αθηνών

Διευθύντρια  
Frontier Science Foundation-Hellas (FSF-H)



---

# Μάθημα: Κλινικές Δοκιμές – Πρώτη Διάλεξη

*Αθήνα, 16 Μάρτη 2015*

- Στόχοι Μαθήματος
- Εισαγωγή
- Τύποι Κλινικών Δοκιμών
- Σχεδιασμός Κλινικών Δοκιμών

---

# Πρόγραμμα Κλινικών Δοκιμών 2015

16/03 θεωρία 4-7 μ.μ.

17/03 θεωρία 4-7 μ.μ.

18/03 άσκηση 4-6 μ.μ.

23/03 θεωρία 4-7 μ.μ.

26/03 άσκηση 4-6 μ.μ.

30/03 θεωρία 4-7 μ.μ.

31/03 θεωρία 4-7 μ.μ.

01/04 άσκηση 4-6 μ.μ.

## ΠΑΣΧΑ

21/04 θεωρία 4-7 μ.μ.

22/04 άσκηση 4-8 μ.μ.

27/04 θεωρία 4-8 μ.μ.

28/04 θεωρία 4-8 μ.μ.

---

# Μάθημα: Κλινικές Δοκιμές

## ■ Διδάσκουσα

Καθ. Ουρανία Δαφνή

[Urania Dafni \[udafni@nurs.uoa.gr\]](mailto:Urania.Dafni@nurs.uoa.gr)

[udafni@frontier-science.gr](mailto:udafni@frontier-science.gr)

210-7710902

210-7461456

## ■ Βοηθός Διδασκαλίας

Γεωργία Βουρλή

Georgia Vourli [gvourli@med.uoa.gr]

210-7462030

---

# Μάθημα: Κλινικές Δοκιμές

- Διαλέξεις
- Εργαστηριακές Ασκήσεις
- Παρουσιάσεις ομάδων φοιτητών μίας δημοσίευσης από την πρόσφατη ιατρική βιβλιογραφία
  
- Εστίαση σε παραδείγματα
- **Στόχος Μαθήματος:** Εισαγωγή στις έννοιες, τις αρχές και τις μεθόδους που χρησιμοποιούνται στις **Κλινικές Δοκιμές**

---

## Κλινικές Δοκιμές - Θέματα

- Εισαγωγή στις Κλινικές Δοκιμές, Σχεδιασμός
- Μελέτες Υπεροχής – Ισοδυναμίας
- Ηθικά Θέματα
- Μεροληψία και Ακρίβεια
- Πρωτόκολλο Κλινικής Δοκιμής
- Μέγεθος Δείγματος (εργαστήριο)
- Ενδιάμεσες Αναλύσεις και εργαστήριο
- Ανάλυση Δεδομένων
- Κριτική αξιολόγηση ιατρικών δημοσιεύσεων
- Παρουσιάσεις Φοιτητών

---

## Παρουσιάσεις Φοιτητών: Papers

1. Paclitaxel plus Bevacizumab versus Paclitaxel Alone for Metastatic Breast Cancer (Miller et al, NEJM, 357:2666-76, 2007).
2. Treatment of Hypertension in Patients 80 Years of Age or Older (Beckett et al, NEJM, 358:1887-98, 2008)
3. Telmisartan, Ramipril, or Both in Patients at High Risk for Vascular Events (The ONTARGET Investigators, NEJM, 358:1547-59, 2008).
4. Effect of Herpes Simplex Suppression on Incidence of HIV among Women in Tanzania (Watson-Jones, et al, NEJM, 358:1560-71, 2008).
5. Vasopressin versus Norepinephrine Infusion in Patients with Septic Shock (Russell et al, NEJM, 358: 877-87, 2008).

---

## Παρουσιάσεις Φοιτητών: Papers

6. Risk of cesarean delivery and neuraxial analgesia (Wong et al, NEJM, 352: 655-665, 2005)
7. Calcium plus Vitamin D and fracture risk (Jackson et al, NEJM, 354:669-683, 2006)
8. Surgical versus Nonsurgical Therapy for Lumbar Spinal Stenosis (Weinstein et al, NEJM, 358:794-810, 2008)
9. Thrombus Aspiration during Primary Percutaneous Coronary Intervention (Svilaas et al, NEJM, 358: 557-67, 2008)
10. Telbivudine versus Lamivudine in Patients with Chronic Hepatitis B (Lai et al, NEJM, 357: 2576-88, 2007).

**Θα δοθεί περίληψη θεμάτων για κάθε δημοσίευση**



---

## Μάθημα: Κλινικές Δοκιμές

- Σημειώσεις Διαλέξεων
- Βιβλιογραφικές αναφορές
- Βιβλία (προαιρετικά):
  - S. Pocock: Clinical Trials-A Practical Approach, Wiley, 1983
  - S. Piantadosi; Clinical Trials; Wiley, 1997
  - ICH-E10 guidelines  
(<http://www.fda.gov/cder/guidance/4155fnl.htm>)



ΜΠΣ Βιοστατιστικής - Μάθημα: Κλινικές Δοκιμές

## Σχεδιασμός Κλινικών Δοκιμών (& Τύποι Κλινικών Δοκιμών)

Ουρανία Δαφνή  
Καθηγήτρια, Διευθύντρια  
Εργαστήριο Βιοστατιστικής  
Πανεπιστήμιο Αθηνών

Διευθύντρια  
Frontier Science Foundation-Hellas (FSF-H)



---

## Στόχος Βιοστατιστικής

- Εξαγωγή συμπερασμάτων σε σχέση με μία αρχική υπόθεση, βιολογικής ή κλινικής φύσης

*ΠΑΡΑΤΗΡΗΣΗ-ΔΕΙΓΜΑ (ΚΛΙΝΙΚΗ ΔΟΚΙΜΗ) →  
→ ΑΠΟΔΕΙΞΗ → ΣΥΜΠΕΡΑΣΜΑ →  
→ ΠΛΗΘΥΣΜΟΣ*

- Βασίζεται στην ορθή εφαρμογή των μεθόδων της Μαθηματικής Επιστήμης μέσω της Θεωρίας Πιθανοτήτων και των Αρχών της Στατιστικής

---

## Σκοπός και στόχος

- Να συζητήσουμε διάφορες βασικές στατιστικές έννοιες που εμπλέκονται και αφορούν το σχεδιασμό και την ερμηνεία των κλινικών δοκιμών
- Να δούμε κάποια παραδείγματα από τη βιβλιογραφία για την καλύτερη κατανόηση των εννοιών

# Τι αντιμετωπίζουν οι επιστήμονες υγείας;

- την ανάγκη για ορθή και ακριβή ερμηνεία διαφόρων διαγνωστικών μεθόδων που χρησιμοποιεί στην κλινική πράξη
- την αμφιβολία για την αποτελεσματικότητα των θεραπευτικών παρεμβάσεων
- την ανάγκη για ισορροπία μεταξύ του οφέλους και της ζημίας που μπορεί να προκαλέσει μια προτεινόμενη φαρμακευτική αγωγή
- την αγωνία για την πορεία της νόσου στους ασθενείς του και το κόστος των παρεμβάσεων του (φάρμακα, διαγνωστικά κριτήρια, διαδικασίες)

---

# Η επιστήμη των αποδείξεων

- Η επιστήμη των αποδείξεων (κατά άλλους ενδείξεων, μετάφραση του αγγλικού όρου *evidence based*) είναι μια επιστημονική προσέγγιση στην καθημερινή κλινική πράξη κατά την οποία ο επιστήμονας του χώρου της υγείας είναι **ενήμερος γύρω από τις μελέτες (κλινικές και επιδημιολογικές)** που υπάρχουν γύρω από το θέμα που διαπραγματεύεται και οι οποίες υποστηρίζουν την κλινική του απόφαση για μια ιατρική πράξη.

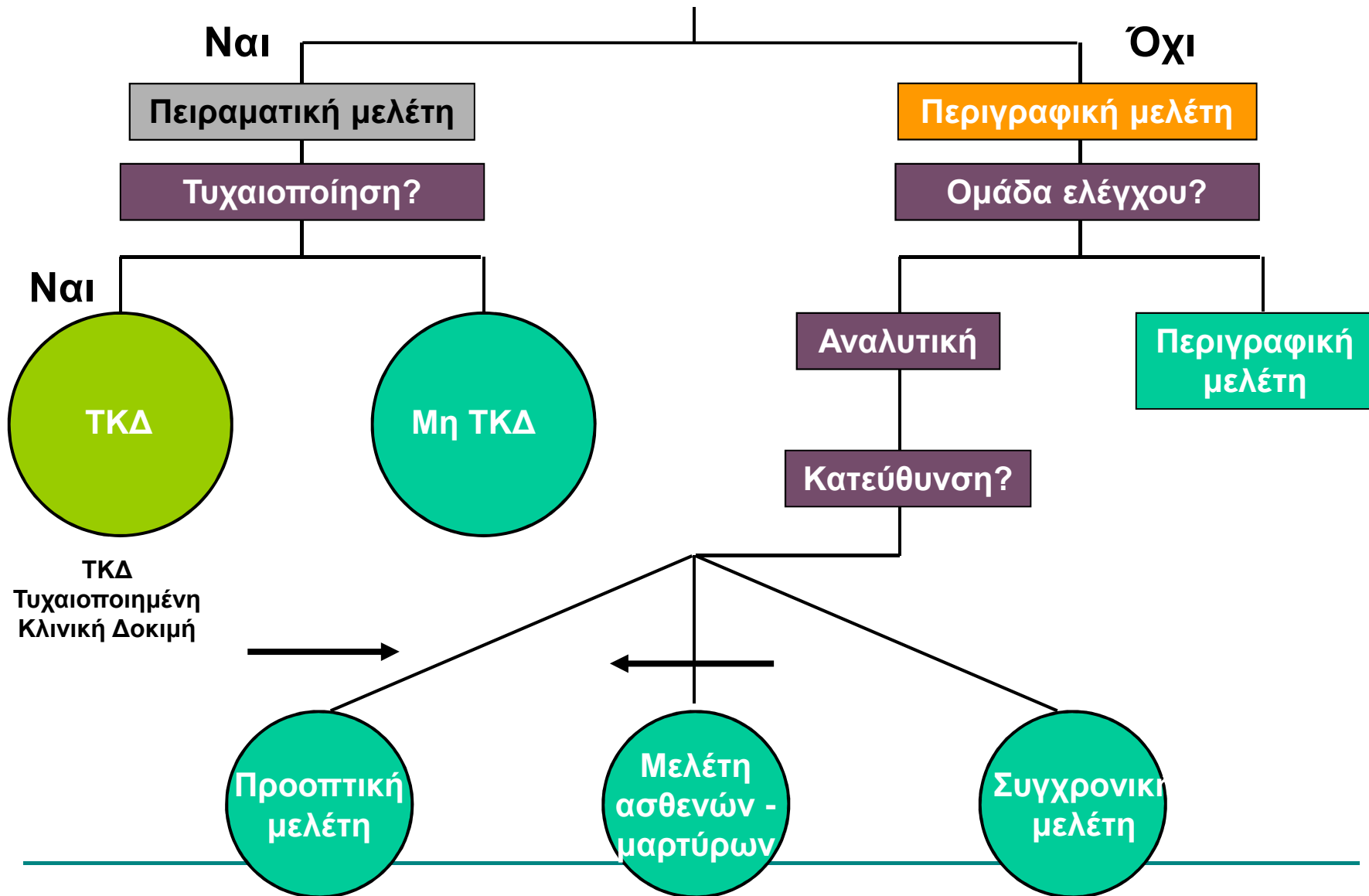
---

# Η επιστήμη των αποδείξεων

Τα βασικά βήματα της επιστήμης των αποδείξεων είναι:

1. Σαφής διατύπωση της κλινικής ερώτησης - προβλήματος
2. Εμπεριστατωμένη διερεύνηση της βιβλιογραφίας σε έγκυρα επιστημονικά περιοδικά και ειδικότερα στις μετα-αναλύσεις (*χρήση αντικειμενικών ποσοτικών μεθόδων για τη σύνοψη δεδομένων κατά τρόπο αντικειμενικό και αδιάβλητο*)
3. Επιλογή των κατάλληλων άρθρων (*αυτών που κρίνεται ότι έχουν τηρήσει τις αρχές της μεθοδολογίας της έρευνας*)
4. Κριτική αξιολόγηση των ευρημάτων των ερευνητικών άρθρων που επελέγησαν
5. Εφαρμογή των πορισμάτων της προαναφερθείσας προσέγγισης στους ασθενείς

# Είδη Μελετών: Υπάρχει ομάδα παρέμβασης;





---

# Types of Clinical Investigation

Case Reports: detailed description of 1/few patients

Case Series: similar to above, but with more patients

Database Analyses: typically, on existing databases

Observational Studies (no treatment interventions imposed)

Cross-sectional Studies: Observe each patient at 1 time point

Case-Control Studies: Subjects selected by disease status and compared with controls for possible causes.

Cohort Studies: Subjects observed over time. Cohort studies can focus on treatment as a determinant of course of disease.

Controlled Clinical Trials: The treatment assignment is by design. Endpoints and analyses planned in advance

---

---

# Examples of Observational Studies

- Cross Sectional Study: Examine a group of teenage school children for nutritional status and academic performance to assess associations.
- Case-Control Study: Select 100 women recently diagnosed with breast cancer in your hospital; select 100 'control' women matched for age, race, and educational status. Survey both groups to ascertain information on use of oral contraceptives (OC) earlier in their lives. Are there higher rates of OC use in group with breast cancer?
- Cohort Study: Identify 500 homosexual men aged 18-30. Ascertain rates and types of sexual activity and HIV status every 6 months for 5 years. How does the risk of becoming infected with HIV depend on the specific type of sexual behavior?
- => Observational studies often aim to identify causes or predictors of disease or some other clinical outcome.

---

## Ιεράρχηση τεκμηρίων

**Μετά - ανάλυση**

**Τυφλά Τυχαιοποιημένες Κλινικές Δοκιμές**

**Τυχαιοποιημένες Κλινικές Δοκιμές**

**Προοπτικές μελέτες**

**Αναδρομικές μελέτες: ασθενών-μαρτύρων**

**Συγχρονικές μελέτες – με ομάδα ελέγχου**

**Περιγραφικές μελέτες – χωρίς ομάδα ελέγχου**

**Σειρά παρατηρήσεων – χωρίς ομάδα ελέγχου**

**Σποραδικές παρατηρήσεις**

---

# Τι είναι μια κλινική δοκιμή

## «Κλασσικός» ορισμός

Μια προοπτική μελέτη παρέμβασης σε ομάδα ανθρώπων, η οποία συγκρίνει τα ευρήματα αναφορικά με την έκβαση μιας νόσου ή μιας κατάστασης, με μια άλλη ομάδα ελέγχου

- Ερευνητική μελέτη σε ανθρώπους
- Εξετάζει την ορθότητα μιας ερευνητικής υπόθεσης
- Προτείνει τρόπους για πρόληψη ή θεραπεία
- Κατάλληλη για αιτιολογικές συσχετίσεις

---

# Other Definitions of a Clinical Trial

A clinical trial is a scientific research activity in human subjects undertaken to determine prospectively the effect and value of preventive, diagnostic, and therapeutic agents, devices, regimens, and procedures.

Hopwood MD, Mabry JC, Sibley WL;  
Rand Corporation, 1980

A clinical trial is any form of planned experiment which involves patients and is designed to elucidate the most appropriate treatment of future patients.

Trials: A Practical Approach.  
Wiley, 1983.

Pocock SJ, Clinical

---

## Other Definitions (continued)

A scientific experiment that generates clinical data for the purpose of evaluating one or more therapies on a patient population. Implicit in the trial is the fact that the clinical investigator has control of the process by which a treatment assigned to a patient.

Gelber RD, Zelen M: Planning and reporting of clinical trials.  
In: Calabresi P (ed). Medical Oncology. Macmillan, 1985.

---

# Clinical Trials Possess Important Properties of the Scientific Method

- Objective and Reproducible Measurement
- Externalizing of Plans and Procedures
- Control of Extraneous Factors
- Submission of Work for External Review and Replication or Disproof

---

# The Principles of Design of Clinical Trials are Easily Stated

- Specify a well-defined population of patients
- Enroll every eligible patient
- Randomly assign study treatment regimens to enrolled patients
- Complete the treatment of every patient according to the protocol



---

## The Principles of Design of Clinical Trials are Easily Stated (continued)

- Retain every patient in follow-up and complete every scheduled follow-up visit and examination
- Include every randomized patient in the analysis
- Use pre-specified analyses to draw conclusions about study hypotheses

# CTs are Not Idealized Experiments

## Idealized Experiment

Experimental units are identical

Treatment exactly reproducible from occasion to occasion

Experimental units are controlled by investigator

Measurement error the only source of variability other than effects of treatments

Measurement error small

## Clinical Trial

Experimental units are human subjects; free to refuse participation

Treatments not exactly reproducible

Patients may withdraw

Not all factors affecting outcome are controlled, or even known

Measurement errors and patient heterogeneity often large relative to treatment effects

---

# Experiments Answer a Scientific Question by Isolating the Intervention and the Outcome from Extraneous Influences

- **Goals** (more later):
  - Eliminate systematic error (Bias)
  - Minimize random error (Precision)
  - Ensure the generalizability of study results
- **Study Design** is the methodology for achieving these goals

---

## How Are These Goals Achieved?

- Specify patient population and method of sampling in a generalizable way
- Pre-specify study hypotheses using clear criteria and endpoints
- Standardize diagnostic, staging, assessment, and follow up procedures
- Describe the treatment protocol

---

## How Are These Goals Achieved (cont)?

- Utilize statistical design to minimize bias and variability
  - Randomization, stratification, blinding, choice of design
- Establish a sample size sufficient to achieve study goals

---

# Clinical Trials are Part of a Family of Research Methods

- Studies of disease mechanisms
- Controlled studies of disease or health care management (most clinical trials)
- Descriptive studies of clinical populations
- Studies of technology or diagnostic procedures
- Observational studies of health care delivery

---

## Δημοσιεύσεις

Οι επιστήμονες υγείας πρέπει να βασίζονται στη δημοσιευμένη αρθρογραφία για θεραπευτικές αποφάσεις

Τελικό Μέτρο Σημαντικότητας Κλινικής Δοκιμής:  
**Επίπτωση πάνω στην Κλινική Πράξη**

---

# OBJECTIVE

The purpose of biomedical research involving human subjects must be to improve diagnostic, therapeutic and prophylactic procedures and the understanding of the etiology and pathogenesis of disease.

(Declaration of Helsinki)



---

# DEFINITION OF A CLINICAL TRIAL

A clinical trial is a scientific experiment involving human subjects in which treatment is initiated specifically for therapy evaluation.

## ELEMENTS OF A CLINICAL TRIAL

Objectives  
Patient Population  
Treatment Options  
Design  
Conduct  
Endpoints

---

# Clinical Trials Need Not Involve Randomization

- Uncontrolled trials
  - Historical controls
  - Concurrent nonrandomized controls
  - Randomized controls
- Each design has its role in clinical investigation.  
These designs represent a hierarchy in terms of  
degree of control

---

# Είδη Κλινικών Δοκιμών

## ■ *Φάσης I*

- Η φαρμακοκινητική και οι βιολογικές επιδράσεις του φαρμάκου εξετάζονται για πρώτη φορά σε ανθρώπους
- Στόχοι: Διερεύνηση Ασφάλειας  
Καθορισμός Ασφαλούς Δόσης

## ■ *Φάσης II ή I/II*

- Μικρές μελέτες ασφάλειας και αποτελεσματικότητας
- Στόχος: Επιλογή λίγων δραστικών θεραπειών για περαιτέρω μελέτη σε δοκιμή Φάσης III

---

# Είδη Κλινικών Δοκιμών

## ■ *Φάσης III*

- Μεγάλου μεγέθους συγκριτικές μελέτες της πειραματικής θεραπείας έναντι της καθιερωμένης θεραπείας ή εφόσον δεν υπάρχει τέτοια θεραπεία, έναντι placebo (*εικονική θεραπεία*)
- Στόχοι: Διερεύνηση αποτελεσματικότητας  
Έγκριση Φαρμάκου

## ■ *Φάσης IV*

- Μελέτες παρατήρησης σε μεγάλο αριθμό ασθενών μετά την κυκλοφορία του φαρμάκου στην αγορά
- Στόχος: Αναγνώριση πιθανών ανεπιθύμητων ενεργειών

---

## Phase II Drug Trials

- 1) Initial dose is from Phase I study(ies).
- 2) Doses usually not escalated; may be lowered for a patient if the patient has toxicity (but that lowered dose does not affect anyone else's dose).
- 3) In diseases where nothing has worked, it is becoming standard practice to treat 14 to 20 patients in the first, "stage" of the study (two-stage Phase II Trials):  
Example: Treat 14 patients. If none respond, then stop!  
Prob (0 responses out of 14|true response rate = 20%)  $\leq 0.05$ . If at least one responds, add N more patients, then stop and do a conditional confidence interval on the response rate.  
(N usually equals 10 to 25 patients)
- 4) For diseases where many things work, Phase II study may be a short time on a new drug before a switch to a drug which is known to work.

