ΚΛΙΝΙΚΟ ΦΡΟΝΤΙΣΤΗΡΙΟ: ΣΤΑΤΙΣΤΙΚΗ-ΕΠΙΔΗΜΙΟΛΟΓΙΑ
ΕΥΑΓΓΕΛΙΣΜΟΣ, ΝΟΕΜΒΡΙΟΣ 2016

Πως Διαβάζουμε μία Μελέτη: Ποιοτικός Έλεγχος και Κριτική Εκτίμηση μιας Εργασίας

Γιώτα Τουλούμη
Καθηγήτρια Βιοστατιστικής και Επιδημιολογίας
Εργ. Υγιεινής, Επιδημιολογίας και Ιατρικής Στατιστικής
Ιατρική Σχολή Πανεπιστημίου Αθήνας

gtouloum@med.uoa.gr

Γιατί να δημοσιεύουμε τα αποτελέσματα της έρευνας μας?

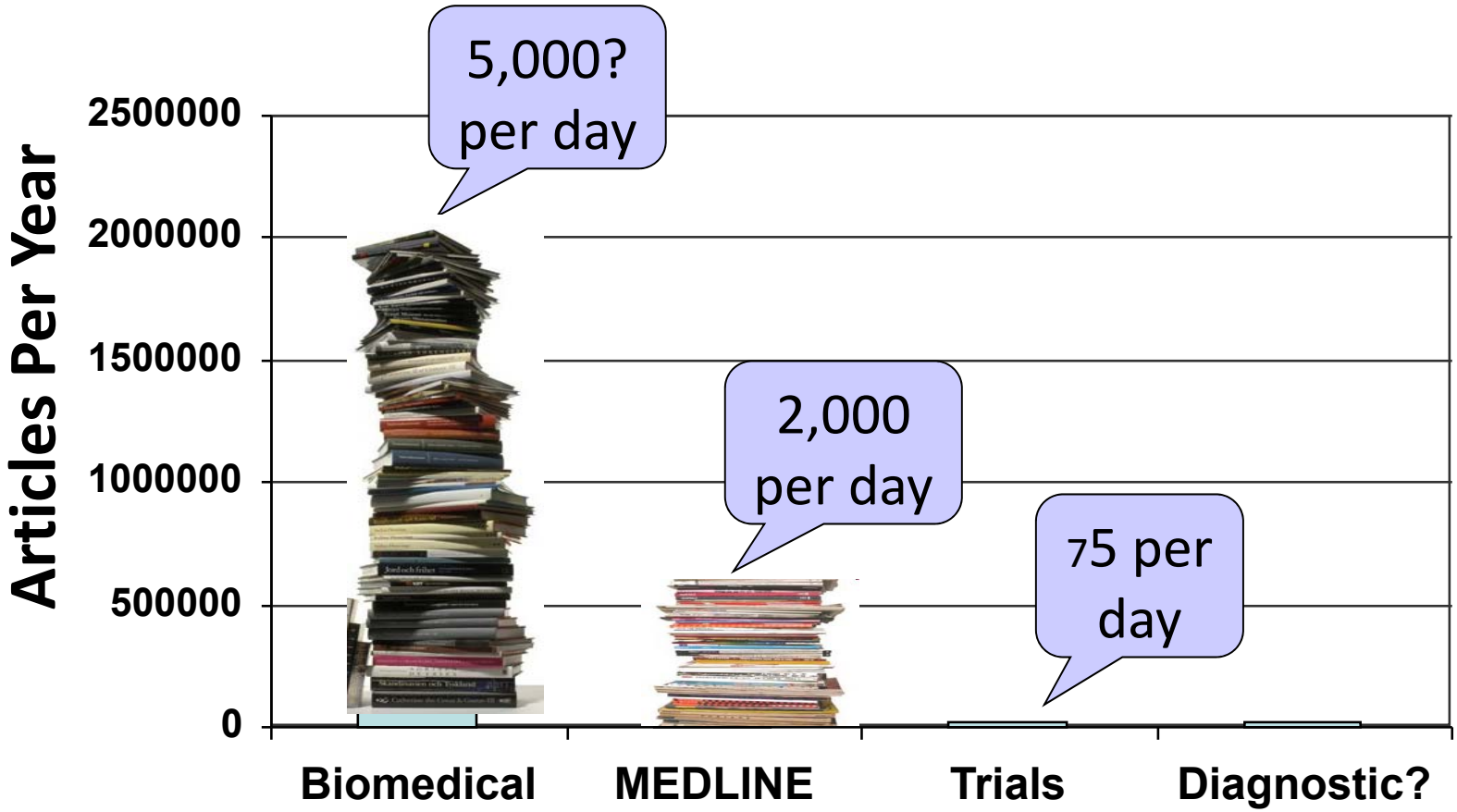
- Ο βασικός και κύριος στόχος της πραγματοποίησης μιας έρευνας είναι η πρόοδος της γνώσης σχετικά με το αντικείμενό της
- Το μεγαλύτερο μέρος των προληπτικών, διαγνωστικών και θεραπευτικών αποφάσεων σήμερα στηρίζεται στα «τεκμήρια», στα αποτελέσματα των ερευνών που δημοσιεύονται και αξιολογούνται ως σημαντικά
- (τεκμηριωμένη ιατρική, evidence-based medicine)

The scandal of poor medical research

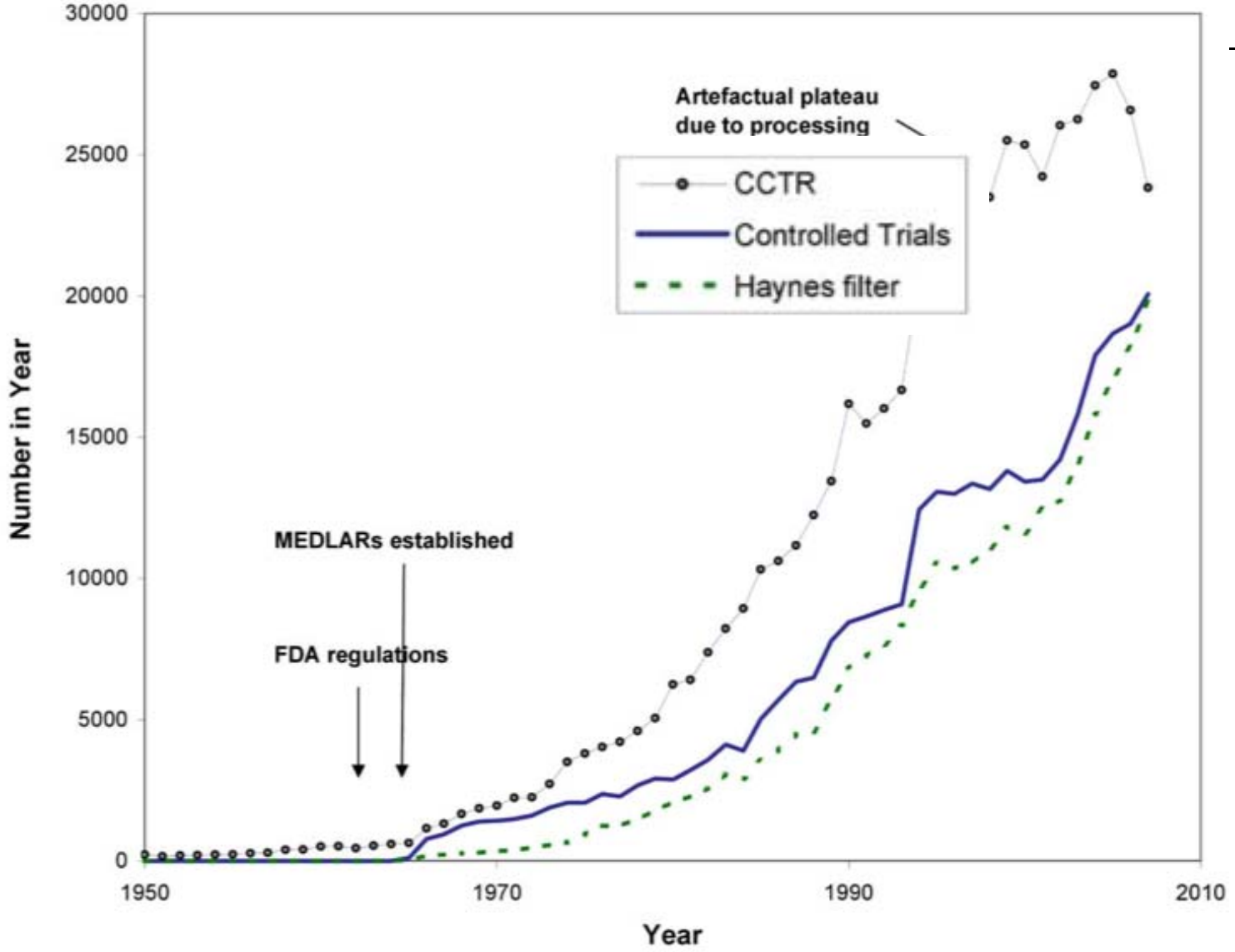
"When I tell friends outside medicine that many papers published in medical journals are misleading because of methodological weaknesses they are rightly shocked....Why are errors so common? Put simply, much poor research arises because researchers feel compelled for career reasons to carry out research that they are ill equipped to perform, and nobody stops them."