---

## Phase II Drug Trials (Cont)

5) Response or other efficacy measure is the main endpoint, but gathering additional toxicity data is still important.

6) May involve “slightly” different patient population than Phase I study.

Usually have to have a disease in which response (however defined) can be objectively assessed.

7) Usually restricted to a single disease. Separate Phase II studies are done in each disease that might benefit from the treatment.

---

# Phase III Studies-Randomized Trials of Treatment ‘Z’: many possible designs

- Direct tests (usually when no known treatment for this disease)
  - Z versus placebo (P)
- Active control ‘W’ (standard treatment for this disease)
  - Z versus W
- Fairly direct tests of value of adding Z to a standard therapy ‘Y’
  - Y + Z versus Y + P
- Amount, Timing of ‘Z’
  - Low dose versus high dose of Z
  - Z initially versus Z delayed
  - Z intermittently versus Z continuously

## Φάσης III - Τυχαιοποίηση-Τυφλότητα

- Οι ομάδες πρέπει ιδανικά να διαφέρουν μόνο ως προς την παρέμβαση
- Η κατανομή των θεραπειών στις ομάδες σύγκρισης πρέπει να γίνεται με τρόπο «τυφλό»
- Με την προϋπόθεση ότι το μέγεθος της μελέτης είναι αρκετά μεγάλο, **η τυχαιοποίηση διασφαλίζει** ότι οι γνωστοί και οι άγνωστοι συγχυτικοί παράγοντες κατανέμονται εξίσου στις ομάδες σύγκρισης
- ΑΡΑ: Η διαφορά στην έκβαση μπορεί να αποδοθεί στις διαφορετικές θεραπευτικές αγωγές



---

# Σχεδιασμός Κλινικών Δοκιμών

Μία καλά σχεδιασμένη κλινική δοκιμή έχει τα ακόλουθα χαρακτηριστικά:

- Σαφώς καθορισμένους στόχους
- Καλώς ορισμένες τελικές εκβάσεις που προκύπτουν από τους τεθέντες στόχους
- Εκ των προτέρων καθορισμένα κριτήρια επιτυχίας ή αποτυχίας της πειραματικής θεραπείας βάσει στατιστικών ελέγχων επί των τελικών εκβάσεων
- Σαφή υπολογισμό του μεγέθους δείγματος και της αντίστοιχης στατιστικής ισχύος
- Πλήρη περιγραφή των κριτηρίων ένταξης και αποκλεισμού ασθενών - τυχαιοποίηση

---

# Στόχοι έναντι Τελικών Εκβάσεων (Μέτρα Αποτελεσματικότητας)

Τελική Έκβαση = Μετρήσιμο Μέγεθος

- ❑ Μέση Τιμή Συνεχούς Μεταβλητής
- ❑ Ποσοστό Επιτυχίας
- ❑ Χρόνος μέχρι το γεγονός ενδιαφέροντος

---

## Μέγεθος Δείγματος

- Ο πειραματισμός με περισσότερα άτομα από τον αναγκαίο αριθμό παραβιάζει τα ατομικά ανθρώπινα δικαιώματα των ασθενών
- Ο πειραματισμός με λιγότερα άτομα από τον αναγκαίο αριθμό παραβιάζει τα δικαιώματα των ασθενών στην κοινότητα

---

# Μέγεθος Δείγματος

- **Επαρκές δείγμα:** Μέγεθος Μελέτης ικανό για να αναγνωριστεί με μεγάλη πιθανότητα (Στατιστική Ισχύς) η ελαχίστη κλινικά σημαντική διαφορά ως στατιστικά σημαντική
  - Αύξηση αξιοπιστίας, εγκυρότητας
  - Μείωση δειγματοληπτικού, συστηματικού και τυχαίου σφάλματος
- **Αντιπροσωπευτικό δείγμα**

---

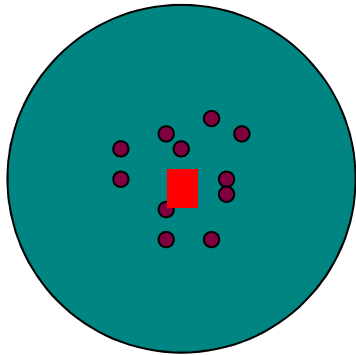
# Μεροληψία και Μεταβλητότητα

**Μεροληψία (*bias*):** οποιαδήποτε επιρροή η οποία καθιστά τα παρατηρούμενα αποτελέσματα μη αντιπροσωπευτικά της θεραπείας.

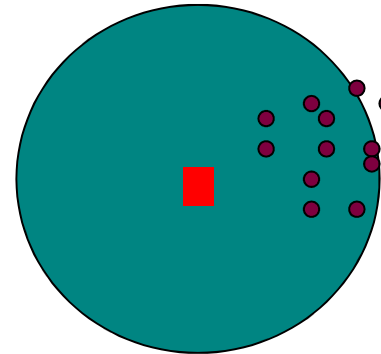
**Μεταβλητότητα (*variability*):** η «ανακρίβεια» των αποτελεσμάτων λόγω της δειγματοληψίας.

# Bias - Variability

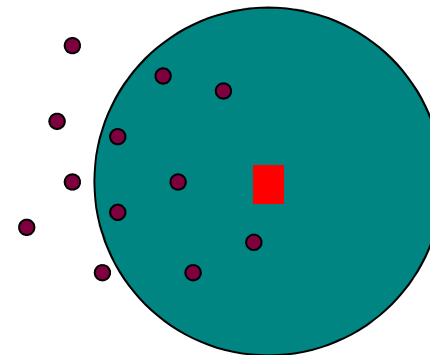
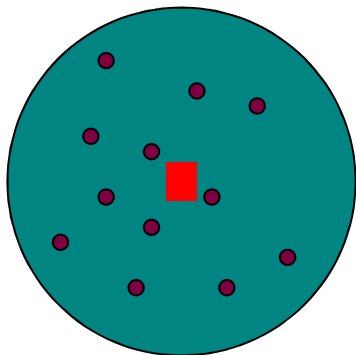
No bias,  
small  
variability



Bias,  
small  
variability



No bias,  
larger  
variability



Bias,  
larger  
variability

---

# Μεροληψία

Η μεροληψία είναι συστηματικό λάθος ή διαφορά

- Οι πιο υγιείς ασθενείς λαμβάνουν την θεραπεία Α και οι λιγότερο υγιείς την Β.
- Η θεραπεία Α είναι πιο γευστική και άρα καταπίνεται μεγαλύτερη ποσότητα από ότι στην θεραπεία Β.
- Η θεραπεία Α είναι «νέα και συναρπαστική» με αποτέλεσμα τόσο οι γιατροί όσο και οι ασθενείς να αναμένουν καλύτερα αποτελέσματα.

---

## Είδη Μεροληψίας

- Μεροληψία επιλογής (προηγούμενη θεραπεία).
- Μεροληψία μέτρησης (ιατρικοί φάκελοι και διαζώσης επικοινωνία).
- Συγκυρτικοί παράγοντες (οι αθλούμενοι έχουν μικρότερο βάρος και δεν καπνίζουν).

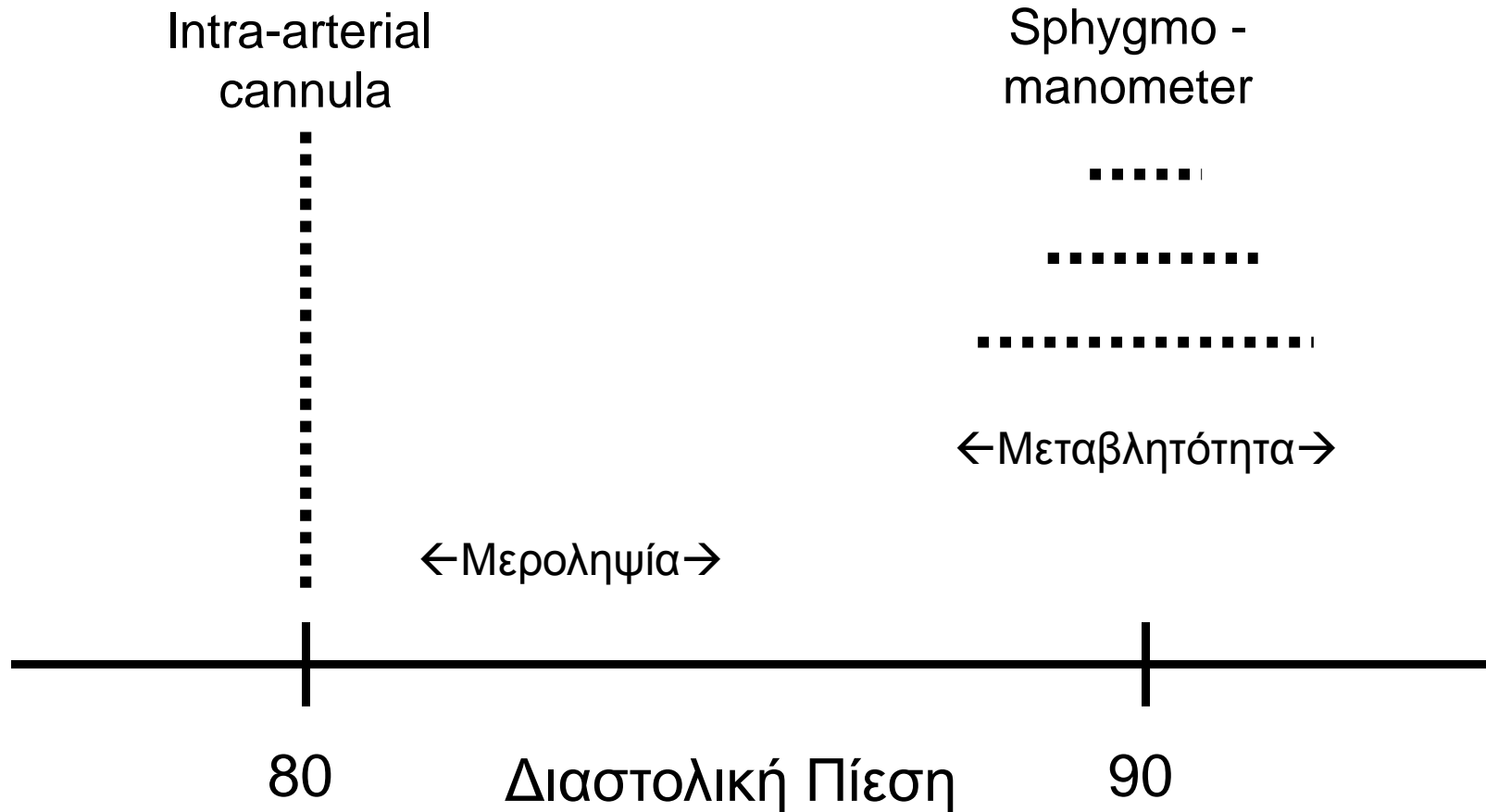


---

# Είδη Μεταβλητότητας

- Μέτρησης
  - Οργάνου
  - Παρατηρητή
- Βιολογική
  - Μεταξύ διαφορετικών ασθενών
  - Εντός του ίδιου ασθενούς

# Τελείως διαφορετικά προβλήματα



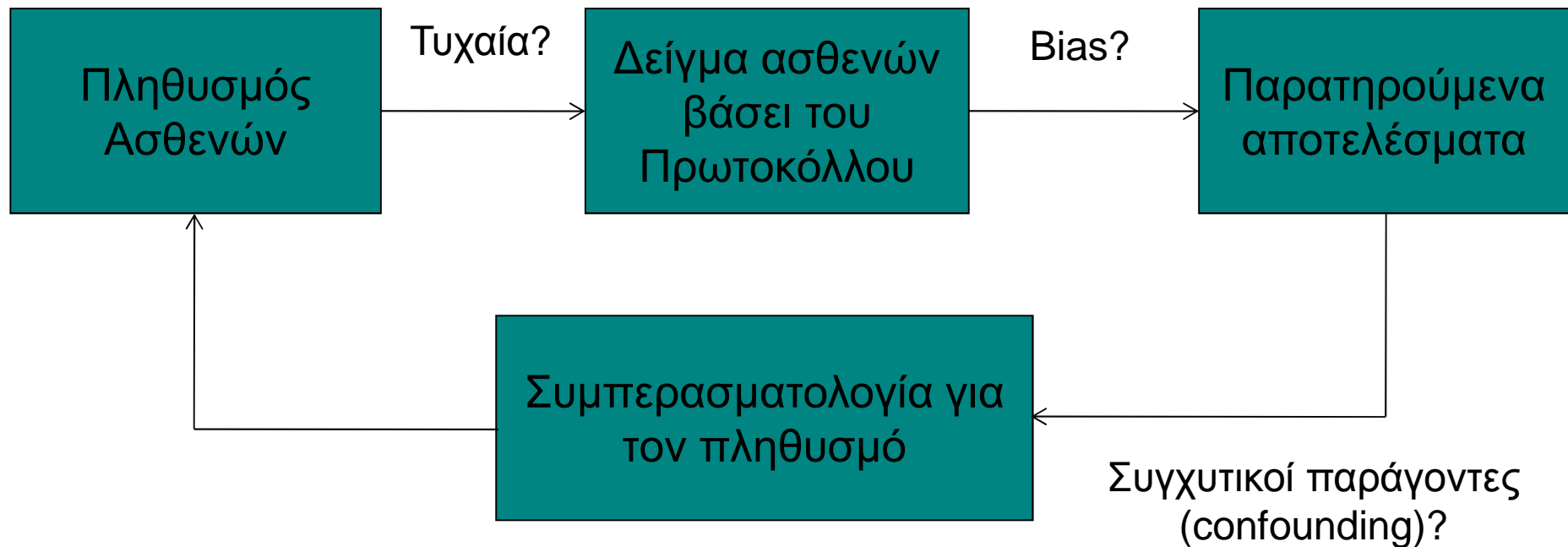
---

# Μεροληψία και μεταβλητότητα

Πώς αντιμετωπίζονται;

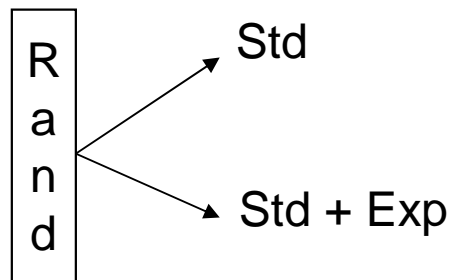
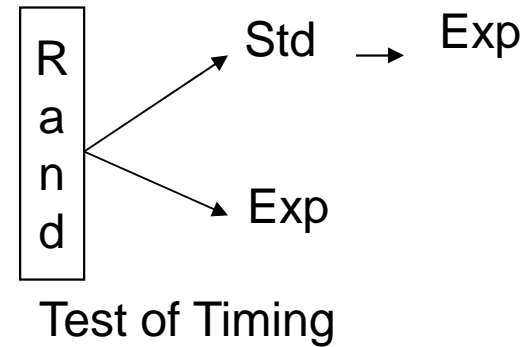
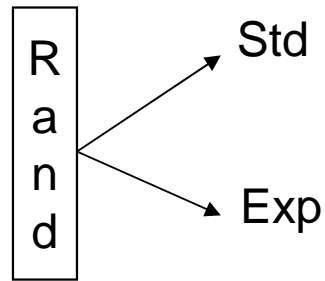
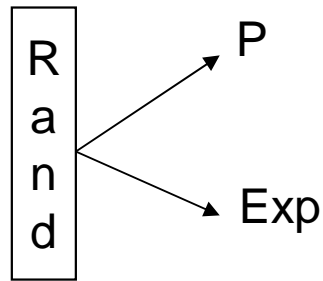
- Και οι δύο με τον ορθό **σχεδιασμό και τη διεξαγωγή** της μελέτης .
- Η χρήση της **κατάλληλης** μεθοδολογίας στατιστικής ανάλυσης μας επιτρέπει να «μετρήσουμε» σωστά τα αποτελέσματα δηλ. να ξεχωρίσουμε το σπάνιο από το σύνηθες ή αναμενόμενο.

# Γενική αλληλουχία στατιστικής συμπερασματολογίας

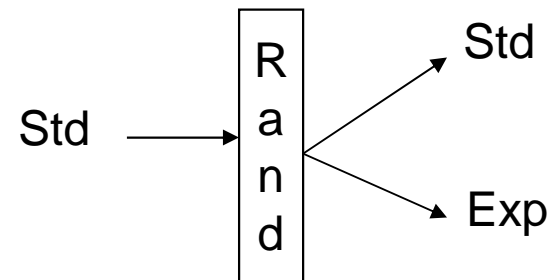


# Basic Phase III Designs

(P=placebo, E=experimental, Std=standard)



Test of Combination Tx



Test of Switching

---

# Factorial Design

- Randomize to each level of several factors
- Example: 2x2 factorial (2 factors, each at 2 levels)
  - Placebo
  - Treatment A
  - Treatment B
  - Treatment A + Treatment B
- Physician's Health Study: 2 by 2
  - 'A'=aspirin, 'B'=betacarotene
- ISIS-4: 2 by 2 by 2
  - Nitroglycerin, Magnesium, and Captopril
  - 8 study 'arms'
- Attempts to assess the value of each factor

---

## Neonatal Network Study of Early Dexamethasone Tx in ELBW Infants

- Population: infants with extremely low birth weight (ELBW)
- In a two-by-two factorial design, tested both dexamethasone treatment and a strategy of minimal ventilatory support (permissive hypercapnia).
- Infants were randomly assigned to one of four groups according to the study medication (dexamethasone or placebo) and ventilatory treatment
  - NEJM, January 11, 2001

---

# ISIS-4

- ISIS-4: a large, simple trial. It enrolled 58,050 acute MI patients at centers around the world and randomly assigned treatment in a 2 X 2 X 2 factorial design (Lancet, 1995, 669-85):
  - 1 Month Oral Mononitrate versus Placebo
  - 1 Month Oral Captopril versus Placebo
  - 24 Hour Intravenous Magnesium Sulfate versus Open Control



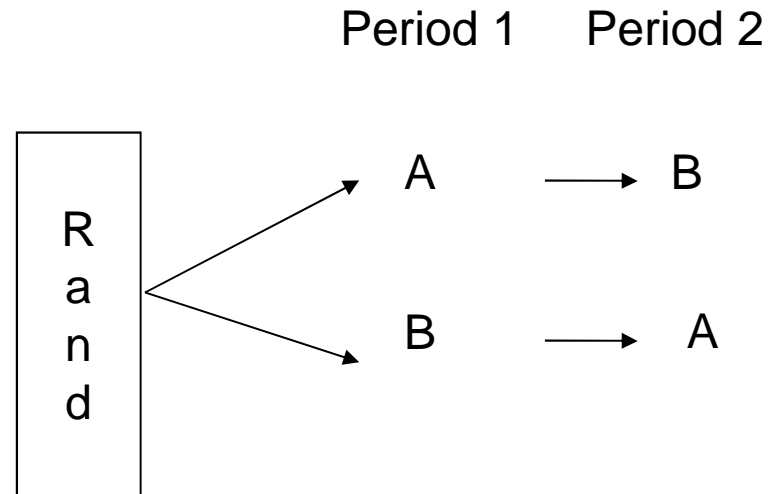
---

# Limitations of Factorial Design

- Factorial designs can be very efficient. They allow simultaneous tests of two hypotheses in the same population (efficacy of aspirin, efficacy of betacarotene).
- Key limitation is **assumption of no interaction**. If the difference in efficacy of regimen A relative to control differs according to the presence or absence of B, the factorial design can produce misleading results