DG Altman, BMJ 1994;308:283

why do we need to use evidence efficiently?



Clinical evidence increasing so rapidly we need better skills to keep up-to-date more efficiently than previous generations of clinicians



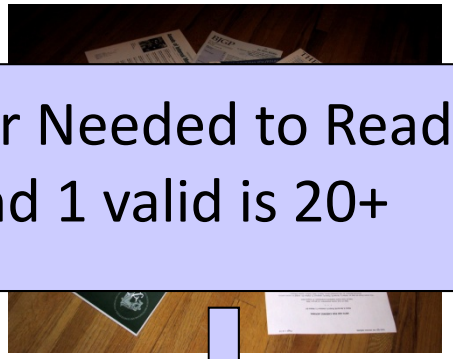
Bastian, Glasziou, Chalmers PLoS 2010 Vol 7 | Issue 9 | e1000326

using evidence more critically

validity: most articles should be ignored

EBM Journal Process

- 140+ journals scanned
  - 60,000 articles
- Is it **valid**? (<5%)
  - Intervention: RCT
  - Prognosis: inception cohort
  - Etc
- Is it **relevant?**
  - 6-12 GPs & specialists asked: Relevant? Newsworthy?
- < 0.5% selected



Number Needed to Read to find 1 valid is 20+

Number Needed to Read to find 1 valid & relevant is 200+

Avoidable waste in the production and reporting of research evidence

Iain Chalmers, Paul Glasziou

Lancet 2009; 374: 86–89

Without accessible and usable reports, research cannot research involving patients have been powerful

bmj.com

The scandal of poor epidemiological research

Erik von Elm and Matthias Egger

BMJ 2004;329;868-869
doi:10.1136/bmj.329.7471.868

Research: increasing value, reducing waste 2

Increasing value and reducing waste in research design, conduct, and analysis

John P A Ioannidis, Sander Greenland, Mark A Hlatky, Muin J Khoury, Malcolm R Macleod, David Moher, Kenneth F Schulz, Robert Tibshirani

Lancet 2014; 383: 166–75

Correctable weaknesses in the design, conduct, and analysis of biomedical and public health research studies can

Currently, many published research findings are false or exaggerated, and an estimated 85% of research resources are wasted.

Ioannidis, PLoS Medicine, 2014:e1001747

but many clinicians cannot tell good from poor quality research

BMJ study of 607 reviewers

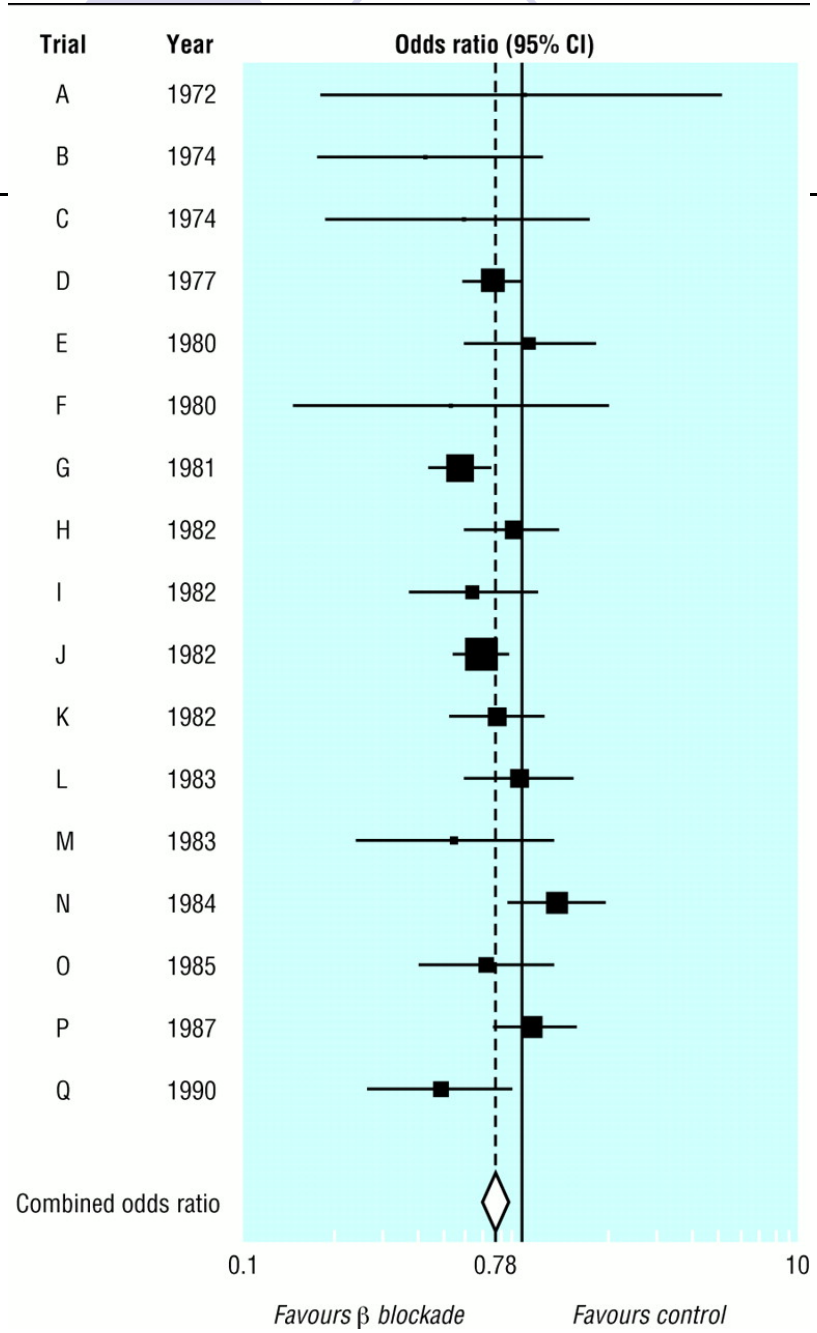
- 14 deliberate errors inserted

Detection rates

- On average <3 of 9 major errors detected
- Poor Randomisation (by name or day) - 47%
- Not intention-to-treat analysis - 22%
- Poor response rate - 41%

Schroter S et al, accepted for Clinical Trials

using evidence more systematically: meta-analysis



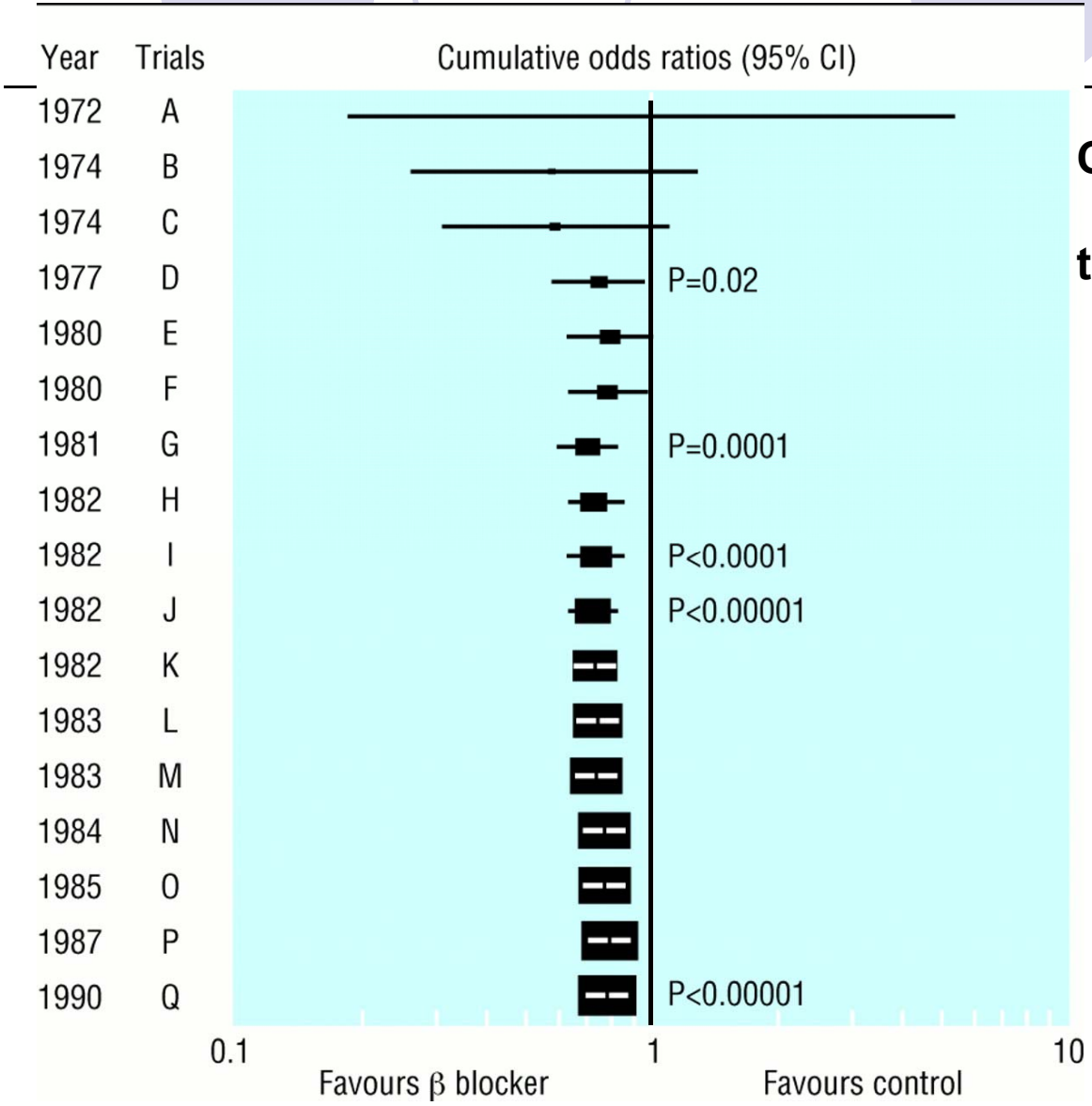
Total mortality from trials of β-blockers in 2° prevention after MI.

Black square & horizontal line correspond to odds ratio (OR) & 95% confidence interval (CI) for each trial. The size of the black square reflects the weight of each trial. The diamond represents the combined OR & 95 CI, showing a 22% reduction in odds of death

Egger, M. et al. BMJ 1997;315:1533-1537

BMJ

using evidence more systematically: cumulative meta-analysis



Cumulative meta-analysis of total mortality results from trials of oral β-blockers after MI.

Size of squares reflect amount of statistical information available at a given point in time

Egger, M. et al. BMJ 1997;315:1371-1374

BMJ

Κύρια Προβλήματα

- Μη Κατάλληλος Σχεδιασμός
- Μη Αντιπροσωπευτικό Δείγμα
- Μικρό Μέγεθος Δείγματος
- Μη Έγκυρη Στατιστική Ανάλυση
- Μη Κατάλληλη Ερμηνεία Αποτελεσμάτων

Επιστημονική Ερώτηση

- **Πρωτότυπη**
- **Προάγει τη γνώση**
- Ψάρεμα ερωτήσεων από δεδομένα ρουτίνας **δεν συνιστάται**
  - Ελλειπείς Πληροφορίες
  - Συστηματικά Σφάλματα / συγχυτικοί παράγοντες

Σφάλματα (Errors)

- **Τυχαία Σφάλματα (Random Errors)**
  - Οφείλονται στη τύχη
- **Συστηματικά Σφάλματα (Systematic Errors, Bias)**
  - Σφάλματα στη Συλλογή, Ανάλυση, Ερμηνεία, Δημοσίευση ή Μετα-ανάλυση δεδομένων που οδηγούν σε συμπεράσματα που συστηματικά διαφοροποιούνται από την «αλήθεια» (Lost 2001)
- **Συγχυτικοί Παράγοντες (Confounding)**
  - Η σχέση μεταξύ του παράγοντα που μελετάμε (έκθεση) και της έκβασης (νόσο) συγχέεται από την παρουσία άλλου (τρίτου) παράγοντα