- 
- These are all variants of the **parallel groups** design. In the parallel groups design, each patient is assigned to one treatment regimen and the experience of groups assigned to different groups is compared
  - An important, though less frequently used design is the *crossover* design

# Crossover Design



This is the two-group, two-period crossover design. Critical issues in all cross-over designs: carry-over effects & patient 'drop outs'

---

## Example of How Designs Vary as Knowledge is Gained: AIDS Therapy

- Early studies compared single drug, e.g., AZT, to placebo
- Once AZT was demonstrated to be effective, studies compared other therapies to AZT
- Investigations of timing of AZT (based on CD4 count) and value of switching
- More recent studies have established the efficacy of combination therapies (HAART)
- Factorial designs have been used to test two questions at the same time
  - A is an antiretroviral agent
  - B is a prophylactic agent for opportunistic infection

# Έλεγχοι Υποθέσεων στις Κλινικές Δοκιμές Φάσης III

- **Μελέτη Υπεροχής (Superiority Trials):** Είναι η νέα θεραπεία καλύτερη από την καθιερωμένη;
  - Ιδέα: Ανάπτυξη πιο αποτελεσματικών θεραπευτικών αγωγών
- **Μελέτη Ισοδυναμίας (Equivalence Trials):** Είναι η νέα θεραπεία όσο καλή είναι η καθιερωμένη;
  - Ιδέα: Ίσως η νέα θεραπεία είναι λιγότερο ακριβή ή έχει λιγότερη τοξικότητα. Τότε, αν είναι εξίσου αποτελεσματική με την καθιερωμένη θεραπεία, θα είναι προτιμητέα

# Κλινική δοκιμή υπεροχής vs κλινική δοκιμή ισοδυναμίας - Ερμηνεία

Έστω 3 θεραπευτικές επιλογές: εικονικό φάρμακο (P), καθιερωμένη θεραπεία (S) και νέα θεραπεία (N). Για να αξιολογήσει κάποιος τη νέα θεραπεία έχει τις εξής επιλογές:

**A:** Να σχεδιάσει μια κλινική δοκιμή υπεροχής η οποία θα:

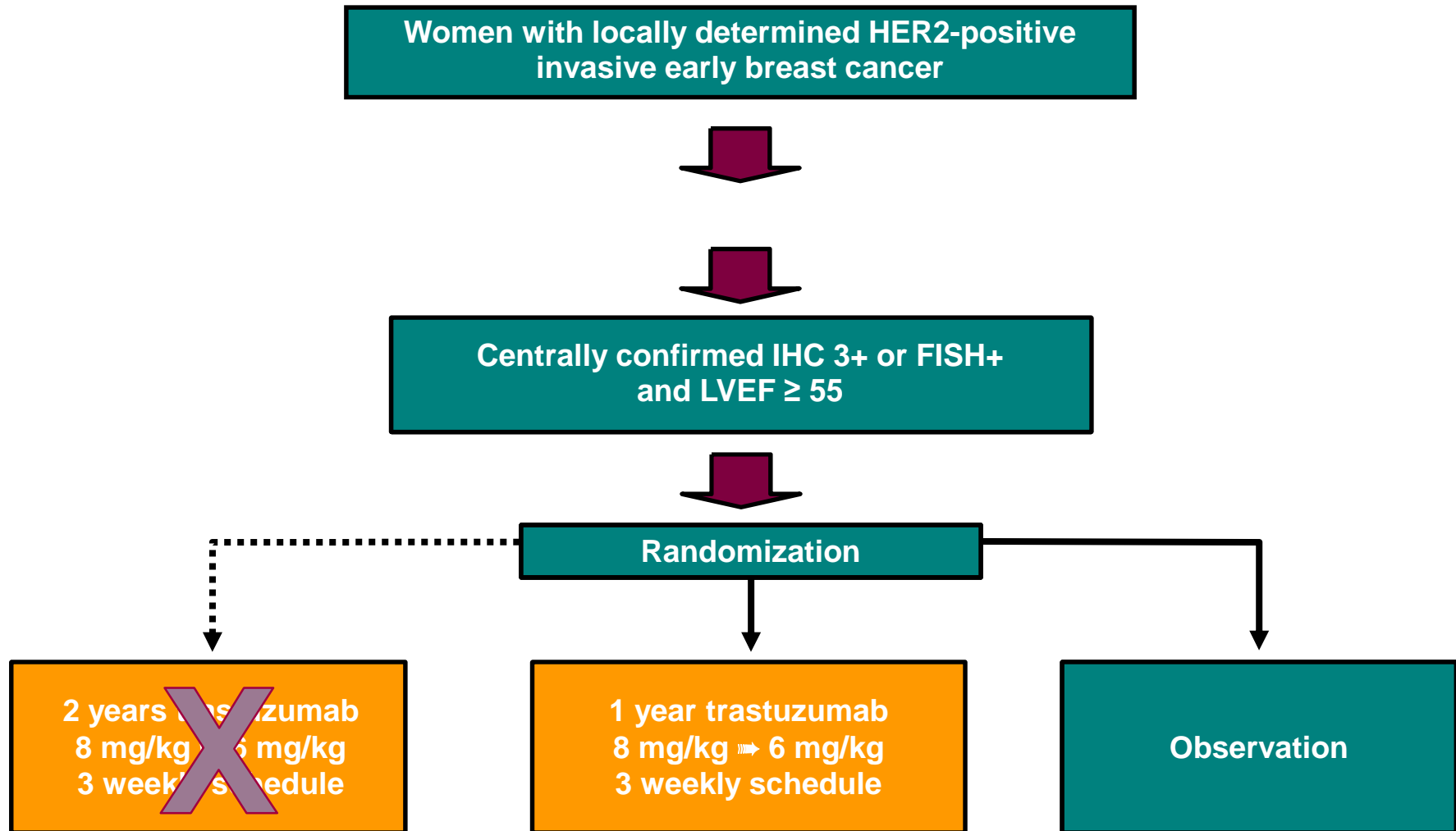
- ✓ Συγκρίνει τη νέα θεραπεία (N) με το εικονικό φάρμακο (P)
- ✓ Συγκρίνει τη νέα θεραπεία (N) σε σχέση με την καθιερωμένη

**B:** Να σχεδιάσει μια κλινική δοκιμή ισοδυναμίας η οποία θα συγκρίνει τη νέα θεραπεία (N) με την καθιερωμένη (S).

## Τι μπορούμε να μάθουμε-1;

- Στη **κλινική δοκιμή υπεροχής** ελέγχουμε τη μηδενική υπόθεση ότι η νέα θεραπεία (T) είναι το ίδιο αποτελεσματική με την καθιερωμένη (S). Στη περίπτωση που απορρίψουμε την  $H_0$  μπορούμε να συμπεράνουμε ότι η νέα θεραπεία (N) είναι αποτελεσματικότερη της καθιερωμένης (S).
- Η κλινική δοκιμή υπεροχής δεν μπορεί να απαντήσει στο ερώτημα αν η νέα θεραπεία (N) είναι ισοδύναμη με την καθιερωμένη (S).

# HERA trial design





---

## Trastuzumab after Adjuvant Chemotherapy in HER2-Positive Breast Cancer

The NEW ENGLAND  
JOURNAL of MEDICINE

ESTABLISHED IN 1812      OCTOBER 20, 2005      VOL. 353 NO. 16

Martine J. Piccart-Gebhart, M.D., Ph.D., Marion Procter, M.Sc., Brian Leyland-Jones, M.D., Ph.D., Aron Goldhirsch, M.D., Michael Untch, M.D., Ian Smith, M.D., Luca Gianni, M.D., Jose Baselga, M.D., Richard Bell, M.D., Christian Jackisch, M.D., David Cameron, M.D., Mitch Dowsett, Ph.D., Carlos H. Barrios, M.D., Günther Steger, M.D., Chiun-Shen Huang, M.D., Ph.D., M.P.H., Michael Andersson, M.D., Dr.Med.Sci., Moshe Inbar, M.D., Mikhail Lichinitser, M.D., István Láng, M.D., Ulrike Nitz, M.D., Hiroji Iwata, M.D., Christoph Thomssen, M.D., Caroline Lohrich, M.D., Thomas M. Suter, M.D., Josef Rüschoff, M.D., Tamás Sütő, M.D., Ph.D., Victoria Greatorex, M.Sc., Carol Ward, M.Sc., Carolyn Strahle, Ph.D., Eleanor McFadden, M.A., M. Stella Dolci, and Richard D. Gelber, Ph.D., for the Herceptin Adjuvant (HERA) Trial Study Team

### Statistical analysis

Enrollment of 4482 patients was planned to detect a 23 percent relative reduction in the risk of a disease-free-survival event with 80 percent power, with the use of a two-sided significance level of 2.5 percent for each comparison: two years of trastuzumab versus observation and one year of trastuzumab versus observation. A total of 951 disease-free-survival events were required for the final analysis. One interim efficacy analysis was planned after 475 events, with a specified significance level of  $P \leq 0.001$  required, with the use of a sequential plan according to the O'Brien-Fleming boundary as implemented by Lan and DeMets. The independent data monitoring committee reviewed data on patient enrollment, deaths, compliance, and safety every six months and conducted the interim cardiac safety and efficacy reviews as preplanned. The efficacy analyses were conducted according to the intention-to-treat principle. Chi-square tests for categorical data and log-rank tests for time-to-event end points provided two-sided P values. Kaplan-Meier curves are presented. Cox proportional hazards regression analysis was used to estimate hazard ratios and 95 percent confidence intervals.

## Trastuzumab after Adjuvant Chemotherapy in HER2-Positive Breast Cancer

The NEW ENGLAND  
JOURNAL of MEDICINE

ESTABLISHED IN 1812 OCTOBER 20, 2005 VOL. 353 NO. 16

Martine J. Piccart-Gebhart, M.D., Ph.D., Marion Procter, M.Sc., Brian Leyland-Jones, M.D., Ph.D., Aron Goldhirsch, M.D., Michael Untch, M.D., Ian Smith, M.D., Luca Gianni, M.D., Jose Baselga, M.D., Richard Bell, M.D., Christian Jackisch, M.D., David Cameron, M.D., Mitch Dowsett, Ph.D., Carlos H. Barrios, M.D., Günther Steger, M.D., Chiun-Shen Huang, M.D., Ph.D., M.P.H., Michael Andersson, M.D., Dr.Med.Sci., Moshe Inbar, M.D., Mikhail Lichinitser, M.D., István Láng, M.D., Ulrike Nitz, M.D., Hiroji Iwata, M.D., Christoph Thomssen, M.D., Caroline Lohrisch, M.D., Thomas M. Suter, M.D., Josef Rüschoff, M.D., Tamás Sütő, M.D., Ph.D., Victoria Greatorex, M.Sc., Carol Ward, M.Sc., Carolyn Strahle, Ph.D., Eleanor McFadden, M.A., M. Stella Dolci, and Richard D. Gelber, Ph.D., for the Herceptin Adjuvant (HERA) Trial Study Team

### Statistical analysis

Enrollment of **4482 patients was planned to detect a 23 percent relative reduction in the risk of a disease-free-survival event with 80 percent power**, with the use of a **two-sided significance level of 2.5 percent for each comparison**: two years of trastuzumab versus observation and one year of trastuzumab versus observation. A total of 951 disease-free-survival events were required for the final analysis. One interim efficacy analysis was planned after 475 events, **with a specified significance level of  $P \leq 0.001$**  required, with the use of a sequential plan according to the O'Brien-Fleming boundary as implemented by Lan and DeMets. The independent data monitoring committee reviewed data on patient enrollment, deaths, compliance, and safety every six months and conducted the interim cardiac safety and efficacy reviews as preplanned. The efficacy analyses were conducted according to the intention-to-treat principle. **Chi-square tests for categorical data and log-rank tests for time-to event end points provided two-sided P values. Kaplan-Meier curves are presented. Cox proportional hazards regression analysis was used to estimate hazard ratios and 95 percent confidence intervals.**

---

## Τι μπορούμε να μάθουμε-2;

- Η μηδενική υπόθεση σε μια **κλινική δοκιμή ισοδυναμίας** είναι ότι η νέα θεραπεία (N) δεν είναι ισοδύναμη με την καθιερωμένη (S). Αν απορρίψουμε τη μηδενική υπόθεση μπορούμε να καταλήξουμε ότι η N είναι ισοδύναμη της S κατά ένα συγκεκριμένο μέγεθος.
- Στις μελέτες ισοδυναμίας δεν υπάρχει η δυνατότητα σύγκρισης της νέας θεραπείας (N) με εικονικό φάρμακο (P).

# Κλινική Δοκιμή Ισοδυναμίας - Παράδειγμα

VOLUME 26 · NUMBER 26 · SEPTEMBER 10 2008

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

## Phase III Study, V-15-32, of Gefitinib Versus Docetaxel in Previously Treated Japanese Patients With Non-Small-Cell Lung Cancer

*Riichiroh Maruyama, Yutaka Nishiwaki, Tomohide Tamura, Nobuyuki Yamamoto, Masahiro Tsuboi, Kazuhiko Nakagawa, Tetsu Shinkai, Shunichi Negoro, Fumio Imamura, Kenji Eguchi, Koji Takeda, Akira Inoue, Keisuke Tomii, Masao Harada, Noriyuki Masuda, Haiyi Jiang, Yohji Itoh, Yukito Ichinose, Nagahiro Saijo, and Masahiro Fukuoka*

# Κλινική Δοκιμή Ισοδυναμίας - Παράδειγμα

VOLUME 26 · NUMBER 26 · SEPTEMBER 10 2008

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Phase III Study, V-15-32, of Gefitinib Versus Docetaxel in Previously Treated Japanese Patients With Non-Small-Cell Lung Cancer

## Methods

The primary objective was to compare overall survival to demonstrate noninferiority for gefitinib relative to docetaxel. An unadjusted Cox regression model was used for the primary analysis.

- Noninferiority was to be concluded if the upper CI limit was  $\leq 1.25$ . Superiority was concluded if the upper CI limit was less than 1.
- A total of 296 death events were required for 90% power to demonstrate noninferiority, with the assumption that gefitinib had better overall survival than docetaxel (median survival, 14 v 12 months), and the study plan was to recruit 484 patients.

---

# Ερώτηση Ι

**Μείωση του επιπέδου σημαντικότητας  $\alpha$ , διατηρώντας όλα τα άλλα αλληλεπιδρώντα μεγέθη σταθερά, οδηγεί σε:**

- **Α.** Αύξηση Σφάλματος Τύπου Ι
- **Β.** Μείωση Μεγέθους Δείγματος
- **Γ.** Αύξηση της Ισχύος
- **Δ.** Αύξηση Σφάλματος Τύπου ΙΙ

---

## Ερώτηση II

Για να αυξήσουμε την ισχύ μιας κλινικής δοκιμής υπεροχής, βάσει μιας συγκεκριμένης τελικής έκβασης, μπορούμε:

- Α. Να χρησιμοποιήσουμε μικρότερη διασπορά
- Β. Να αυξήσουμε το μέγεθος δείγματος
- Γ. Να αρκεστούμε στο να διακρίνουμε μεγαλύτερη διαφορά από την «ελάχιστη κλινικά σημαντική»
- Δ. Τα Α και Β
- Ε. Τα Α και Γ

## Ερώτηση III

Να σημειώσετε ποιά από τις κατωτέρω δηλώσεις είναι σωστή (μόνο μία):

- Α. Όταν σε ένα στατιστικό έλεγχο υπόθεσης, η  $p$ -value  $< \alpha$ , απορρίπτω την εναλλακτική υπόθεση
- Β. Όταν σε ένα στατιστικό έλεγχο υπόθεσης, η  $p$ -value  $> \alpha$ , δέχομαι την εναλλακτική υπόθεση
- Γ. Όταν σε ένα στατιστικό έλεγχο υπόθεσης, η  $p$ -value  $< \alpha$ , απορρίπτω τη μηδενική υπόθεση
- Δ. Όταν σε ένα στατιστικό έλεγχο υπόθεσης, η  $p$ -value  $> \alpha$ , απορρίπτω την εναλλακτική υπόθεση



---

## Ερώτηση IV

**Σε μία μελέτη υπεροχής, που συγκρίνει ως προς την αποτελεσματικότητά τους δύο θεραπευτικές αγωγές, με τελικό αποτέλεσμα  $p\text{-value} > \alpha$ , το τελικό συμπέρασμα είναι:**

- Α. Η μελέτη δε διαθέτε ικανοποιητική ισχύ να δείξει διαφορά
- Β. Υπάρχει διαφορά μεταξύ των θεραπειών
- Γ. Δέχομαι την ισοδυναμία των δύο θεραπειών
- Δ. Κανένα από τα ανωτέρω

## Ερώτηση V

Σε μία σχετικά μικρού μεγέθους μελέτη υπεροχής, που συγκρίνει ως προς την αποτελεσματικότητά τους δύο Θεραπευτικές αγωγές, με τελικό αποτέλεσμα  $p\text{-value} < \alpha$  ισχύει η δήλωση (μόνο μία):

- Α. Απορρίπτεται η μηδενική υπόθεση με πιθανότητα σφάλματος  $\alpha$
- Β. Η ισχύς είναι πολύ σημαντική και πρέπει να ελεγχθεί
- Γ. Αποδείχτηκε χωρίς αμφιβολία διαφορά μεταξύ των θεραπειών
- Δ. Το μέγεθος δείγματος δεν επαρκεί για εξαγωγή συμπεράσματος

---

# Θέματα

- Έλεγχοι υποθέσεων
- Σφάλματα τύπου I και II
- Στατιστική ισχύς
- Μέγεθος Κλινικής Δοκιμής

# Έλεγχος Υποθέσεων

## ■ Μηδενική Υπόθεση ( $H_0$ ):

Συνήθως η  $H_0$  είναι μία δήλωση «μη ύπαρξης διαφοράς» ή «μη επίδρασης» και γενικά θέλουμε να μην ισχύει

- Ο μέσος του πληθυσμού ισούται με κάποια υποθετική τιμή  $\mu_0$ ,  
 $H_0 : \mu = \mu_0$ .
- Η διάμεση επιβίωση ατόμων που πήραν την θεραπεία Β είναι 3 χρόνια,  $H_0 : \delta=3$
- Το ποσοστό ασθενών που πήραν τη νέα θεραπεία και είχαν κλινικά μετρήσιμη βελτίωση ήταν 20%,  $H_0 : \pi=0.20$

# Έλεγχος Υποθέσεων

## ■ Εναλλακτική Υπόθεση ( $H_A$ ):

Η εναλλακτική υπόθεση είναι η δήλωση που ελπίζουμε να αληθεύει αντί της μηδενικής

- $H_A : \mu \neq \mu_0$ .
- Διάμεση επιβίωση < 3 χρόνια,  $H_A : \delta < 3$ .
- Ποσοστό ασθενών είναι 30%  $H_A : \pi > 0.20, \pi = 0.30$

---

## Trastuzumab after Adjuvant Chemotherapy in HER2-Positive Breast Cancer

Martine J. Piccart-Gebhart, M.D., Ph.D., Marion Procter, M.Sci., Brian Leyland-Jones, M.D., Ph.D., Aron Goldhirsch, M.D., Michael Untch, M.D., Ian Smith, M.D., Luca Gianni, M.D., Jose Baselga, M.D., Richard Bell, M.D., Christian Jackisch, M.D., David Cameron, M.D., Mitch Dowsett, Ph.D., Carlos H. Barrios, M.D., Günther Steger, M.D., Chiun-Shen Huang, M.D., Ph.D., M.P.H., Michael Andersson, M.D., Dr.Med.Sci., Moshe Inbar, M.D., Mikhail Lichinitser, M.D., István Láng, M.D., Ulrike Nitz, M.D., Hiroji Iwata, M.D., Christoph Thomssen, M.D., Caroline Lohrisch, M.D., Thomas M. Suter, M.D., Josef Rüschoff, M.D., Tamás Sütő, M.D., Ph.D., Victoria Greatorex, M.Sc., Carol Ward, M.Sc., Carolyn Strahle, Ph.D., Eleanor McFadden, M.A., M. Stella Dolci, and Richard D. Gelber, Ph.D., for the Herceptin Adjuvant (HERA) Trial Study Team

### ***Statistical analysis***

Enrollment of 4482 patients was planned **to detect a 23 percent relative reduction in the risk of a disease-free–survival event** with 80 percent power, with the use of a two-sided significance level of 2.5 percent for each comparison: two years of trastuzumab versus observation and **one year of trastuzumab versus observation.**

***Κλινικά Σημαντική Διαφορά (Clinical significance): 23%***

---

## Παρουσίαση των Υποθέσεων

- $\lambda_T$  = Κίνδυνος εμφάνισης νόσου (Risk of a disease event) για την ομάδα του trastuzumab
- $\lambda_S$  = Κίνδυνος εμφάνισης νόσου για την ομάδα παρατήρησης

Λόγος Κινδύνων, Hazard Ratio  $\theta = \text{HR} = \lambda_T / \lambda_S$

$\text{HR} > 1$  σημαίνει  $\lambda_T > \lambda_S$ , **trastuzumab χειρότερο**

$\text{HR} = 1$  σημαίνει  $\lambda_T = \lambda_S$ , **καμία διαφορά**

$\text{HR} < 1$  σημαίνει  $\lambda_T < \lambda_S$ , **trastuzumab καλύτερο**

# Παρουσίαση των Υποθέσεων

Λόγος Κινδύνων, Hazard Ratio  $\theta = \text{HR} = \lambda_T / \lambda_S$

- Μηδενική Υπόθεση (Null):

$$H_0: \quad \text{HR} = 1 \quad \text{OR}$$

$$\lambda_T = \lambda_S \quad \text{OR}$$

$$\text{Median DFS}_T = \text{Median DFS}_S$$

- Εναλλακτική Υπόθεση (Alternative): **trastuzumab καλύτερο**

$$H_A: \quad \text{HR} \neq 1, \text{HR} < 1 \quad \text{OR}$$

$$\lambda_T < \lambda_S \quad \text{OR}$$

$$\text{Median DFS}_T > \text{Median DFS}_S$$



---

## Παρουσίαση των Υποθέσεων

Μηδενική Υπόθεση:  $H_0: HR=1$  ( $\lambda_T = \lambda_S$ )

Εναλλακτική Υπόθεση:  $H_a: HR \neq 1, HR < 1$  ( $\lambda_T < \lambda_S$ )

**Κλινικά Σημαντική Διαφορά : 23%**

Εάν στην πραγματικότητα HR είναι μικρότερος κατά 23%,  
θα θέλαμε να υπάρχει βάσει της μελέτης **ΥΨΗΛΗ**  
**ΠΙΘΑΝΟΤΗΤΑ** να βρεθεί αυτή η διαφορά.

Αυτή η πιθανότητα είναι η

***ΣΤΑΤΙΣΤΙΚΗ ΙΣΧΥΣ, STATISTICAL POWER = 1-β***

# Παρουσίαση των Υποθέσεων

Θα πρέπει να μπορούμε να εντοπίσουμε με μεγάλη πιθανότητα αυτή τη μείωση κατά 23% εάν αυτή πραγματικά υπάρχει, δηλ. με **ΜΕΓΑΛΗ ΣΤΑΤΙΣΤΙΚΗ ΙΣΧΥ**

## Trastuzumab after Adjuvant Chemotherapy in HER2-Positive Breast Cancer

Martine J. Piccart-Gebhart, M.D., Ph.D., Marion Procter, M.Sc., Brian Leyland-Jones, M.D., Ph.D., Aron Goldhirsch, M.D., Michael Untch, M.D., Ian Smith, M.D., Luca Gianni, M.D., Jose Baselga, M.D., Richard Bell, M.D., Christian Jackisch, M.D., David Cameron, M.D., Mitch Dowsett, Ph.D., Carlos H. Barrios, M.D., Günther Steger, M.D., Chiun-Shen Huang, M.D., Ph.D., M.P.H., Michael Andersson, M.D., Dr.Med.Sci., Moshe Inbar, M.D., Mikhail Lichinitser, M.D., István Láng, M.D., Ulrike Nitz, M.D., Hiroji Iwata, M.D., Christoph Thomssen, M.D., Caroline Lohrisch, M.D., Thomas M. Suter, M.D., Josef Rüschoff, M.D., Tamás Sütő, M.D., Ph.D., Victoria Greatorex, M.Sc., Carol Ward, M.Sc., Carolyn Straehle, Ph.D., Eleanor McFadden, M.A., M. Stella Dolci, and Richard D. Gelber, Ph.D., for the Herceptin Adjuvant (HERA) Trial Study Team

## Statistical analysis

Enrollment of 4482 patients was planned to detect a 23 percent relative reduction in the risk of a disease-free–survival event with 80 percent power, with the use of a two-sided significance level of 2.5 percent for each comparison: two years of trastuzumab versus observation and one year of trastuzumab versus observation.

## Παρουσίαση των Υποθέσεων

- Δουλεύουμε με το λογάριθμο του HR ( $\Delta$ ) για να μπορούμε να κάνουμε χρήση της κανονικής κατανομής
- Εκτιμώμενος λόγος κινδύνων στη λογαριθμική κλίμακα [ $\ln(HR)$ ] για εμφάνιση νόσου μεταξύ της ομάδας trastuzumab και της ομάδας παρατήρησης:

$$\Delta = \ln(HR) = \ln(\lambda_T) - \ln(\lambda_S)$$

όπου,  $\Delta$  είναι η αληθινή σχετική ελάττωση του κινδύνου εμφάνισης νόσου για θεραπεία ενός έτους με trastuzumab έναντι απλής παρατήρησης

# Παρουσίαση των Υποθέσεων

- Έχουμε

- Μηδενική Υπόθεση (Null Hypothesis):

$$H_0: \Delta = 0 \text{ (για HR=1)}$$

- Εναλλακτική Υπόθεση (Alternative Hypothesis):

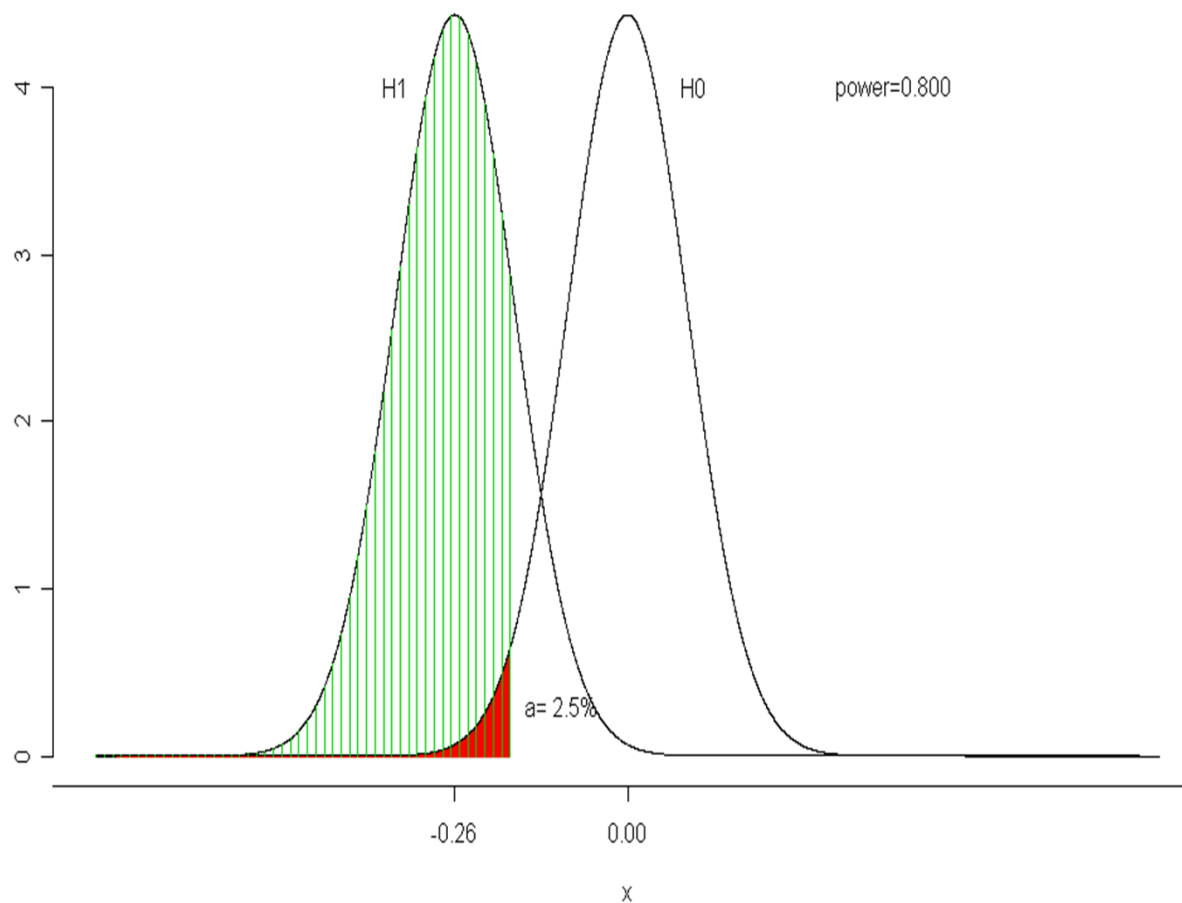
$$H_a: \Delta < 0 \text{ ( for HR<1)}$$

Το σημείο στην εναλλακτική υπόθεση που ενδιαφέρει να διακρίνουμε με ισχύ 80% είναι

$$\Delta = \ln(1 - 0.23) = \ln(0.77) = -0.26$$

Δηλ. η ελάχιστη κλινικά σημαντική διαφορά

# HERA Trial (HR=0.77 vs. HR=1)



$$\Delta = \ln(\text{HR}) = \ln(0.77) = -\ln(1-0.23) = -0.26$$

# Representation of the Hypotheses

## Trastuzumab after Adjuvant Chemotherapy in HER2-Positive Breast Cancer

Martine J. Piccart-Gebhart, M.D., Ph.D., Marion Procter, M.Sci., Brian Leyland-Jones, M.D., Ph.D., Aron Goldhirsch, M.D., Michael Untch, M.D., Ian Smith, M.D., Luca Gianni, M.D., Jose Baselga, M.D., Richard Bell, M.D., Christian Jackisch, M.D., David Cameron, M.D., Mitch Dowsett, Ph.D., Carlos H. Barrios, M.D., Günther Steger, M.D., Chiun-Shen Huang, M.D., Ph.D., M.P.H., Michael Andersson, M.D., Dr.Med.Sci., Moshe Inbar, M.D., Mikhail Lichinitser, M.D., István Láng, M.D., Ulrike Nitz, M.D., Hiroji Iwata, M.D., Christoph Thomssen, M.D., Caroline Lohrisch, M.D., Thomas M. Suter, M.D., Josef Rüschoff, M.D., Tamás Sütő, M.D., Ph.D., Victoria Greatorex, M.Sc., Carol Ward, M.Sc., Carolyn Strahle, Ph.D., Eleanor McFadden, M.A., M. Stella Dolci, and Richard D. Gelber, Ph.D., for the Herceptin Adjuvant (HERA) Trial Study Team

### Statistical analysis

Enrollment of **4482** patients was planned to detect a 23 percent relative reduction in the risk of a disease-free–survival event with 80 percent power, with the use of a two-sided significance level of 2.5 percent for each comparison: two years of trastuzumab versus observation and one year of trastuzumab versus observation.

# Έλεγχος Υποθέσεων

- **Μηδενική Υπόθεση ή  $H_0$ :**

Συνήθως η  $H_0$  είναι μία δήλωση «μη ύπαρξης διαφοράς» ή «μη επίδρασης» και γενικά θέλουμε να μην ισχύει

**Trastuzumab median DFS = observation median DFS, HR=1**

- **Εναλλακτική Υπόθεση ή  $H_A$ :**

Η εναλλακτική υπόθεση είναι η δήλωση που ελπίζουμε να αληθεύει αντί της μηδενικής

**Trastuzumab median DFS > observation median DFS, HR=0.77**

Το τελικό συμπέρασμα είναι ΠΑΝΤΑ ένα εκ των ακόλουθων δύο:

- Απορρίπτουμε τη Μηδενική Υπόθεση
- Αποτύχαμε να απορρίψουμε τη Μηδενική Υπόθεση

# Τύποι Σφαλμάτων

- Στην εφαρμογή της τεχνικής του ελέγχου υποθέσεων, υπάρχουν δύο τύποι σφαλμάτων που μπορεί να διαπραχθούν.
- Το πρώτο καλείται *σφάλμα τύπου I*, και είναι επίσης γνωστό ως *σφάλμα απόρριψης*, ή *α σφάλμα*.
- Η πιθανότητα να διαπράξει κανείς ένα σφάλμα τύπου I καθορίζεται από το επίπεδο σημαντικότητας του ελέγχου.  
$$\alpha = P(\text{απόρριψης της } H_0 \mid H_0 \text{ αληθής}).$$
- Ο δεύτερος τύπος σφάλματος που μπορεί να διαπραχθεί κατά τη διάρκεια ενός ελέγχου υπόθεσης είναι ένα *σφάλμα τύπου II* ή *σφάλμα μη απόρριψης* ή *β σφάλμα*.
- Η πιθανότητα να διαπράξει κανείς ένα σφάλμα τύπου II είναι:
- $\beta = P(\text{μη απόρριψης της } H_0 \mid H_0 \text{ λανθασμένη})$  και υπολογίζεται σε μία συγκεκριμένη τιμή της εναλλακτικής υπόθεσης.



## Ο πίνακας της αλήθειας

		Πραγματικότητα	
		$H_0$ αληθής	$H_0$ όχι αληθής
Απόφαση βάσει μελέτης	<p><i>Δεν απορρίπτουμε <math>H_0</math></i></p> <p><math>\Delta = 0</math>  <math>\ln(\text{HR}) = 0, \text{HR} = 1</math>  <math>\ln(\lambda_T) - \ln(\lambda_S) = 0</math></p>	<p>✓</p> <p>(1-<math>\alpha</math>)</p>	<p>Σφάλμα τύπου II</p> <p>(<math>\beta = 20\%</math>)</p>
	<p><i>Απορρίπτουμε <math>H_0</math></i></p> <p><math>\Delta &lt; 0, \text{HR} &lt; 1</math></p>	<p>Σφάλμα τύπου I</p> <p>(<math>\alpha = 2.5\%</math>)</p>	<p>✓</p> <p>(Ισχύς = 1-<math>\beta = 0.80</math>)</p>

---

# Τρόπος επιλογής Σφαλμάτων

- Η επιλογή βασίζεται στο κόστος των σφαλμάτων
- Επιλέγουμε μικρότερο σφάλμα τύπου I όταν το κόστος να απορρίψουμε τη μηδενική υπόθεση είναι μεγάλο
- Επιλέγουμε μεγαλύτερο σφάλμα τύπου I όταν θέλουμε να απορρίψουμε πιο εύκολα την  $H_0$

---

# Ισχύς

- Η **ισχύς** είναι η πιθανότητα να απορρίψουμε την  $H_0$  όταν η μηδενική υπόθεση είναι λανθασμένη, δηλαδή είναι η πιθανότητα να αποφύγουμε ένα σφάλμα τύπου II.

$$\text{Ισχύς} = P(\text{απόρριψης της } H_0 \mid H_0 \text{ λανθασμένη}) = 1 - \beta$$

- Όπως και η  $\beta$ , η ισχύς πρέπει να υπολογιστεί για μία συγκεκριμένη τιμή για τη παράμετρο του πληθυσμού στην εναλλακτική υπόθεση.

---

## Μελέτες ισοδυναμίας

- Ελέγχουμε αν η πειραματική θεραπεία είναι «τόσο αποτελεσματική» όσο και η συνήθης.
- Αποκαλούνται και μελέτες «απόδειξης της μηδενικής υπόθεσης» ή
- Non-inferiority μελέτες.

Αν η «ισοδύναμη» θεραπεία έχει λιγότερο κόστος ή λιγότερες παρενέργειες το κέρδος είναι προφανές.

## Τι ζητάμε

- Έχουμε 2 θεραπείες τη νέα και την καθιερωμένη.
- Ας υποθέσουμε ότι

$$\theta = \pi_{\text{νέας}} - \pi_{\text{καθ}}$$

- Θέλουμε να διακρίνουμε αν  $\theta \leq -\Delta$
- Το  $-\Delta$  με άλλα λόγια προσδιορίζει το κατώτατο όριο ανοχής, να είναι η νέα θεραπεία χειρότερη από την παλιά (**μονόπλευρο διάστημα εμπιστοσύνης**).

### **ΠΡΟΣΟΧΗ:**

- Δε μπορούμε ποτέ με βεβαιότητα να αποδείξουμε ότι  $\Delta=0$ . Αυτό που μπορούμε να αποδείξουμε είναι ότι το  $\Delta$  παίρνει κάποια τιμή πολύ κοντά στο 0.

---

## Μελέτες ισοδυναμίας και διάστημα εμπιστοσύνης

*Ποιο είναι το μέγεθος δείγματος σε κάθε σκέλος ώστε να εξασφαλίσουμε με πιθανότητα τουλάχιστον  $(1-\beta)$  ότι το κατώτερο όριο του  $(1-\alpha)$  διαστήματος εμπιστοσύνης είναι μεγαλύτερο της μείωσης στο αποτέλεσμα της νέας θεραπείας όταν αυτές είναι όντως ισοδύναμες;*

# Κλινική Δοκιμή Ισοδυναμίας - Παράδειγμα

VOLUME 26 · NUMBER 26 · SEPTEMBER 10 2008

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

## Phase III Study, V-15-32, of Gefitinib Versus Docetaxel in Previously Treated Japanese Patients With Non-Small-Cell Lung Cancer

*Riichiroh Maruyama, Yutaka Nishiwaki, Tomohide Tamura, Nobuyuki Yamamoto, Masahiro Tsuboi, Kazuhiko Nakagawa, Tetsu Shinkai, Shunichi Negoro, Fumio Imamura, Kenji Eguchi, Koji Takeda, Akira Inoue, Keisuke Tomii, Masao Harada, Noriyuki Masuda, Haiyi Jiang, Yohji Itoh, Yukito Ichinose, Nagahiro Saijo, and Masahiro Fukuoka*

# Κλινική Δοκιμή Ισοδυναμίας - Παράδειγμα

VOLUME 26 · NUMBER 26 · SEPTEMBER 10 2008

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

## Phase III Study, V-15-32, of Gefitinib Versus Docetaxel in Previously Treated Japanese Patients With Non-Small-Cell Lung Cancer

### Methods

The primary objective was to compare overall survival to demonstrate noninferiority for gefitinib relative to docetaxel. An unadjusted Cox regression model was used for the primary analysis.

- Noninferiority was to be concluded if the upper CI limit was  $\leq 1.25$ . Superiority was concluded if the upper CI limit was less than 1.
- A total of 296 death events were required for 90% power to demonstrate noninferiority, with the assumption that gefitinib had better overall survival than docetaxel (median survival, 14 v 12 months), and the study plan was to recruit 484 patients.