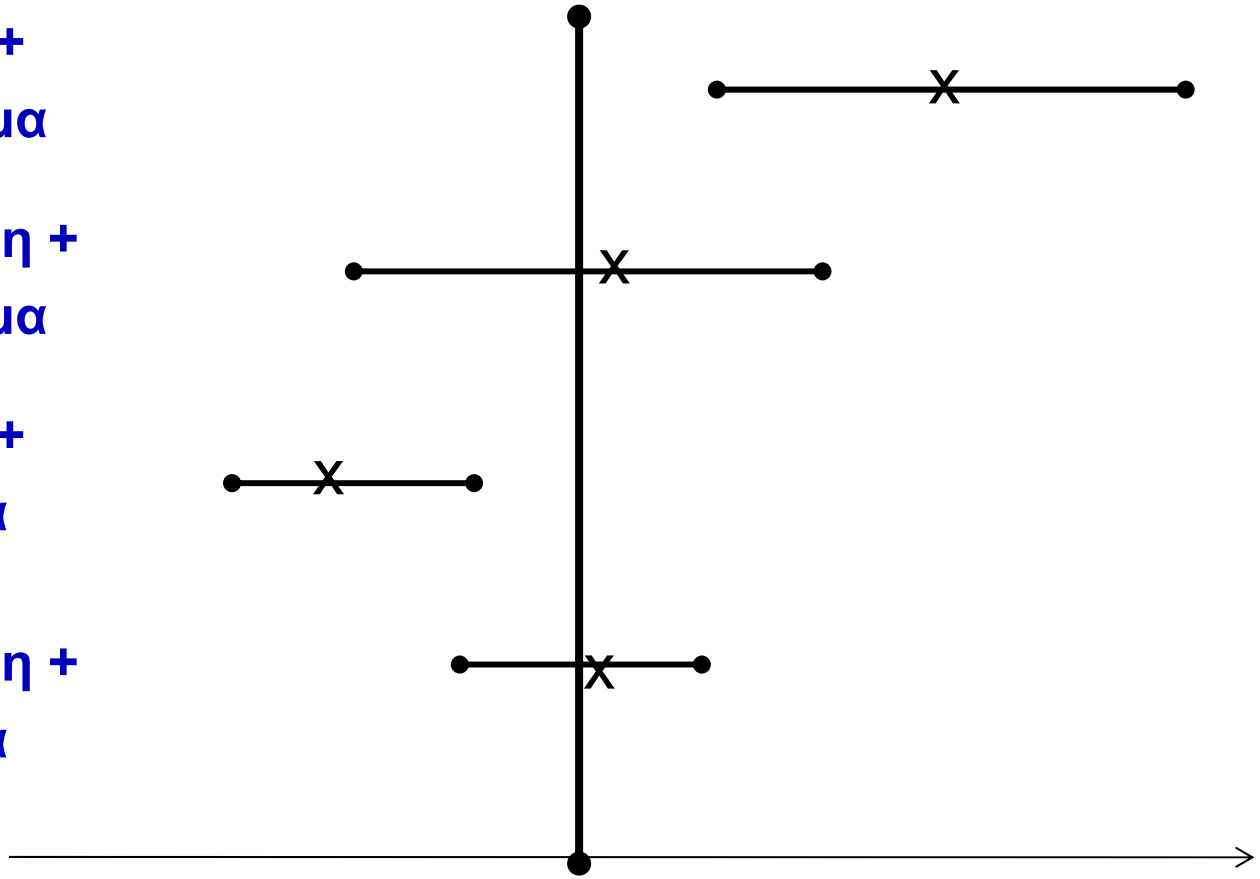
Η έννοια του τυχαίου σφάλματος και του συστηματικού σφάλματος

Συστηματικό σφάλμα +
Μεγάλο πιθανό σφάλμα

Ανεπηρέαστη εκτίμηση +
Μεγάλο πιθανό σφάλμα

Συστηματικό σφάλμα +
Μικρό πιθανό σφάλμα

Ανεπηρέαστη εκτίμηση +
Μικρό πιθανό σφάλμα

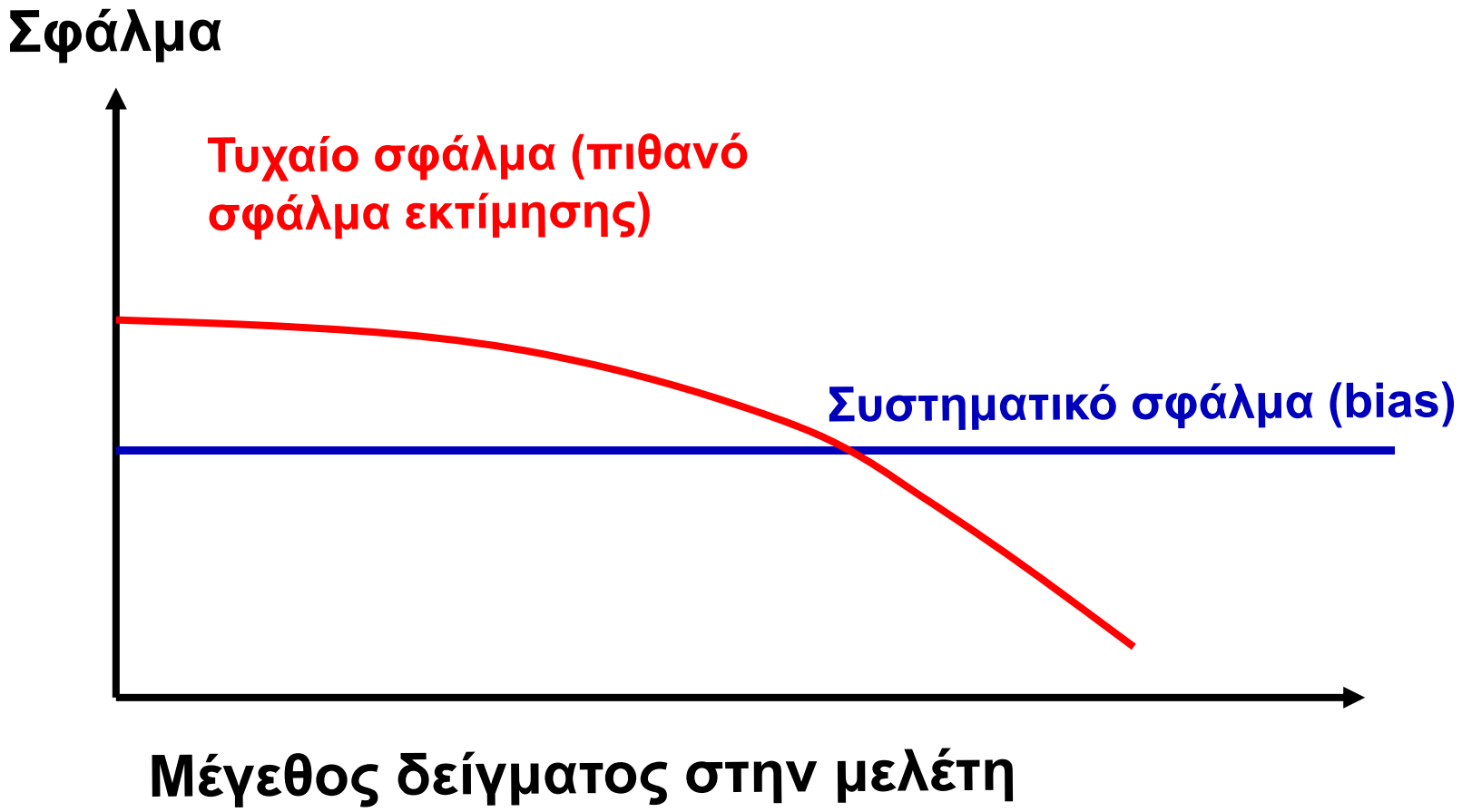


Πραγματική σχέση

Είδη Συστηματικών Σφαλμάτων

- **Σφάλμα Επιλογής (selection bias)**
  - Συστηματικές διαφορές στα χαρακτηριστικά συμμετεχόντων και μη συμμετεχόντων στη μελέτη
- **Σφάλμα Παρατήρησης ή Πληροφορίας (Information bias)**
  - Σφάλμα στη μέτρηση της έκβασης ή της έκθεσης με αποτέλεσμα διαφορετική ακρίβεια μετρήσεων στις συγκρινόμενες ομάδες
- **Δυσταξινόμηση (missclassification)**
  - Σφάλμα στη κατάταξη της έκθεσης ή της νόσου