# Equivalence and non-inferiority trial for DIFFERENCE in INCIDENCE

Estimated DIFFERENCE between the new treatment and the standard treatment:  $d$

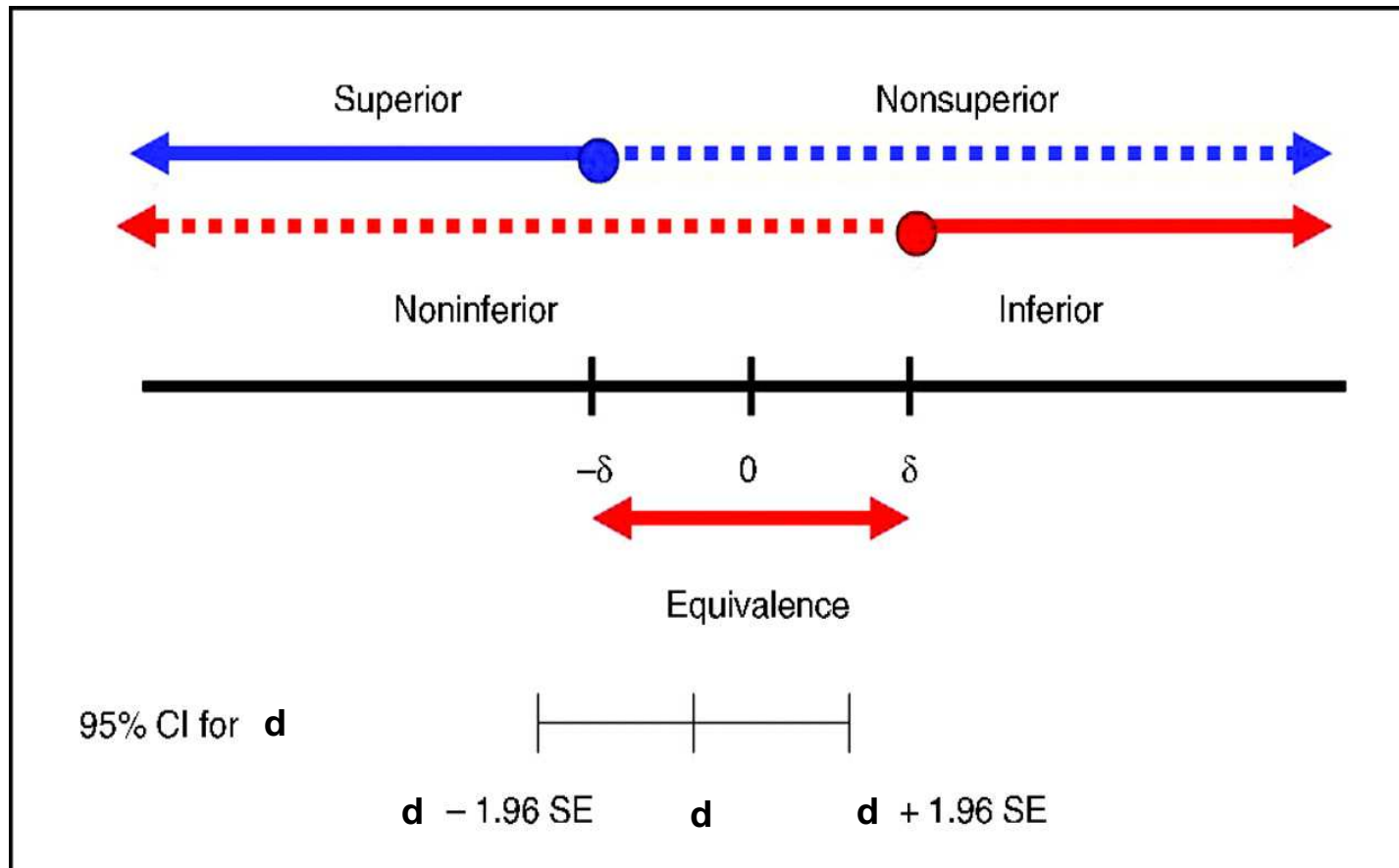
**Equivalence:** the 95% CI for the DIFFERENCE in INCIDENCE lies completely within the equivalence limits (i.e., between  $-\delta$  and  $\delta$ )

**Non-inferiority:** the new treatment is not inferior to the standard treatment if the upper 95% confidence limit is less than the non-inferiority margin  $\delta$

JOURNAL OF CLINICAL ONCOLOGY

Zee, B. C.-Y. J Clin Oncol; 24:1026-1028 2006

# Equivalence and non-inferiority trial for DIFFERENCE in INCIDENCE



JOURNAL OF CLINICAL ONCOLOGY

Zee, B. C.-Y. J Clin Oncol; 24:1026-1028 2006

---

# Κλινική Δοκιμή Ισοδυναμίας - Παράδειγμα

**Phase III Study, V-15-32, of Gefitinib Versus Docetaxel in Previously Treated Japanese Patients With Non-Small- Cell Lung Cancer**

## Results

At data cutoff for overall survival (October 31, 2006)

- Noninferiority in overall survival was not achieved (HR, 1.12; 95.24% CI, 0.89 to 1.40) according to the predefined criterion (upper CI limit for  $HR \leq 1.25$ ).
- However, no statistically significant difference in overall survival was apparent (Wald p-value=.330).

---

# Representation of the Hypotheses

## Trastuzumab after Adjuvant Chemotherapy in HER2-Positive Breast Cancer

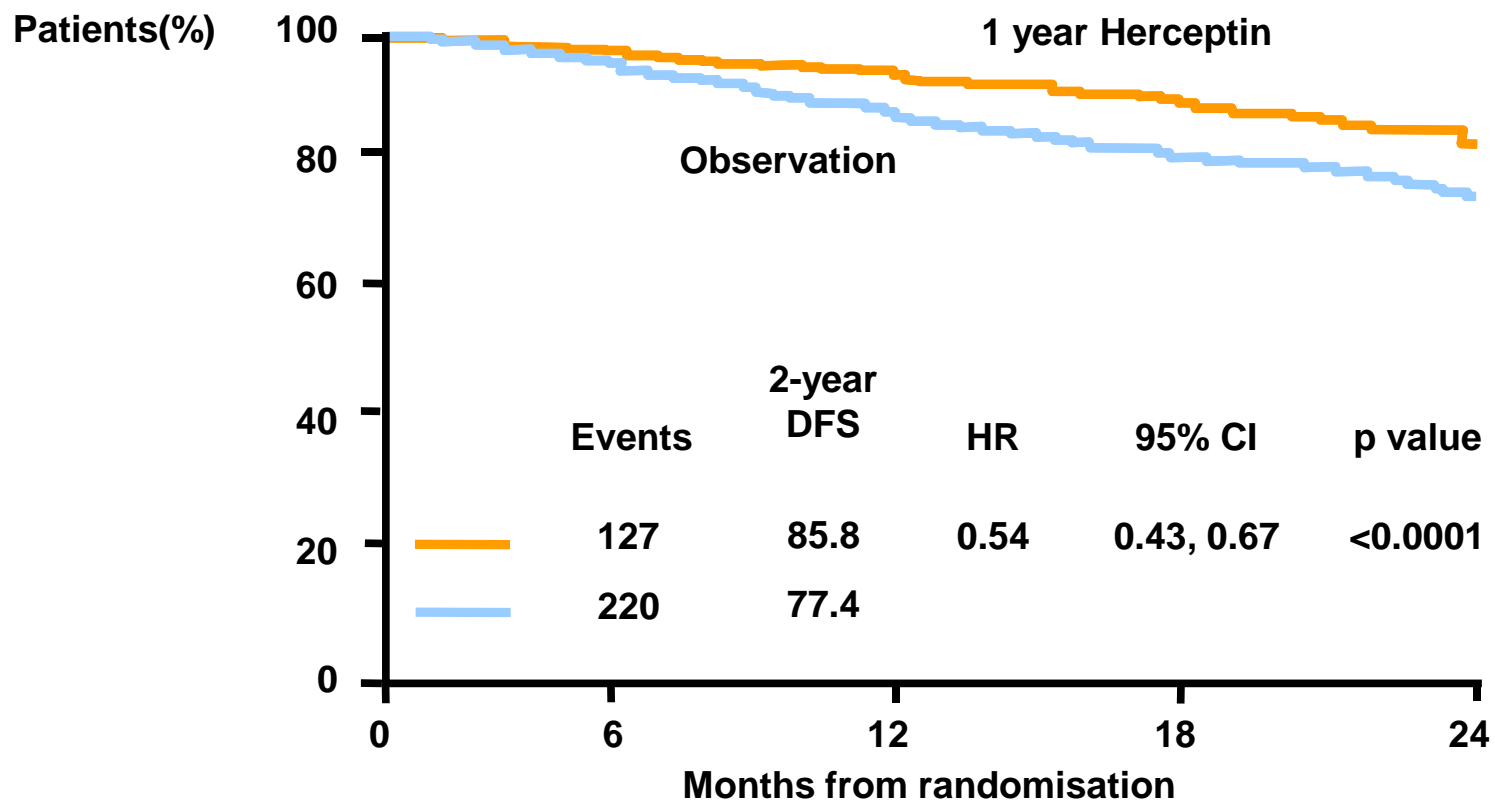
Martine J. Piccart-Gebhart, M.D., Ph.D., Marion Procter, M.Sci., Brian Leyland-Jones, M.D., Ph.D., Aron Goldhirsch, M.D., Michael Untch, M.D., Ian Smith, M.D., Luca Gianni, M.D., Jose Baselga, M.D., Richard Bell, M.D., Christian Jackisch, M.D., David Cameron, M.D., Mitch Dowsett, Ph.D., Carlos H. Barrios, M.D., Günther Steger, M.D., Chiun-Shen Huang, M.D., Ph.D., M.P.H., Michael Andersson, M.D., Dr.Med.Sci., Moshe Inbar, M.D., Mikhail Lichinitser, M.D., István Láng, M.D., Ulrike Nitz, M.D., Hiroji Iwata, M.D., Christoph Thomssen, M.D., Caroline Lohrisch, M.D., Thomas M. Suter, M.D., Josef Rüschoff, M.D., Tamás Sütő, M.D., Ph.D., Victoria Greatorex, M.Sc., Carol Ward, M.Sc., Carolyn Straehle, Ph.D., Eleanor McFadden, M.A., M. Stella Dolci, and Richard D. Gelber, Ph.D., for the Herceptin Adjuvant (HERA) Trial Study Team

### Statistical analysis

Enrollment of 4482 patients was planned to detect a 23 percent relative reduction in the risk of a disease-free–survival event with 80 percent power, with the use of a two-sided significance level of 2.5 percent for each comparison: two years of trastuzumab versus observation and one year of trastuzumab versus observation.

---

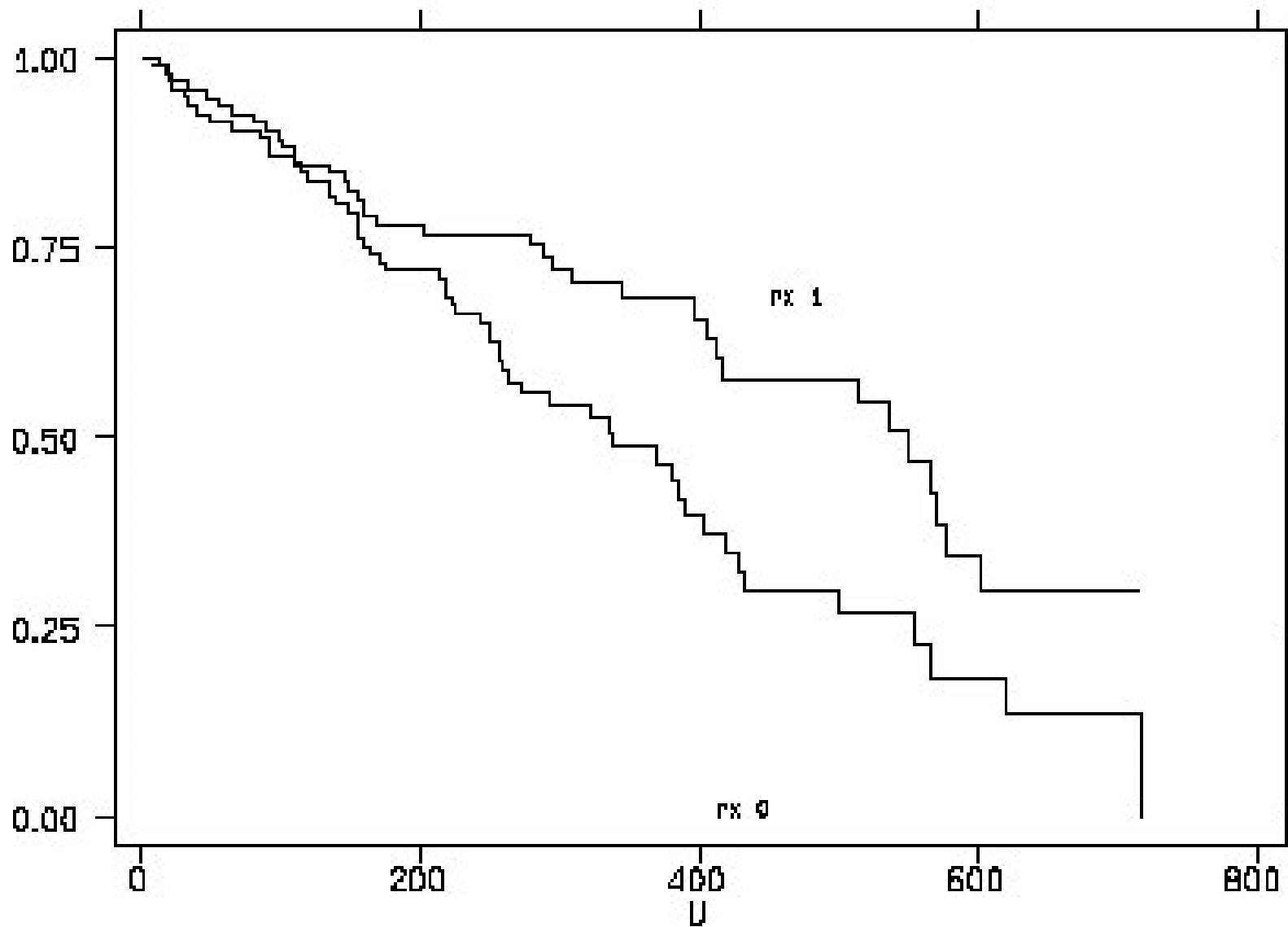
# HERA: DFS



No. at risk	0	6	12	18	24
1 year Herceptin	1694	1172	885	532	268
Observation	1693	1108	767	445	224

Median follow-up: 1 year; DFS, disease-free survival; HR, hazard ratio; CI, confidence interval

Kaplan-Meier survival estimates, by rx



# Trastuzumab after Adjuvant Chemotherapy in HER2-Positive Breast Cancer

The NEW ENGLAND  
JOURNAL of MEDICINE

ESTABLISHED IN 1812

OCTOBER 20, 2005

VOL. 353 NO. 16

Martine J. Piccart-Gebhart, M.D., Ph.D., Marion Procter, M.Sc., Brian Leyland-Jones, M.D., Ph.D., Aron Goldhirsch, M.D., Michael Untch, M.D., Ian Smith, M.D., Luca Gianni, M.D., Jose Baselga, M.D., Richard Bell, M.D., Christian Jackisch, M.D., David Cameron, M.D., Mitch Dowsett, Ph.D., Carlos H. Barrios, M.D., Günther Steger, M.D., Chiun-Shen Huang, M.D., Ph.D., M.P.H., Michael Andersson, M.D., Dr.Med.Sci., Moshe Inbar, M.D., Mikhail Lichinitser, M.D., István Láng, M.D., Ulrike Nitz, M.D., Hiroji Iwata, M.D., Christoph Thomssen, M.D., Caroline Lohrisch, M.D., Thomas M. Suter, M.D., Josef Rüschoff, M.D., Tamás Sütő, M.D., Ph.D., Victoria Greatorex, M.Sc., Carol Ward, M.Sc., Carolyn Strahle, Ph.D., Eleanor McFadden, M.A., M. Stella Dolci, and Richard D. Gelber, Ph.D., for the Herceptin Adjuvant (HERA) Trial Study Team

## Results

**The unadjusted hazard ratio** for an event in the trastuzumab group, as compared with the observation group, **was 0.54 (95 percent confidence interval, 0.43 to 0.67; P<0.0001** by the log-rank test, crossing the interim analysis boundary), representing an absolute benefit in terms of disease-free survival at two years of 8.4 percentage points.

# Έλεγχος Υποθέσεων

- **Παρατηρούμενο επίπεδο σημαντικότητας του ελέγχου**  
(*p-value*) καλείται π.χ., η πιθανότητα να πάρουμε μια τιμή τόσο ακραία ή πιο ακραία από το παρατηρούμενο δειγματικό HR, δεδομένου ότι η μηδενική υπόθεση  $H_0: HR = HR_0$  είναι αληθής.
- Το παρατηρούμενο επίπεδο σημαντικότητας συγκρίνεται με το προκαθορισμένο επίπεδο σημαντικότητας  $\alpha$  για να αποφασίσουμε εάν η μηδενική υπόθεση πρέπει να απορριφθεί.
  - Εάν το  $p\text{-value} \leq \alpha$ , απορρίπτουμε την μηδενική υπόθεση,  $H_0$  και δεχόμαστε την  $H_A$ .
  - Εάν το  $p\text{-value} > \alpha$ , **ΔΕΝ** απορρίπτουμε την μηδενική υπόθεση,  $H_0$ .



# Trastuzumab after Adjuvant Chemotherapy in HER2-Positive Breast Cancer

The NEW ENGLAND  
JOURNAL of MEDICINE

ESTABLISHED IN 1812

OCTOBER 20, 2005

VOL. 353 NO. 16

Martine J. Piccart-Gebhart, M.D., Ph.D., Marion Procter, M.Sc., Brian Leyland-Jones, M.D., Ph.D., Aron Goldhirsch, M.D., Michael Untch, M.D., Ian Smith, M.D., Luca Gianni, M.D., Jose Baselga, M.D., Richard Bell, M.D., Christian Jackisch, M.D., David Cameron, M.D., Mitch Dowsett, Ph.D., Carlos H. Barrios, M.D., Günther Steger, M.D., Chiun-Shen Huang, M.D., Ph.D., M.P.H., Michael Andersson, M.D., Dr.Med.Sci., Moshe Inbar, M.D., Mikhail Lichinitser, M.D., István Láng, M.D., Ulrike Nitz, M.D., Hiroji Iwata, M.D., Christoph Thomssen, M.D., Caroline Lohrisch, M.D., Thomas M. Suter, M.D., Josef Rüschoff, M.D., Tamás Sütő, M.D., Ph.D., Victoria Greatorex, M.Sc., Carol Ward, M.Sc., Carolyn Strahle, Ph.D., Eleanor McFadden, M.A., M. Stella Dolci, and Richard D. Gelber, Ph.D., for the Herceptin Adjuvant (HERA) Trial Study Team

## Discussion

This study shows that trastuzumab can benefit women with HER2-positive breast cancer when given after completion of adjuvant chemotherapy. As compared with observation after primary therapy (including surgery with or without radiotherapy and neoadjuvant or adjuvant chemotherapy), **trastuzumab given after primary therapy reduced the rate of recurrence, particularly distant recurrence, by approximately 50 percent.** This degree of benefit in early breast cancer is the largest to be reported since the introduction of tamoxifen in hormone-receptor–positive disease.

---

# Τελικό Συμπέρασμα σε μία Μελέτη Υπεροχής

- Απορρίπτουμε τη Μηδενική Υπόθεση:  
απόρριψη «ισοδυναμίας»  
αποδοχή στατιστικά σημαντικής διαφοράς
- Αποτύχαμε να απορρίψουμε τη Μηδενική Υπόθεση:  
Ερώτημα: Ποια ήταν η ισχύς

**ΠΡΟΣΟΧΗ:** Δεν πρέπει να γίνεται σύγχυση μεταξύ της **ΑΠΟΔΕΙΞΗΣ ΙΣΟΔΥΝΑΜΙΑΣ** και της **ΑΠΟΤΥΧΙΑΣ ΑΠΟΡΡΙΨΗΣ** της Μηδενικής Υπόθεσης

## 4 αλληλοεξαρτώμενες ποσότητες: $\alpha$ , $\beta$ , $\Delta$ , $N$

- $\alpha$  = ρυθμός ψευδώς θετικών (σφάλμα τύπου I)
- $\beta$  = ρυθμός ψευδώς αρνητικών (σφάλμα τύπου II)
- $\Delta$  = «η ελάχιστη κλινικά σημαντική διαφορά»
- $N$  = το μέγεθος του δείγματος.

Δεδομένης της διαδικασίας ελέγχου, οι ποσότητες αυτές αλληλοεξαρτώνται π.χ.

- $\Delta$  αυξάνεται  $\rightarrow$  πιο εύκολα το ανακαλύπτουμε  $\rightarrow$  το  $1-\beta$  αυξάνεται.