Σφάλματα



Πηγή: Rothman, 2002

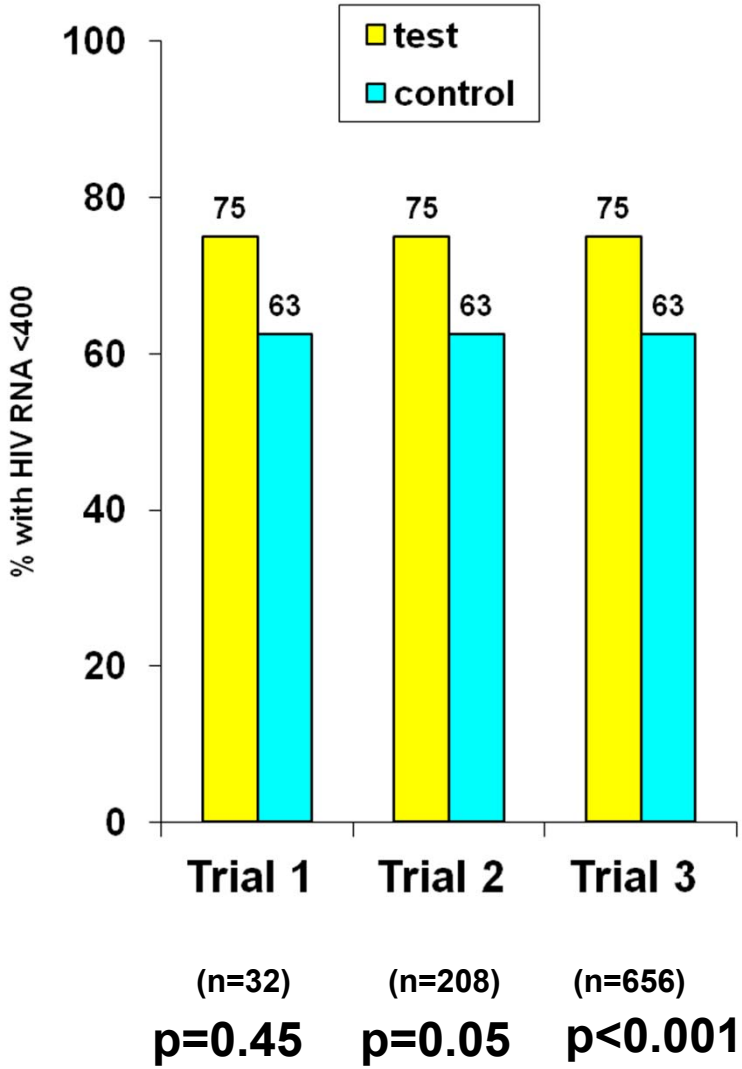
Υπολογισμός μεγέθους δείγματος (sample size)

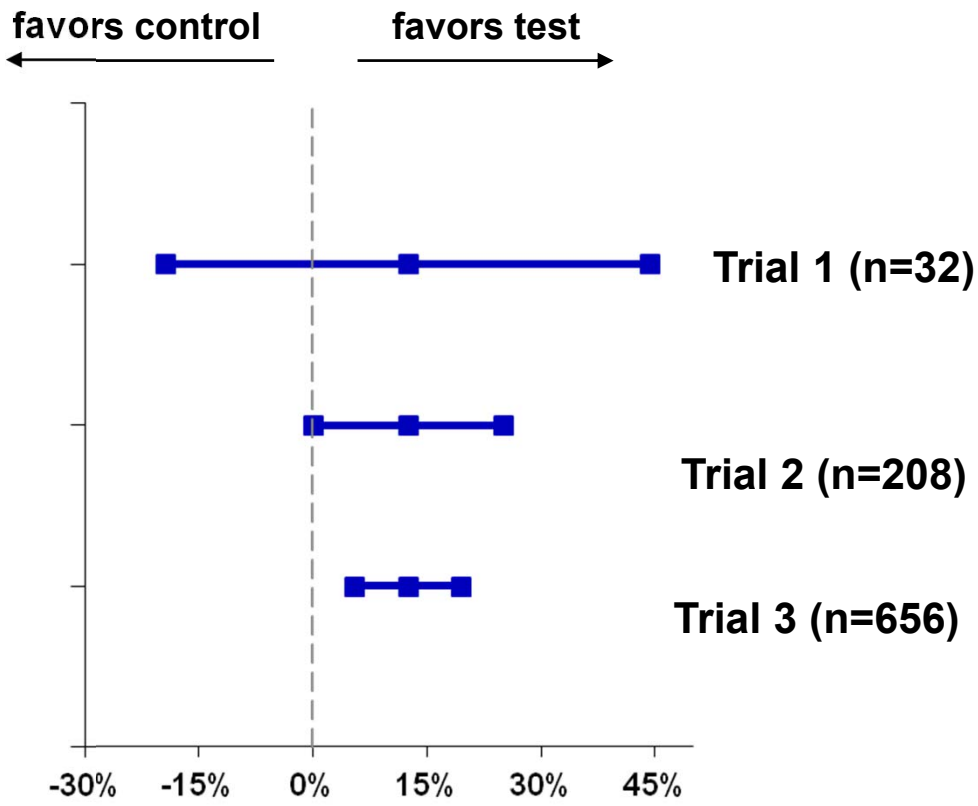
Γιατί πρέπει να εκτιμήσουμε το απαιτούμενο μέγεθος δείγματος **ΠΡΙΝ** την έναρξη της μελέτης;

- **Μικρό δείγμα:**
  - Η μελέτη μπορεί να αποτύχει να απαντήσει το ερώτημα
- **Μεγάλο δείγμα:**
  - Κόστος, χρόνος
- Ο σκοπός είναι να εκτιμηθεί ο «κατάλληλος» αριθμός ασθενών που απαιτούνται για τη μελέτη

P-value και όρια αξιοπιστίας ανάλογα με το μέγεθος του δείγματος

ΙΣΧΥΣ: Πιθανότητα να ανιχνεύσω μια διαφορά δεδομένου ότι αυτή υπάρχει





95% CI for difference in response rates

Άλλα σφάλματα σχεδιασμού

- Τυχαιοποίηση
  - Σημαντικές διαφορές στα βασικά χαρακτηριστικά των συγκρινόμενων ομάδων
- Τυφλότητα
- Ορισμός των εκβάσεων
  - Υποκειμενική
  - Κλινική έκβαση / αναπληρωτής δείκτης (surrogate marker)
  - Απλή ή σύνθετη έκβαση
  - % ανταπόκρισης / Μέση διαφορά

Θέματα Ηθικής

- Έντυπο ενημέρωσης και συναίνεσης συμμετεχόντων
- Έγκριση από κατάλληλη επιτροπή Ηθικής και Δεοντολογίας
- Απαραίτητες προϋποθέσεις σε όλες τις μελέτες (όχι μόνο σε RCTs)
- Προϋπόθεση για τη δημοσίευση των αποτελεσμάτων

Υπερβολική έμφαση στα p-values

- Συχνά δίνονται οι τιμές των p-values χωρίς πολλές φορές να δίνεται καν το είδος της σχέσης
  - π.χ. Response to treatment was associated with age (p=0.002)

- Η πληροφορία για τη σχέση θα πρέπει να δίνεται με στοιχεία που να επιτρέπουν στον αναγνώστη να καταλάβει το είδος και το μέγεθος της σχέσης (π.χ. το odds ratio με το 95% CI)
  - π.χ. Patients older than 40 years old had 2.1 (95% CI: 1.4, 3.3) higher risk of relapse as compared to patients <40 yrs (p=0.002)

P-values και 95% CIs

- **Uniform Requirements for Manuscripts Submitted to Biomedical Journals:** "When possible, quantify findings and present them with appropriate indicators of measurement error or uncertainty (such as confidence intervals).

  Avoid relying solely on statistical hypothesis testing, such as *P values, which fail to convey important information* about effect size"

- **Consort statement:** "Authors should accompany this [the estimate of the treatment effect] by a confidence interval for the estimated effect, which indicates a central range of uncertainty for the true treatment effect"

Πολλαπλές συγκρίσεις

Treatment	Response	
	Yes	No
Drug A	180 (60.0%)	120 (40.0%)
Drug B	170 (56.7%)	130 (43.3%)
Drug C	130 (43.3%)	170 (56.7%)
Drug D	130 (43.3%)	170 (56.7%)

$p<0.001$

Αυτό σημαίνει ότι υπάρχει στατιστικά σημαντική σχέση μεταξύ θεραπείας και έκβασης ΑΛΛΑ δεν δίνει πληροφορία για τις επιμέρους συγκρίσεις π.χ.