---

## Πού είναι το πρόβλημα;

- Το «α» και το «β» είναι ανταγωνιστικά
  - Με δεδομένο το  $\Delta$
  - Και δεδομένο το  $N$
- Μειώνοντας το «α» αυξάνεται το «β»
- Μειώνοντας το «β» αυξάνεται το «α»

**Συμπέρασμα: πιθανότητα λάθους θα παραμείνει ό,τι και να κάνουμε**

---

## Τι μπορεί να γίνει; (1)

- Να αυξήσουμε το μέγεθος του δείγματος κρατώντας το  $\Delta$  σταθερό.
- Με ποιο τρόπο;
  - Να μειώσουμε το «α» κρατώντας το «β» σταθερό
  - Να μειώσουμε το «β» κρατώντας το «α» σταθερό
  - Να τα μειώσουμε και τα δύο=μεγάλη αύξηση του δείγματος

## Τι μπορεί να γίνει; (2)

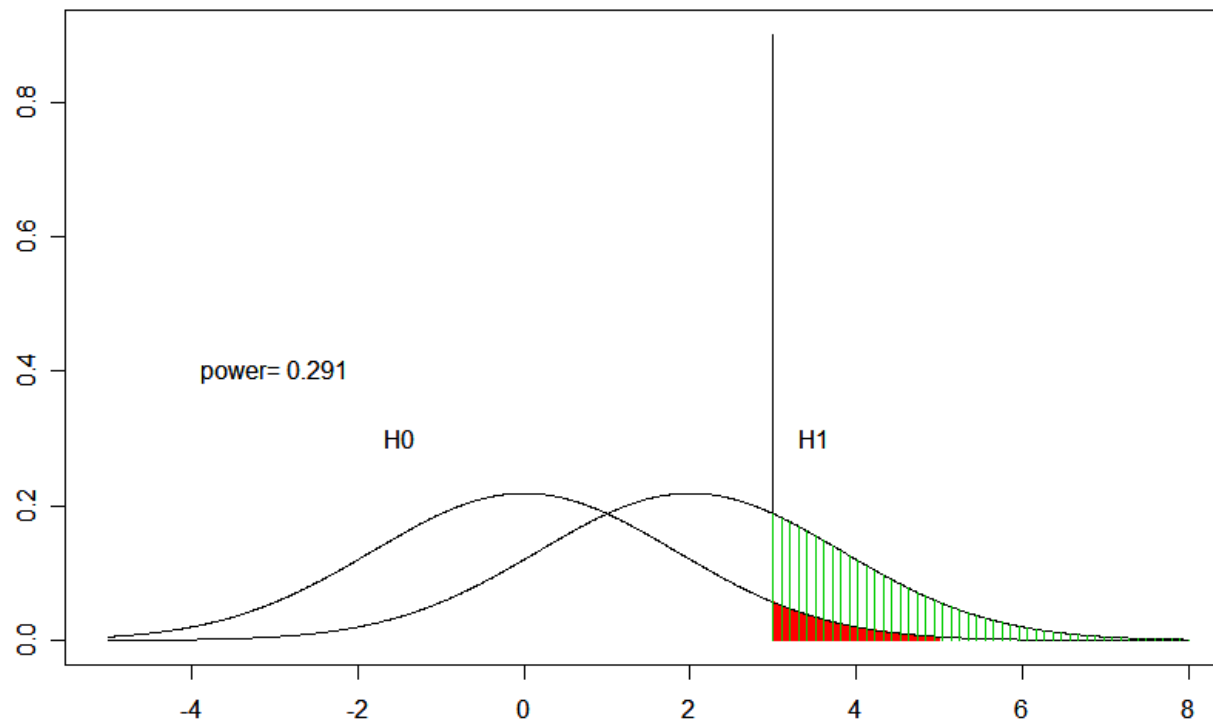
- Να αυξήσουμε το  $\Delta$  κρατώντας το μέγεθος του δείγματος σταθερό.
- Ποιο είναι το αποτέλεσμα;
  - Ας θυμηθούμε ότι το «β» είναι συνάρτηση του  $\Delta$
  - Όταν το  $\Delta$  αυξάνεται το «β» μειώνεται
- «Κερδίζουμε» ότι το «α» παραμένει σταθερό
- Ποιο είναι το πρόβλημα; Η απόφαση αυτή είναι τεχνητή. Είναι λογικό να υποθέσουμε μεγάλο  $\Delta$ ; π.χ. είναι λογικό να πούμε ότι η νέα θεραπεία θα αυξήσει το RR κατά 50%, ενώ όλοι ξέρουν ότι αυτό δεν είναι λογικό!

# Sample size = 30, Power = 29%

Sampling distn under  $H_0: \mu_1 - \mu_2 = 0$

Sampling distn under  $H_1: \mu_1 - \mu_2 = 2$

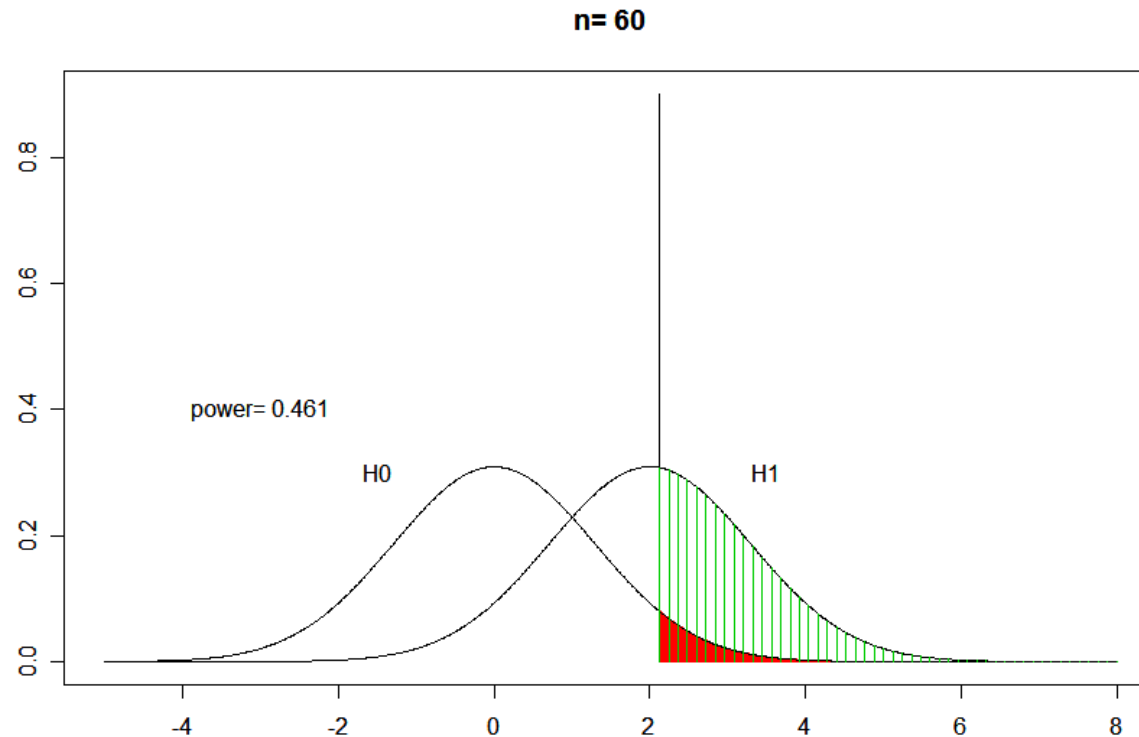
$n = 30$



# Sample size = 60, Power = 46%

Sampling distn under  $H_0: \mu_1 - \mu_2 = 0$

Sampling distn under  $H_1: \mu_1 - \mu_2 = 2$

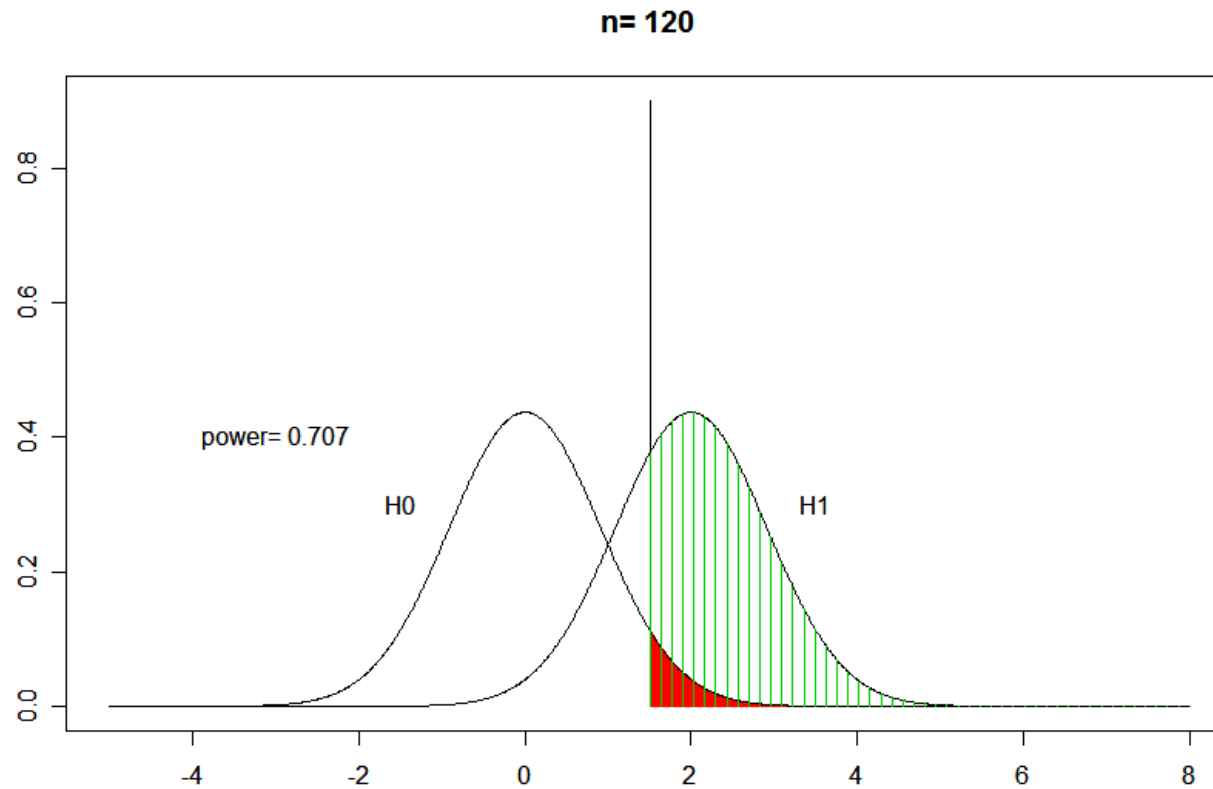




# Sample size = 120, Power = 71%

Sampling distn under  $H_0: \mu_1 - \mu_2 = 0$

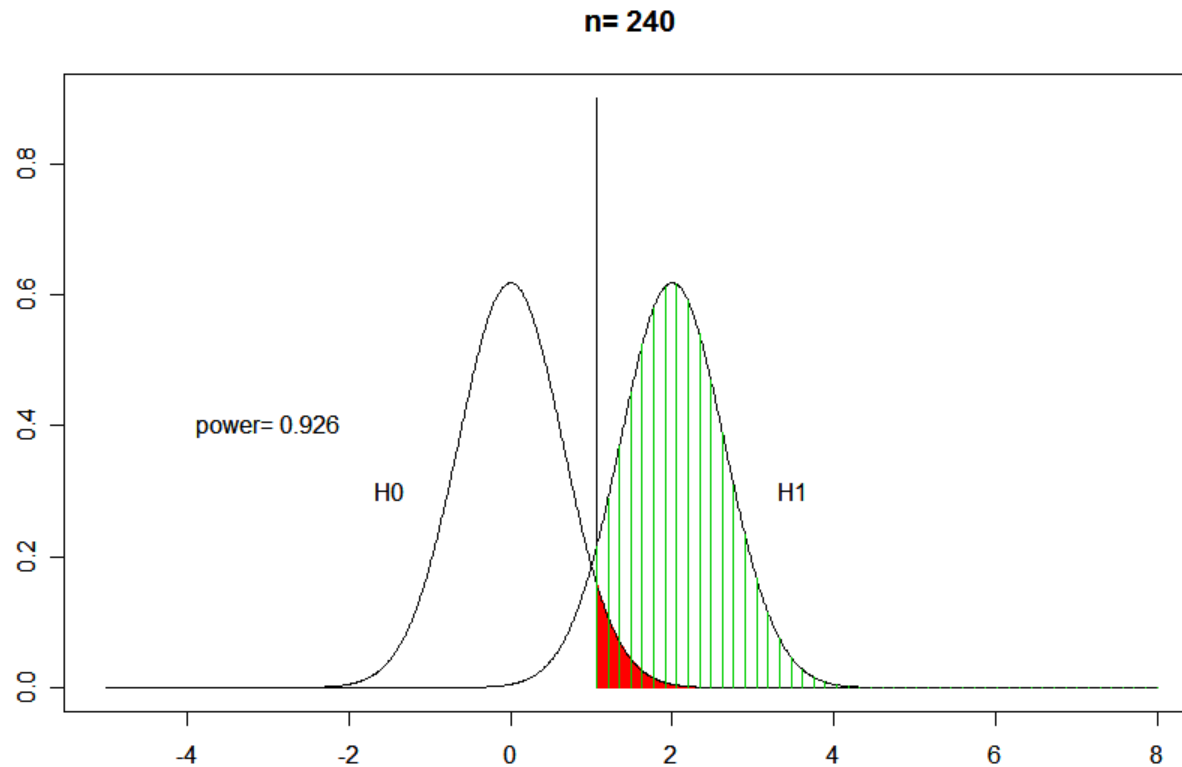
Sampling distn under  $H_1: \mu_1 - \mu_2 = 2$



# Sample size = 240, Power = 93%

Sampling distn under  $H_0: \mu_1 - \mu_2 = 0$

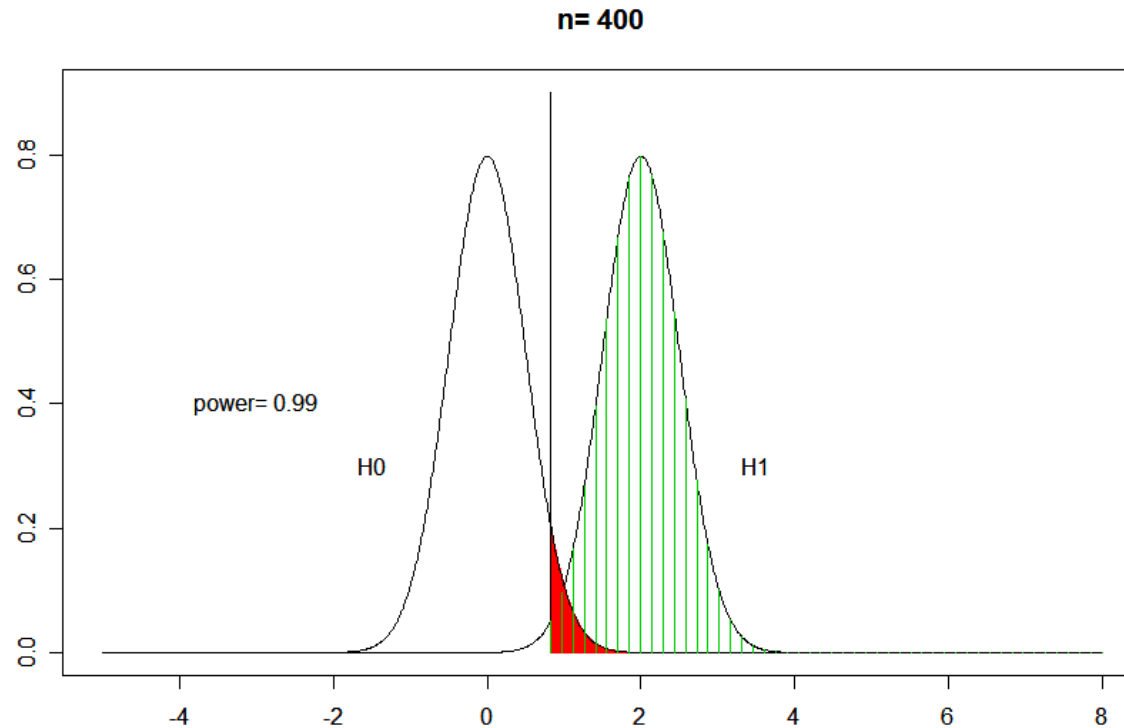
Sampling distn under  $H_1: \mu_1 - \mu_2 = 2$



# Sample size = 400, Power > 99%

Sampling distn under  $H_0: \mu_1 - \mu_2 = 0$

Sampling distn under  $H_1: \mu_1 - \mu_2 = 2$



---

# Επομένως

Άρα υπάρχει ο κίνδυνος

- Να απορρίψουμε τη νέα θεραπεία ενώ αυτή είναι καλή
- Να δεχτούμε τη θεραπεία ως καλή ενώ στην πραγματικότητα δεν είναι

**Συνεπώς πρέπει να κρατήσουμε υπό έλεγχο τις πιθανότητες τέτοιων σφαλμάτων με κατάλληλο σχεδιασμό**

---

# ΕΡΩΤΗΣΕΙΣ ΑΠΑΝΤΗΣΕΙΣ



# Ερμηνεία & Αξιολόγηση Αποτελεσμάτων Κλινικών Δοκιμών

Ουρανία Δαφνή  
Καθηγήτρια, Διευθύντρια  
Εργαστήριο Βιοστατιστικής  
Πανεπιστήμιο Αθηνών

Διευθύντρια  
Frontier Science Foundation-Hellas (FSF-H)

<b>Table 1. Baseline Characteristics of the Patients, Tumors, and Primary Treatments (Intention-to-Treat Groups).*</b>		
<b>Variable</b>	<b>1 Yr of Trastuzumab (N= 1694)</b>	<b>Observation (N= 1693)</b>
<b>Region — no. (%)</b>		
Western and Northern Europe, Canada, South Africa, Australia, New Zealand	1208 (71.3)	1222 (72.2)
Asia Pacific, Japan	203 (12.0)	202 (11.9)
Eastern Europe	189 (11.2)	175 (10.3)
Central and South America	94 (5.5)	94 (5.6)
<b>Race — no. (%) †</b>		
White	1411 (83.3)	1406 (83.0)
Asian	210 (12.4)	207 (12.2)
Black	12 (0.7)	8 (0.5)
Other	57 (3.4)	67 (4.0)
Missing	4 (0.2)	5 (0.3)
<b>Age — no. (%)</b>		
<35 yr	128 (7.6)	123 (7.3)
35–49 yr	751 (44.3)	739 (43.7)
50–59 yr	538 (31.8)	553 (32.7)
≥60 yr	274 (16.2)	275 (16.2)
Missing	3 (0.2)	3 (0.2)
<b>Median age — yr</b>		
	49	49
<b>Menopausal status at randomization (after completion of chemotherapy) — no. (%)</b>		
Premenopausal	272 (16.1)	261 (15.4)
Uncertain	642 (37.9)	629 (37.2)
Postmenopausal	779 (46.0)	797 (47.1)
Missing	1 (0.1)	6 (0.4)

---

# Τυχαιοποίηση-Στρωματοποίηση

- Με την τυχαιοποίηση ελέγχεται:
  - Η μεταβλητότητα του θεραπευτικού αποτελέσματος
  - Το συστηματικό σφάλμα επιλογής (selection bias)
- Η στρωματοποίηση εξισορροπεί γνωστούς παράγοντες κινδύνου μεταξύ των θεραπειών υπό διερεύνηση
  - Η τυχαιοποίηση διεξάγεται εντός κάθε στρώματος