A vs. B A vs. C

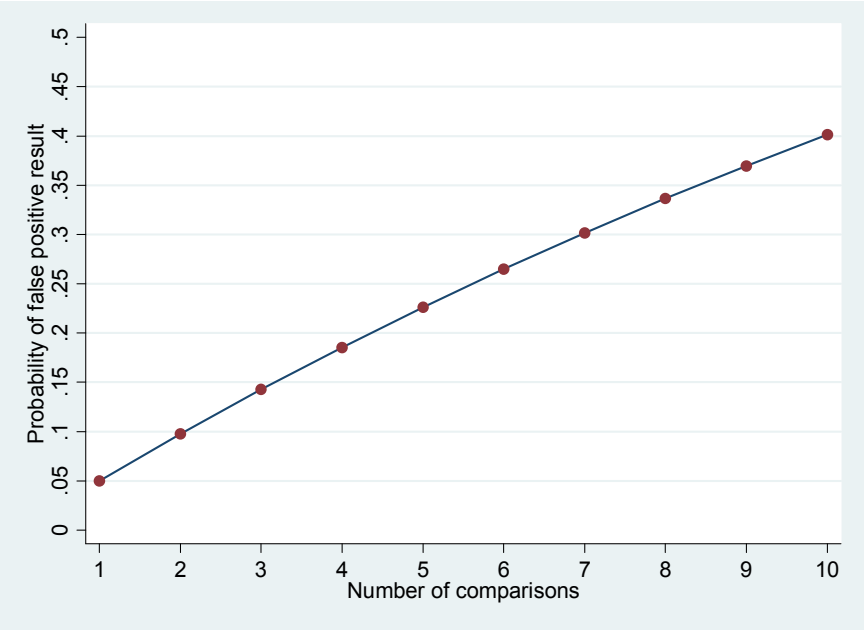
A vs. D B vs. C

B vs. D C vs. D

Πολλαπλές συγκρίσεις

- Επίπεδο σημαντικότητας 0.05 σημαίνει ότι η πιθανότητα σε ένα στατιστικό τεστ να βρεθεί μία στατιστικά σημαντική σχέση κατά τύχη είναι 5%

  - Αν πραγματοποιήσουμε περισσότερα από ένα τεστ, η πιθανότητα να βρεθεί μία στατιστικά σημαντική σχέση κατά τύχη είναι >5%



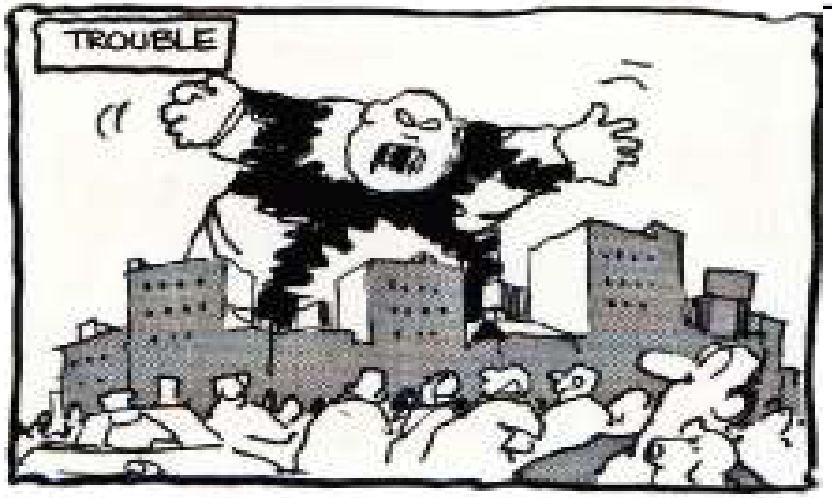
  - Αποφεύγουμε τις πολλαπλές συγκρίσεις ή π.χ. πολλαπλές συγκρίσεις με κατάλληλο επίπεδο σημαντικότητας (π.χ. Bonferroni correction, αν και αρκετά συντηρητική μέθοδος)

Άλλα σφάλματα κατά τη στατιστική ανάλυση

- Κατάλληλη στατιστική μεθοδολογία
  - Time to response vs % ανταπόκρισης vs μέση διαφορά
  - Απλή vs σύνθετη έκβαση
  - Εξαρτάται από το είδος της έκβασης

- Χειρισμός ελλειπουσών τιμών / Μη-συμμόρφωσης / Αλλαγών στη θεραπεία

Ελλείπουσες τιμές: Γιατί είναι σημαντικές;



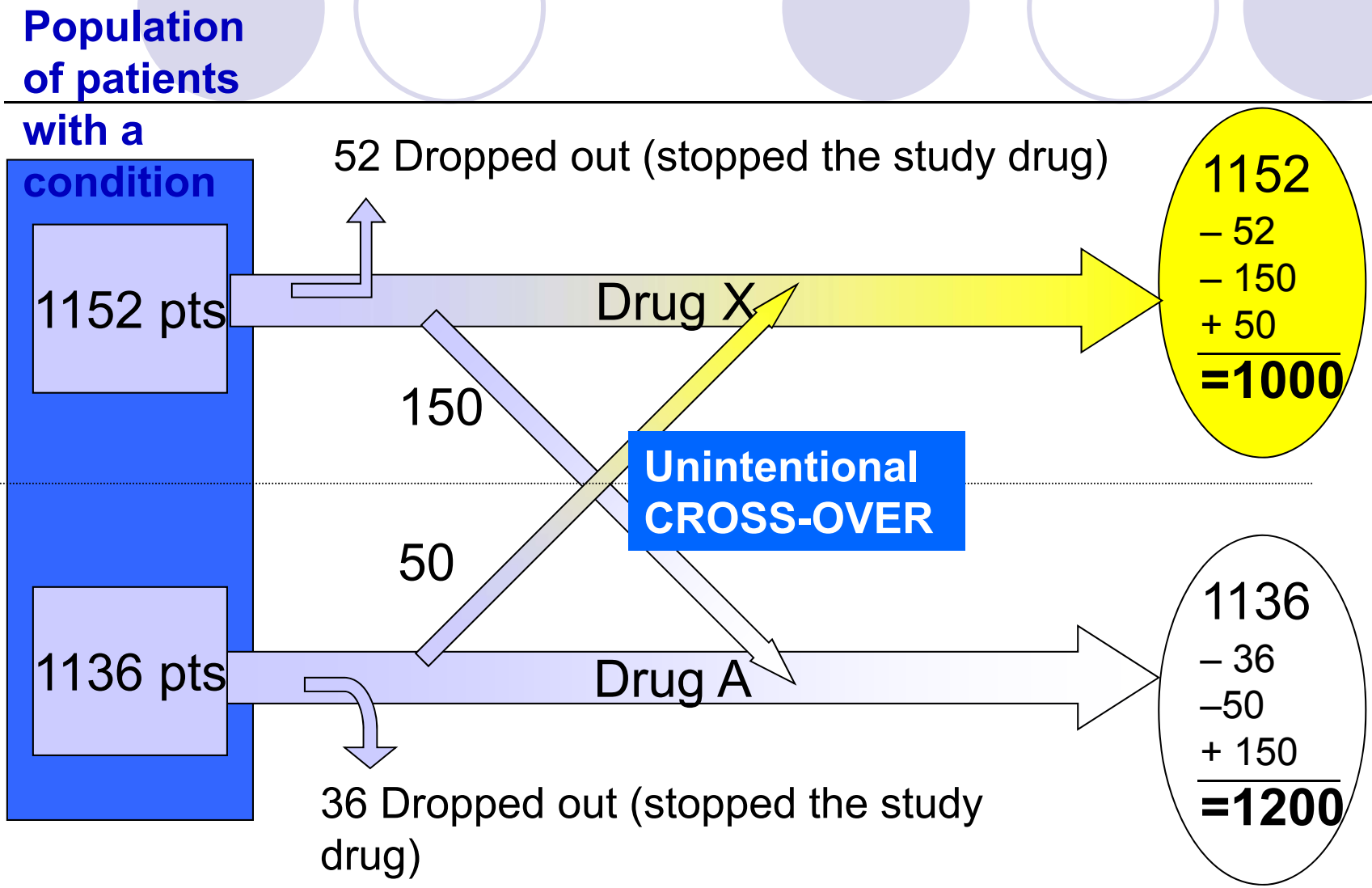
Συχνές σε βιοϊατρικά δεδομένα

- Μη-ανταπόκριση, διακοπή παρακολούθησης
- Μη καταγραφή επιμέρους δεδομένων
- Μελέτες πεδίου (π.χ. επιπολασμού)
  - Άρνηση απάντησης σε συγκεκριμένες ερωτήσεις
  - Άρνηση συμμετοχής στη μελέτη

Γιατί οι ελλείπουσες τιμές είναι πρόβλημα;

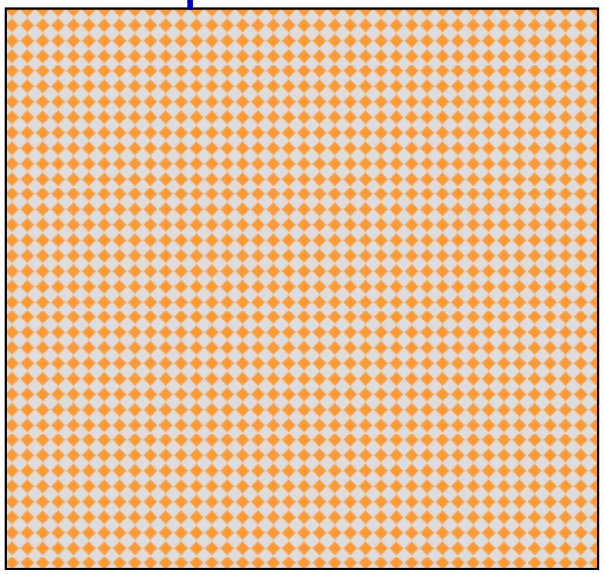
1. Μείωση Ισχύος
   - Δεν μπορεί να επανακτηθεί κατά την ανάλυση
2. Οποιαδήποτε ανάλυση στηρίζεται σε μη-ελεγχόμενες (untestable) υποθέσεις για τις ελλείπουσες τιμές
   - Λάθος υποθέσεις $\Rightarrow$ εσφαλμένες εκτιμήσεις
3. Μερικές δημοφιλείς μέθοδοι ανάλυσης εκτιμούν εσφαλμένα πιθανά σφάλματα (Standard Errors)
   - Ως αποτέλεσμα τα επίπεδα σημαντικότητας (P-values) και τα διαστήματα εμπιστοσύνης (confidence intervals) είναι εσφαλμένα
4. Μερικές δημοφιλείς μέθοδοι ανάλυσης είναι μη-επαρκείς
   - Διαστήματα εμπιστοσύνης ευρύτερα από ότι χρειάζεται

Per-Protocol (Βάση Πρωτοκόλλου)



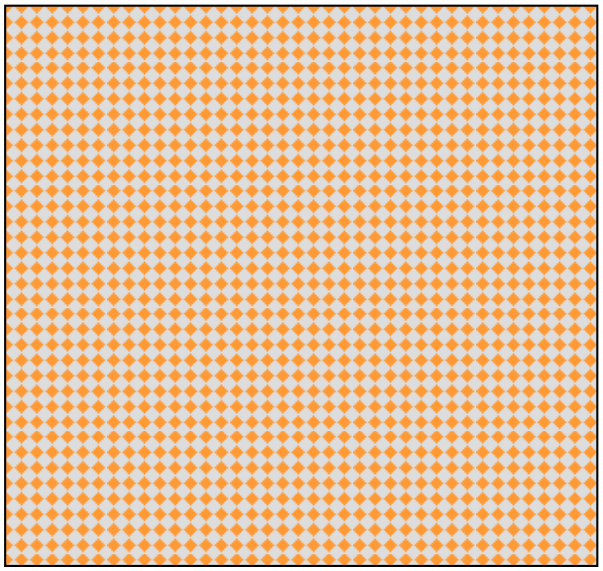
Drop outs change prognosis of original group

Experimental Group



Ave Age = 58

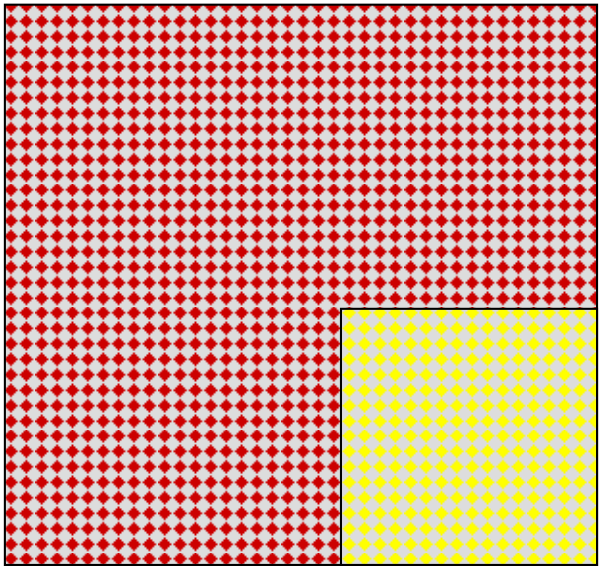
Control Group



Ave Age = 57

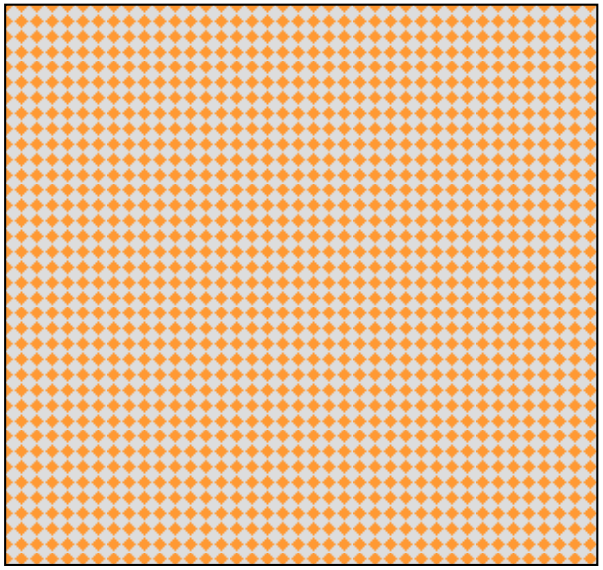
Drop outs change prognosis of original group