# Ανισορροπίες μπορούν να οδηγήσουν σε Ψευδώς Θετικά Αποτελέσματα

## Χαμηλού Κινδύνου

Θεραπεία	Επιτυχία	Αποτυχία	Σύνολο
A	36 (90%)	4	40
B	9 (90%)	1	10

## Υψηλού Κινδύνου

Θεραπεία	Επιτυχία	Αποτυχία	Σύνολο
A	1 (10%)	9	10
B	4 (10%)	36	40

## Συνολικά Αποτελέσματα

Θεραπεία	Επιτυχία	Αποτυχία	Σύνολο
A	37 (74%)	13	50
B	13 (26%)	37	50

# Ανισορροπίες μπορούν να οδηγήσουν σε Ψευδώς Αρνητικά Αποτελέσματα

## Χαμηλού Κινδύνου

Θεραπεία	Επιτυχία	Αποτυχία	Σύνολο
A	15 (75%)	5	20
B	24 (60%)	16	40

## Υψηλού Κινδύνου

Θεραπεία	Επιτυχία	Αποτυχία	Σύνολο
A	10 (33%)	20	30
B	1 (10%)	9	10

## Συνολικά Αποτελέσματα

Θεραπεία	Επιτυχία	Αποτυχία	Σύνολο
A	25 (50%)	25	50
B	25 (50%)	25	50

# Αξιολόγηση αναφορών κλινικών μελετών

Υπάρχει μεγάλη ανομοιομορφία στην ποιότητα των μελετών

- Επιστημονικά περιοδικά
    - Ποιος είναι ο δείκτης απόρριψης υποβαλλόμενων άρθρων;
    - Ποιοι απαρτίζουν τη Συντακτική Επιτροπή;
    - Είναι η δημοσίευση αξιολογημένη από επιστημονικούς κριτές (refereed?)
    - Βαθμός Απήχησης (Impact Factor)
  - Πρακτικά συνεδρίων;
  - Abstract
  - Άρθρο στα μαζικά μέσα ενημέρωσης;
- 
- Για καμία από τους ανωτέρω τρόπους δημοσίευσης δεν απαιτείται τα δεδομένα να είναι διαθέσιμα για εξωτερική αξιολόγηση
  - Ακόμα και στα καλύτερα περιοδικά τα κριτήρια δημοσίευσης μπορούν να βελτιωθούν

---

# Κατάλογος συγγραφέων

- Ένας μεγάλος κατάλογος συγγραφέων δεν συνεπάγεται πάντα ότι το κείμενο έχει τη σύμφωνη γνώμη όλων
- Λείπουν συγγραφείς; (π.χ. ακτινολόγος ή στατιστικός)
- Διεπιστημονική Προσέγγιση

---

# Πηγή χρηματοδότησης/Μεροληψία/Κίνητρα

- Ποιος χρηματοδότησε τη μελέτη;
- Τα ευρήματα, αντικατοπτρίζουν τη μεροληψία των ερευνητών ή την αντίθετη άποψη (η τελευταία είναι γενικά πιο αξιόπιστη);
- Γιατί δημοσίευσαν τώρα; (γιατί ωρίμασε η μελέτη, είναι μία περιοδική αναφορά ή μία εκκρεμούσα αίτηση για χρηματοδότηση;)
- Γιατί έκαναν την μελέτη (επιβεβαίωση παλαιότερων αποτελεσμάτων με μεγαλύτερη μελέτη ή μία μικρή μελέτη μετά από μία άλλη μικρή μελέτη)

## «Αληθώς Θετικά» vs. «ψευδώς Θετικά» αποτελέσματα

- Διεξαγωγή ανεξάρτητης μελέτης επιβεβαίωσης των αποτελεσμάτων
- Αυτό μπορεί να μην είναι δυνατό λόγω ηθικών, οικονομικών και λόγων κλινικού ενδιαφέροντος
- Ακόμη και αν πραγματοποιηθεί, μία μελέτη επιβεβαίωσης αποτελεσμάτων μπορεί να μην είναι δημοσιεύσιμη ως «επαναληπτική» (γνωστών αποτελεσμάτων)

---

# Ποιότητα Κλινικής Δοκιμής

- Σχεδιασμός Μελέτης
  - Πρωτόκολλο
  - Φόρμες Συλλογής στοιχείων
- Τρόπος Διεξαγωγής μιας Μελέτης
  - Έλεγχος Ποιότητας
  - Παρακολούθηση (Monitoring)
  - Εξωτερική Ιατρική Αξιολόγηση
- Αξιολόγηση και αναφορά αποτελεσμάτων
  - Ανάλυση
  - Εξαιρέσεις
  - Δημοσιεύσεις

---

# ΠΡΟΒΛΗΜΑΤΑ Ι

- Μη ικανοποιητικός αριθμός ασθενών 1960-1975, 71 αρνητικές RCTs

Frieman SA *et al.* *N. Engl. J. Med.* 1978; 299: 690-694

Moher D *et al.* *JAMA* 1994; 272: 122-124

- Μη αναφορά στο μέγεθος του δείγματος 40/45
- Έλλειψη Διαστημάτων Εμπιστοσύνης 39/45
- Αναφορά στατιστικά σημαντικών αποτελεσμάτων σε μεγαλύτερο ποσοστό στις περιλήψεις.

Pocock SJ *et al.* *N. Engl. J. Med.* 1987; 317: 426-432



---

## ΠΡΟΒΛΗΜΑΤΑ II

- Μη αναφορά του τύπου τυχαιοποίησης 48/80 (1987-1988, 80 RCTs)

Altman DG, Dore CJ. *Lancet* 1990; 335: 149-153

- Μη αναφορά του τρόπου τυχαιοποίησης 140/206
- Μη αναφορά του τρόπου απόκρυψης θεραπείας πριν την αρχή της με συνέπεια μεγαλύτερες εκτιμήσεις επίδρασης θεραπείας κατά 30-40% 159/206 (1990-1991, Αγγλία και ΗΠΑ)

Schulz UF, *et al.* *JAMA* 1994; 272: 125-128

---

# Οδηγίες Άρθρων Αναφοράς Κλινικών Δοκιμών

Οκτώβριος 1993

Standards of Reporting Trials  
(SORT) *JAMA* 1995; 273-276

Μάρτιος 1994

Asilomar Working group on  
Recommendations for  
Reporting of Clinical Trials  
in the Biomedical Literature

Σεπτέμβριος 1995

*Annals of Internal Medicine*  
1994; 844-895 CONSORT  
group Εκδότες/κλινικοί  
επιδημιολόγοι/στατιστικοί

---

# Standards for reporting

- The CONSORT statement, Consolidated standards of Reporting Trials: Begg C, Cho M, Eastwood S, Horton R, Moher D, *et al.* **Improving the quality of reporting of randomized controlled trials.** The CONSORT statement. *JAMA* 1996; 276: 637-639

Moher D, Schulz KF, Altman DG *et al.* The CONSORT statement: Revised Recommendations for Improving the Quality of Reports of Parallel-Group Randomized Trials. *Annals of Internal Medicine* 2001; 134: 657-662

- The TREND statement, Transparent Reporting of Evaluations with Non-randomized Designs: Standards for reporting non-randomized evaluations of behavioral and public health interventions:

Des Jarlais DC, Lyles C & Crepaz N. **Improving the reporting quality of non-randomized evaluations of behavioral and public health interventions.** The TREND statement. *American Journal of Public Health* 2004; 94: 360-366

---

# The CONSORT statement

## Κύριες Ενότητες και Υποενότητες

- Τίτλος
- Περίληψη
- Εισαγωγή
- Μέθοδοι
  - Πρωτόκολλο
  - Τυχαιοποίηση
  - Μέθοδοι Απόκρυψης Θεραπευτικής Ομάδας
- Αποτελέσματα
  - Ροή Συμμετεχόντων στη Μελέτη και Παρακολούθησης
  - Ανάλυση
- Σχόλια/Συζήτηση

---

# The CONSORT statement

Μετάδοση πληροφορίας όσον αφορά:

- Σχεδίαση
- Διεξαγωγή
- Ανάλυση
- Δυνατότητα γενίκευσης των αποτελεσμάτων

---

# Περιγραφή Ενοτήτων - CONSORT statement

- Τίτλος
  - Αναγνώριση της μελέτης ως τυχαιοποιημένη δοκιμή
  
- Περίληψη
  - Χρήση μιας δομημένης μορφής
  
- Εισαγωγή
  - Δήλωση όλων των προοπτικά καθορισμένων υποθέσεων, των κλινικών στόχων και των σχεδιασμένων αναλύσεων υποομάδων ή αναλύσεων με επεξηγηματικές μεταβλητές

---

# Περιγραφή Ενότητας Μεθόδων - CONSORT

## Πρωτόκολλο

- ❑ Μελετώμενος πληθυσμός σύμφωνα με τη σχεδίαση συμπεριλαμβανομένων των κριτηρίων εισαγωγής και αποκλεισμού
- ❑ Προγραμματισμένες παρεμβάσεις και χρονοδιάγραμμα
- ❑ Μέτρα κύριας τελικής έκβασης (και δευτερευουσών) και οι ελάχιστες σημαντικές διαφορές ενδιαφέροντος
- ❑ Τρόπος υπολογισμού μεγέθους δείγματος (λεπτομερής περιγραφή)
- ❑ Λογική και περιγραφή των μεθόδων στατιστικής ανάλυσης, παρέχοντας λεπτομέρειες για τις κύριες συγκριτικές μεθόδους και αν ολοκληρώθηκαν σύμφωνα με την αρχή “intention-to-treat”
- ❑ Περιγραφή κανόνων διακοπής της δοκιμής που έχουν οριστεί κατά τη σχεδίαση

---

# Περιγραφή Ενότητας Μεθόδων - CONSORT

## Τυχαιοποίηση

- ❑ Περιγραφή Μονάδας Τυχαιοποίησης (άτομο, συρροή ατόμων, γεωγραφική μονάδα)
- ❑ Περιγραφή Μεθόδου που χρησιμοποιήθηκε για την τυχαιοποίηση
- ❑ Περιγραφή Μεθόδου που χρησιμοποιήθηκε για την απόκρυψη του μηχανισμού τυχαιοποίησης και του χρόνου τυχαιοποίησης
- ❑ Περιγραφή Μεθόδου διαχωρισμού μεταξύ του παραγωγού της τυχαιοποίησης και του ατόμου που την εκτελεί



---

# Περιγραφή Ενότητας Μεθόδων - CONSORT

## Μέθοδοι Απόκρυψης Θεραπευτικής

### Ομάδας(Blinding)

- ❑ Περιγραφή Μηχανισμού (πχ., κάψουλες, δισκία),
- ❑ Περιγραφή ομοιότητας στα χαρακτηριστικά του φαρμακευτικού σκευάσματος (εμφάνιση, γεύση),
- ❑ Περιγραφή ελέγχου της απόκρυψης του μηχανισμού τυχαιοποίησης (φύλαξη κατά τη διάρκεια της μελέτης και πότε έγινε διαθέσιμος)
- ❑ Περιγραφή αποδείξεων για την επιτυχή απόκριση μεταξύ των συμμετεχόντων, του ατόμου που κάνουν παρέμβαση, των αξιολογητών των τελικών εκβάσεων, και των αναλυτών των δεδομένων

---

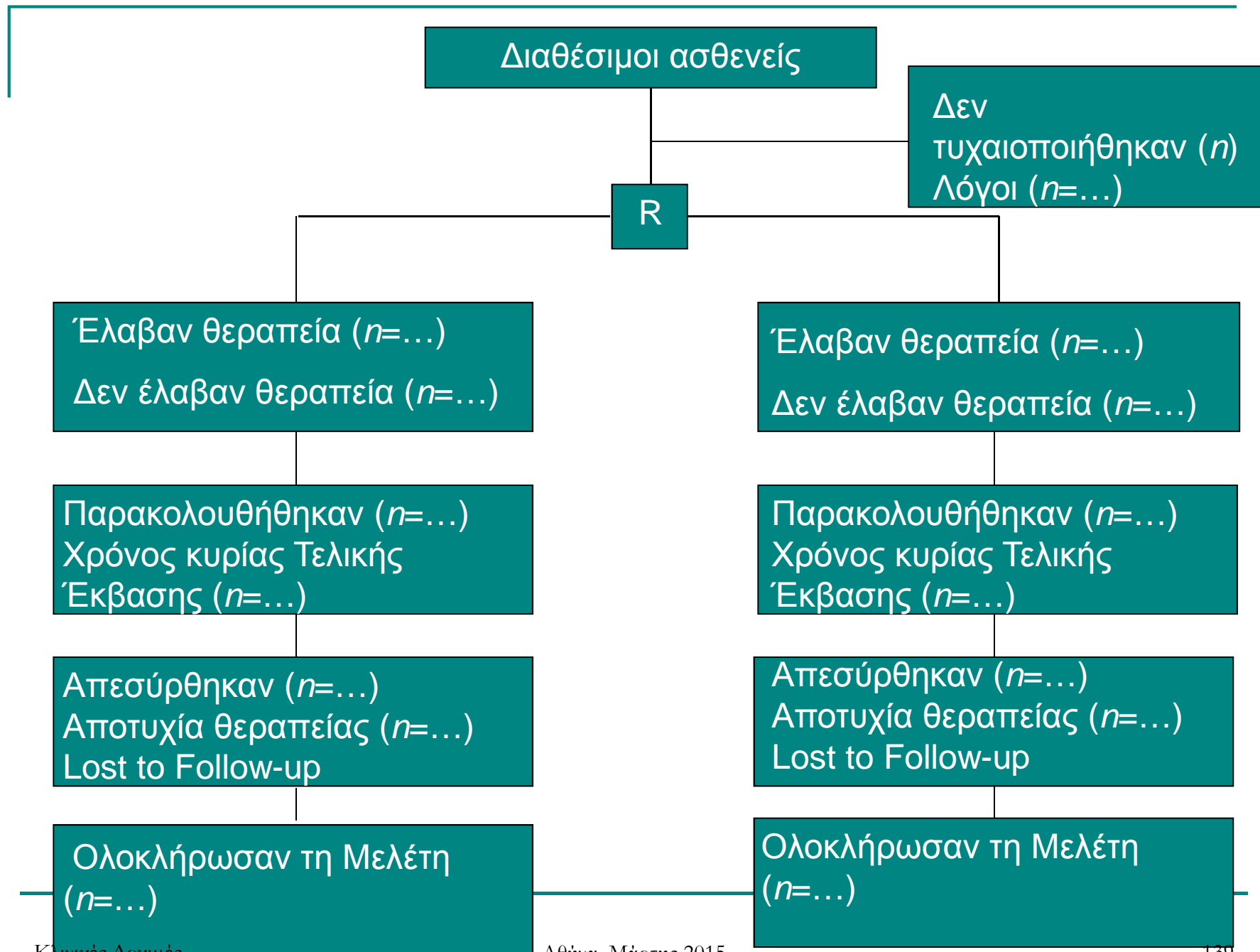
# Περιγραφή Ενότητας Αποτελεσμάτων

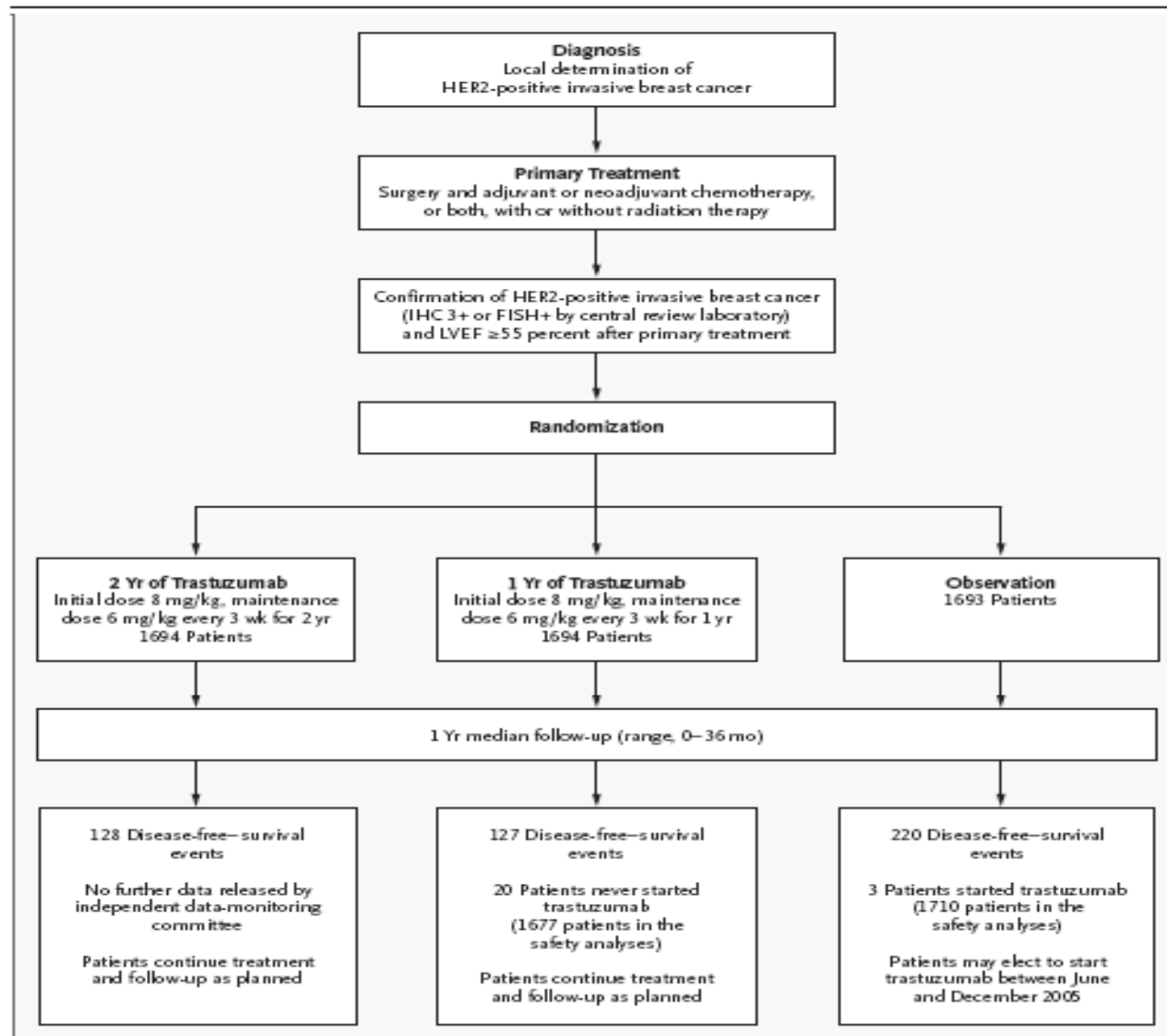
## The CONSORT statement

### Ροή Συμμετεχόντων στη Μελέτη και Παρακολούθηση

Παρουσίαση της ροής των ασθενών μέσα στη μελέτη (Διάγραμμα), των αριθμών και τη χρονική δομή της διαδικασίας της τυχαιοποίησης, των παρεμβάσεων και των μετρήσεων για κάθε τυχαιοποιημένη ομάδα

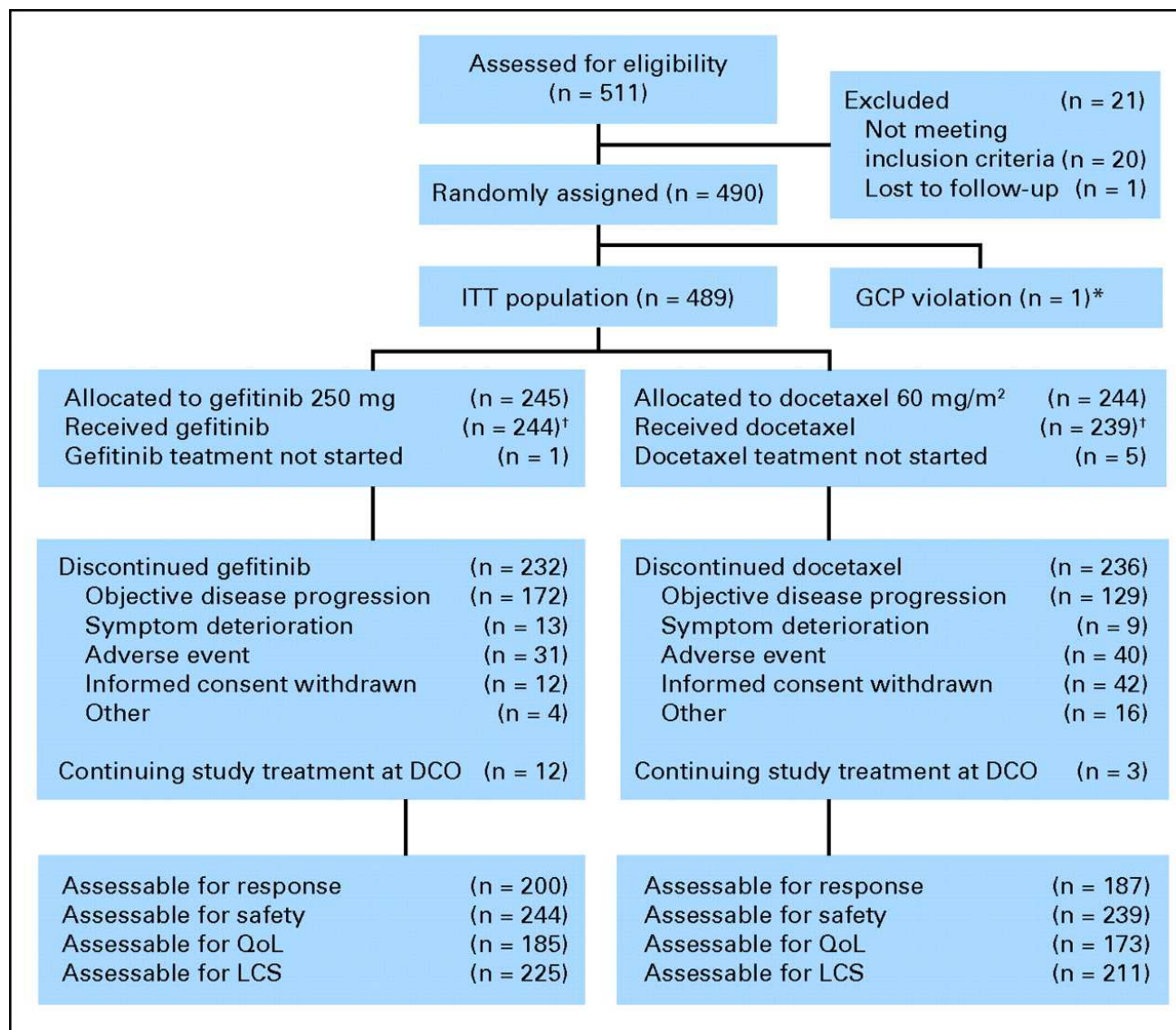
**INTENT TO TREAT (ITT):** Ανάλυση βάσει Τυχαιοποίησης και ΟΧΙ βάσει της θεραπευτικής αγωγής που έλαβαν τελικά με πλήρη συμμόρφωση οι ασθενείς





**Figure 1.** Consolidated Standards of Reporting Trials (CONSORT) Chart of the Herceptin Adjuvant (HERA) Trial.

A total of 5090 women were enrolled between December 2001 and March 2005, for 5081 of whom information was available for analysis. IHC denotes immunohistochemistry, FISH fluorescence in situ hybridization, and LVEF left ventricular ejection fraction.



# Περιγραφή Ενότητας Αποτελεσμάτων-CONSORT

## Ανάλυση

- Δήλωση εκτίμησης της επίδρασης της παρέμβασης στο κύριο μέτρο τελικής έκβασης και στα δευτερεύοντα μέτρα, περιλαμβάνοντας και σημειακές εκτιμήτριες και μέτρα ακρίβειας (Διάστημα εμπιστοσύνης).
- Δήλωση αποτελεσμάτων σε απόλυτους αριθμούς όταν είναι δυνατό (πχ., 10/20, όχι 50%)
- Παρουσίαση περιληπτικών μέτρων και κατάλληλων περιγραφικών και στατιστικών μέτρων συμπερασματολογίας σε ικανοποιητική λεπτομέρεια που να επιτρέπει εναλλακτικές αναλύσεις και επανάληψη της ανάλυσης

---

# Περιγραφή Ενότητας Αποτελεσμάτων-CONSORT

## Ανάλυση (συν.)

- ❑ Περιγραφή προγνωστικών παραγόντων ανά θεραπευτική ομάδα και οποιαδήποτε προσπάθεια να ληφθούν υπόψη στην ανάλυση
- ❑ Περιγραφή τυχόν αποκλίσεων του πρωτοκόλλου από το σχεδιασμό της μελέτης, και τους αντίστοιχους λόγους

---

# Περιγραφή Ενότητας Σχολίων - CONSORT

## Σχόλια

- Δήλωση συγκεκριμένης ερμηνείας των αποτελεσμάτων της μελέτης, συμπεριλαμβάνοντας τις πιθανές πηγές μεροληψίας και ανακρίβειας (εσωτερική εγκυρότητα) και συζήτηση εξωτερικής εγκυρότητας, συμπεριλαμβάνοντας κατάλληλες ποσοτικές μετρήσεις όποτε αυτό είναι δυνατό
  
- Δήλωση γενικότερης ερμηνείας των δεδομένων λαμβάνοντας υπόψη το σύνολο της υπαρκτής γνώσης



---

# Ποιότητα Κλινικής Δοκιμής

- Σχεδιασμός Μελέτης
  - Πρωτόκολλο
  - Φόρμες Συλλογής στοιχείων
- Τρόπος Διεξαγωγής μιας Μελέτης
  - Έλεγχος Ποιότητας
  - Παρακολούθηση (Monitoring)
  - Εξωτερική Ίατρική Αξιολόγηση
- Αξιολόγηση και αναφορά αποτελεσμάτων
  - Ανάλυση
  - **Εξαιρέσεις**
  - Δημοσιεύσεις

# Εξαιρέση Περιπτώσεων από την Ανάλυση

## Βασικοί Λόγοι

- ❑ Ασθενείς που δεν πληρούν τα κριτήρια εισαγωγής στη Μελέτη (Ineligible)
- ❑ Ασθενείς που δεν έλαβαν «ικανοποιητική» θεραπεία σύμφωνα με το πρωτόκολλο (protocol variations) ή την καταστρατήγησαν (protocol violations)
- ❑ Ασθενείς με ελλιπή, μη επαρκή στοιχεία (π.χ. ο ασθενής χάθηκε από παρακολούθηση, Lost to Follow-up)
- ❑ Ασθενείς που αποσύρθηκαν από τη μελέτη – ακυρώσεις και αρνήσεις (early withdrawals)

---

## Εναλλακτικοί Τρόποι Ανάλυσης για κάθε Περίπτωση Εξαίρεσης

- Η ανάλυση περιλαμβάνει όλες τις Περιπτώσεις χωρίς Εξαίρεση
- Η ανάλυση ΔΕΝ περιλαμβάνει καμία από αυτές τις περιπτώσεις
- Επιλεκτικά εξαιρούνται κάποιες περιπτώσεις

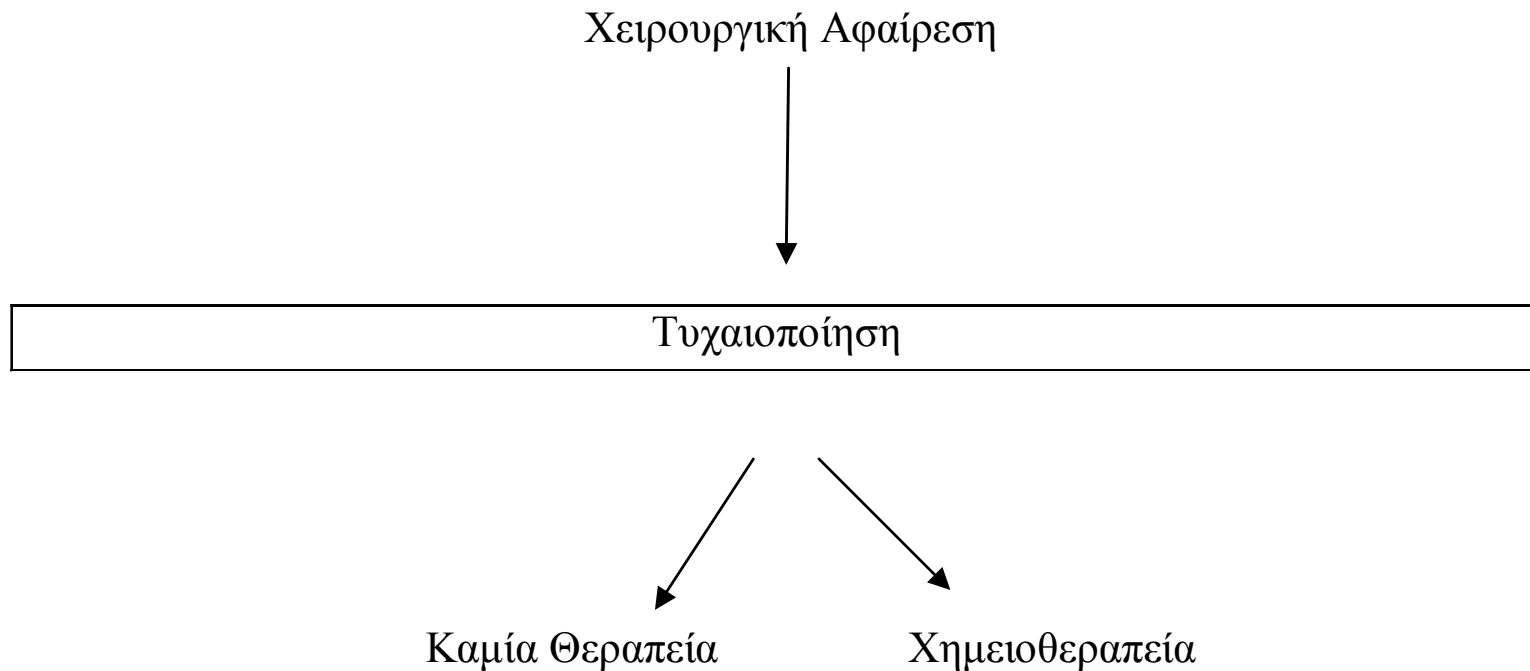
## Ασθενείς που δεν πληρούν τα κριτήρια εισαγωγής στη Μελέτη (Ineligible)

Τα κριτήρια καταλληλότητας των ασθενών σύμφωνα με το πρωτόκολλο είναι απαραίτητα για τον ξεκάθαρο ορισμό του πληθυσμού των ασθενών. Με αυτόν τον τρόπο παρέχεται η δυνατότητα γενίκευσης των συμπερασμάτων στον κατάλληλο πληθυσμό

- Κριτήρια κατά την εισαγωγή (χαρακτηριστικά που είναι γνωστά πριν την είσοδο στη μελέτη – ηλικία, στάδιο νόσου)
- Κριτήρια με κάποια καθυστέρηση (αναμενόμενη επιβίωση τουλάχιστον 3 μηνών), μπορεί να προκαλέσουν μεροληψία (bias)

# Παράδειγμα: Μεροληψία λόγω «καθυστερημένου» κριτηρίου εισαγωγής

Το πρωτόκολλο απαιτεί οι ασθενείς που παίρνουν χημειοθεραπεία να είναι ικανοί να αρχίζουν θεραπεία εντός 60 ημερών



---

## Παράδειγμα: Μεροληψία λόγω «καθυστερημένου» κριτηρίου εισαγωγής

**ΠΡΟΒΛΗΜΑ:** Η ομάδα «μη θεραπείας» δεν απαιτείται να πληροί αυτό το κριτήριο

**ΜΕΡΟΛΗΨΙΑ:** Οι ασθενείς με κακή φυσική κατάσταση εισέρχονται στην ομάδα «μη θεραπείας» αλλά όχι στην ομάδα «χημειοθεραπείας»

**ΛΥΣΗ:** Η τυχαιοποίηση διεξάγεται όταν ο ασθενής είναι έτοιμος για χημειοθεραπεία

---

## Ασθενείς που αποσύρθηκαν από τη μελέτη (early withdrawals)

- Συνήθεις Λόγοι

**Ακύρωση** – απόσυρση μεταξύ τυχαιοποίησης και έναρξης θεραπείας

- Θάνατος
- Χειροτέρευση Κατάστασης
- Αναγνώριση με καθυστέρηση μη πλήρωσης κριτηρίων εισαγωγής
- Απόσυρση συγκατάθεσης ασθενούς

**Άλλες Αποσύρσεις** – μεταξύ έναρξης και ολοκλήρωσης θεραπείας

- Απόσυρση συγκατάθεσης ασθενούς
- Περιπλοκές Θεραπείας
- Υποψία Αποτυχίας Θεραπείας

---

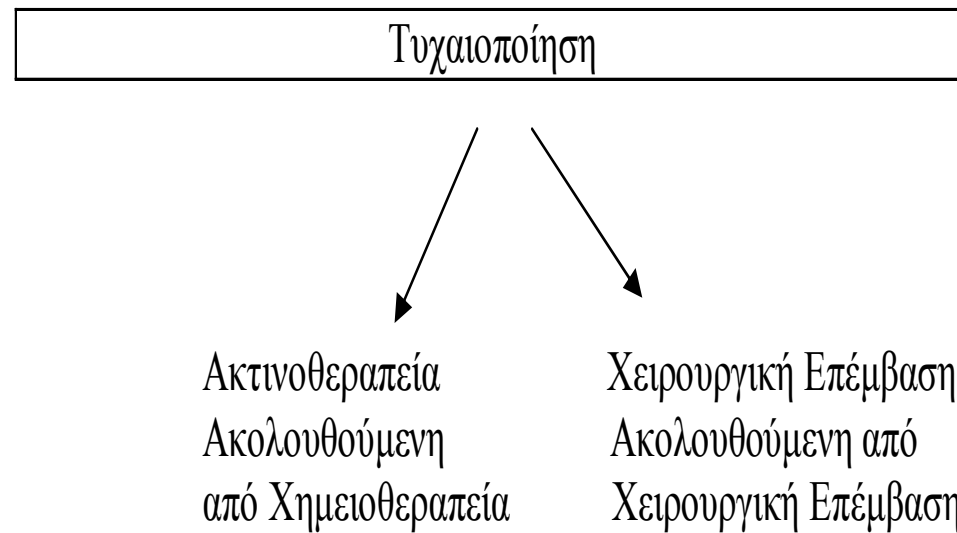
## Ασθενείς που αποσύρθηκαν από τη μελέτη (early withdrawals)

### ΟΔΗΓΙΕΣ:

- ❑ Οι ακυρώσεις συνήθως εξαιρούνται από την ανάλυση. Στις περισσότερες των περιπτώσεων οδηγούν σε αμελητέα μεροληψία.
- ❑ Εξαιρέσεις άλλων αποσύρσεων όμως συνήθως οδηγούν σε μεροληπτικά αποτελέσματα.
- ❑ Ακόμη και η εξαίρεση των ακυρώσεων θα μπορούσε να οδηγήσει σε μεροληψία.



# Ασθενείς που αποσύρθηκαν από τη μελέτη (early withdrawals)



---

## Ασθενείς που αποσύρθηκαν από τη μελέτη (early withdrawals)

### Παράδειγμα: Καρκίνος Κεφαλής-Τραχήλου

#### Προβλήματα:

1. Η άρνηση θεραπείας είναι πιθανόν να διαφέρει μεταξύ ομάδων τυχαιοποίησης
2. Οι μηχανισμοί άρνησης είναι διαφορετικοί μεταξύ των δύο ομάδων

#### Συμπέρασμα:

Οι ασθενείς τυχαιοποιούνται σε θεραπευτικό πρόγραμμα και όχι θεραπευτική αγωγή. Όλοι οι ασθενείς που τυχαιοποιούνται παρακολουθούνται και περιλαμβάνονται στη τελική αναφορά

# Περίληψη και Οδηγίες

- Όλες οι εξαιρέσεις πρέπει να αναφέρονται με λεπτομέρεια – Συγκρίσεις έναντι εκτιμήσεων
- Οι κανόνες εξαιρέσεων πρέπει να αναφέρονται και στο πρωτόκολλο για να αποφεύγεται διαφορετική τακτική
- Το να συμπεριληφθεί ένας ασθενής στην ανάλυση δεν πρέπει να εξαρτάται από τη θεραπεία ή την έκβαση
- Εξαίρεση προβληματικών περιπτώσεων μπορεί να εισαγάγει μεροληψία
- Επιλεκτική εξαίρεση είναι υποκειμενική
- Μεγάλος αριθμός εξαιρέσεων δημιουργεί αμφιβολίες για τη συνολική ποιότητα της μελέτης

---

*Κατευθυντήριες οδηγίες δημοσίευσης κλινικών δοκιμών ογκολογίας: Ευθύνες των εκδοτών και ερευνητών (Zelen M.)*

- Πληθυσμός ασθενών
- Θεραπεία (που έχει δοθεί στην πραγματικότητα)
- Σχεδιασμός της μελέτης
- Καταμέτρηση ασθενών (registration system)
- Παρακολούθηση (follow-up)
- Ποιότητα Δεδομένων
- Τελικά σημεία και περικοπή παρατηρήσεων(endpoints)
- Μεροληψία μη τυχαιοποίησης της μελέτης
- Στατιστικοί μέθοδοι και Ισχύς
- Διακοπή της μελέτης
- Προκαταρκτικά έναντι τελικών αποτελεσμάτων

*J Clin Onc, Vol. 1, No. 2, February, 1983*

---

---

## Σημαντικά σημεία προς έλεγχο κατά την αξιολόγηση των Κλινικών Δοκιμών

- Εξαιρέσεις από την ανάλυση
- Τυχαιοποίηση
- Συμπεράσματα από Μελέτες Υπεροχής Έναντι Μελετών Ισοδυναμίας
- Υποκατάστατα Μέτρα Έκβασης
- Ανάλυση Υποομάδων

---

## Ερώτηση VI

Η φερεγγυότητα μιας Τυχαιοποιημένης Κλινικής Δοκιμής διασφαλίζεται από:

- Α. Περιορισμένο Πληθυσμό Αναφοράς
- Β. Μεγάλο Μέγεθος Δείγματος
- Γ. Στατιστικά σημαντικό αποτέλεσμα
- Δ. Όλα τα ανωτέρω
- Ε. Κανένα από τα ανωτέρω

---

## Ερώτηση VII

Η κύρια ανάλυση μιας Τυχαιοποιημένης Κλινικής Δοκιμής Υπεροχής πρέπει να περιλαμβάνει:

- Α. μόνο ασθενείς που παρακολούθηθηκαν για ικανοποιητικό χρόνο
- Β. μόνο ασθενείς που πήραν τη θεραπεία
- Γ. όλους τους ασθενείς που τυχαιοποιήθηκαν ανεξαρτήτως αν πήραν τη θεραπεία ή παρακολούθηθηκαν για ικανοποιητικό χρόνο
- Δ. Όλα τα ανωτέρω
- Ε. Κανένα από τα ανωτέρω

---

## Ερώτηση VIII

**Η Τυχαιοποίηση για μία αρκετά μεγάλου μεγέθους Κλινική Δοκιμή διασφαλίζει:**

- **Α.** ότι σε κάθε κέντρο θα υπάρχει ακριβώς ο ίδιος αριθμός ασθενών στη θεραπεία Α και στη θεραπεία Β
- **Β.** την ύπαρξη μεροληψίας στην επιλογή ασθενών
- **Γ.** ότι οι άγνωστοι συγχυτικοί παράγοντες κατανέμονται εξίσου στις ομάδες σύγκρισης
- **Δ.** Όλα τα ανωτέρω
- **Ε.** Κανένα από τα ανωτέρω



## Ερώτηση ΙΧ

Να σημειώσετε ποιά από τις κατωτέρω δηλώσεις είναι σωστή (μόνο μία):

- Α. Στην περίληψη μιας Κλινικής Δοκιμής καλό είναι να παρουσιάζουμε μόνο τα στατιστικά σημαντικά αποτελέσματα και όχι τα μη στατιστικά σημαντικά ιδιαίτερα όταν έχουν προκύψει από πολλαπλές αναλύσεις
- Β. Στην περίληψη μιας Κλινικής Δοκιμής καλό είναι να παρουσιάζουμε μόνο τα στατιστικά σημαντικά αποτελέσματα και όχι τα μη στατιστικά σημαντικά ιδιαίτερα όταν έχουν προκύψει από την κύρια ανάλυση
- Γ. Στη συζήτηση, δεν πρέπει να αναφέρονται τυχόν προβλήματα που έχουν εμφανιστεί κατά τη διεξαγωγή της Κλινικής Δοκιμής γιατί μπορεί να μη δημοσιευτεί η μελέτη
- Δ. Όλα τα ανωτέρω
- Ε. Κανένα από τα ανωτέρω