Experimental Group



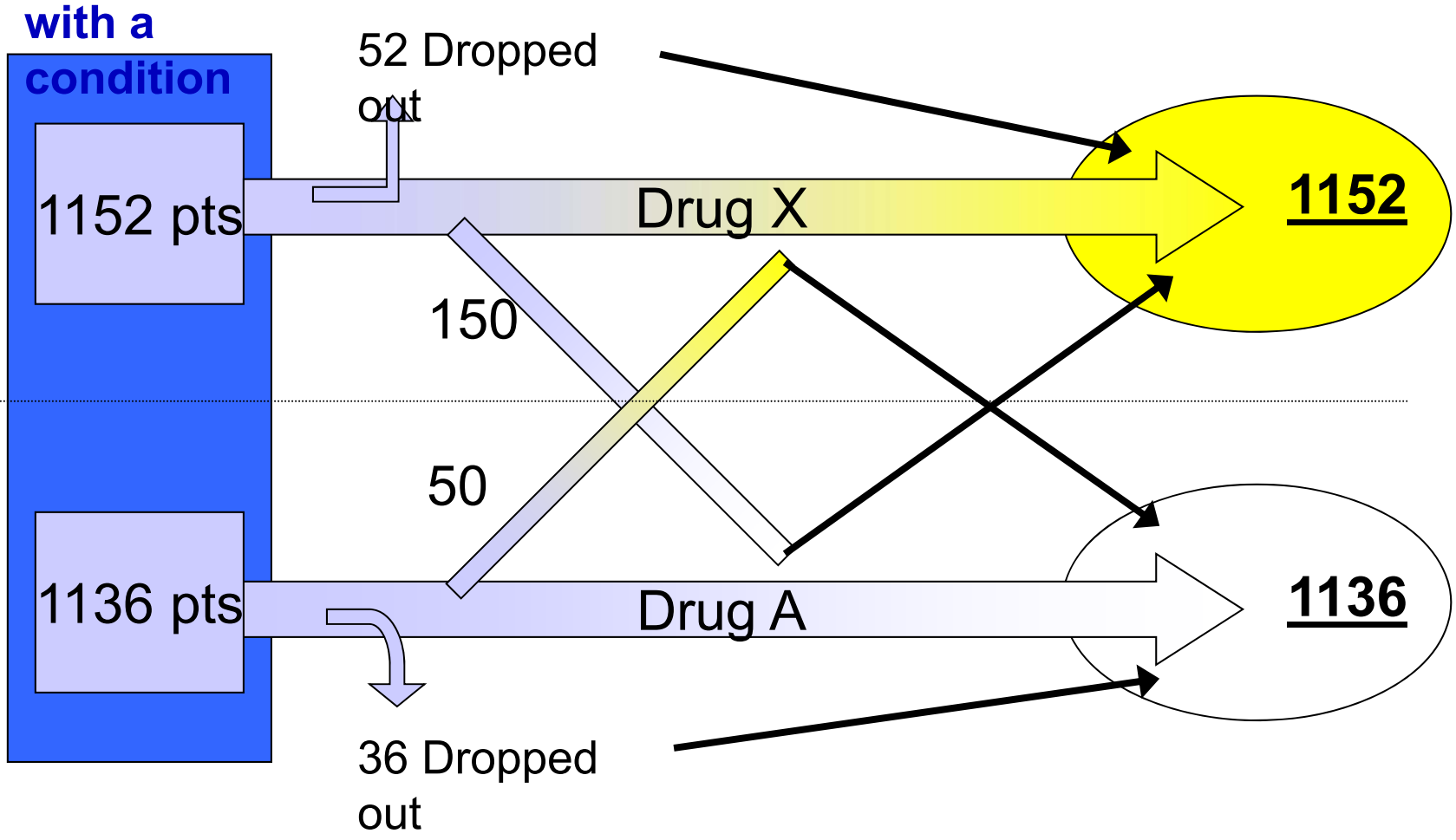
Ave Age = 51 Ave Age = 72

Control Group



Ave Age = 57

Intention-to-Treat (Κατά πρόθεση Θεραπείας)



Στατιστική Ανάλυση: Παράδειγμα

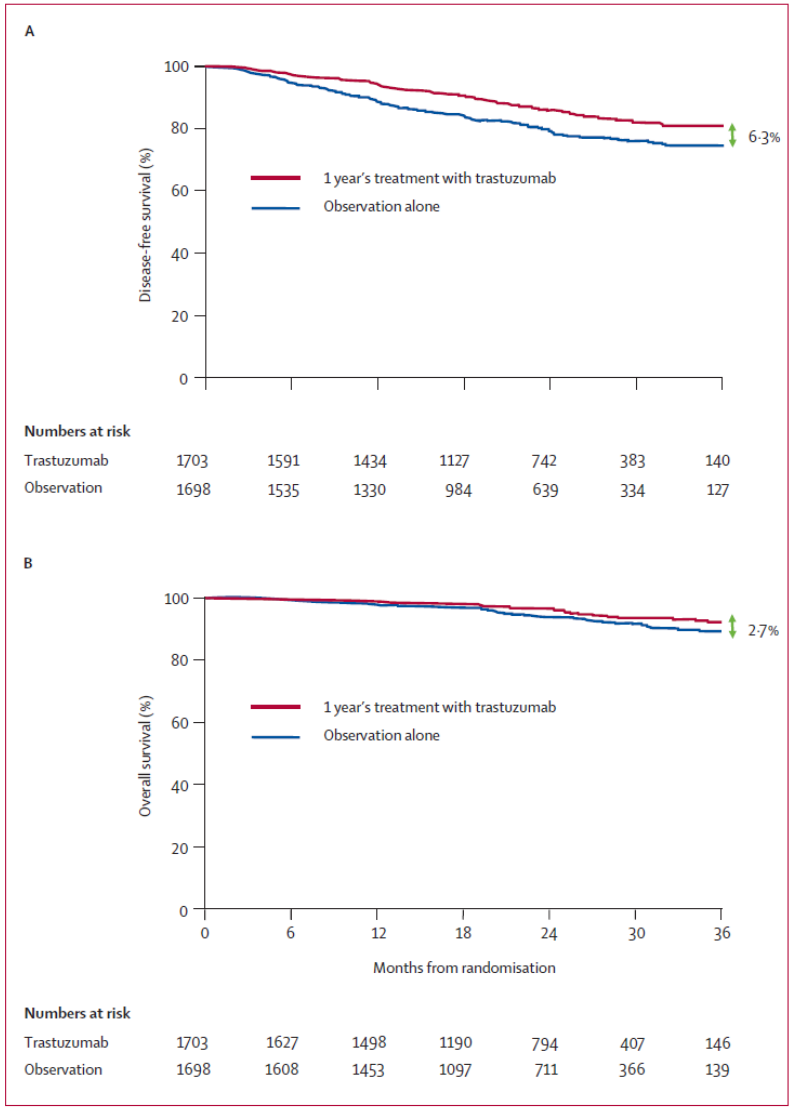


Figure 2: **Kaplan-Meier estimates of disease-free survival and overall survival**
(A) Disease-free survival for 1 year of trastuzumab *vs* observation with a median follow-up of 2 years. (B) Overall survival for 1 year of trastuzumab *vs* observation with a median follow-up of 2 years.

the site of first disease-free survival events. The unadjusted HR for the risk of an event in the trastuzumab group compared with the observation group was 0·64 (95% CI 0·54–0·76; p<0·0001 by the log rank test), which corresponds to an absolute disease-free survival benefit of 6·3% (80·6% *vs* 74·3%) at 3 years (figure 2A).

In addition to the main intent-to-treat analysis, an analysis that censored women at the time of switching to trastuzumab has also been done to compensate for a potential effect of delayed administration of trastuzumab.

Smith I et al. Lancet 2007; 369:29-36

Τι σημαίνει ITT ανάλυση όταν η έκβαση δεν είναι γνωστή για κάποια άτομα;

- In new advice, the European Medicines Agency (2010) takes a more relaxed view:
  - "Full set analysis generally requires the imputation of values or modelling for the unrecorded data"
- The CONSORT (2010) statement says:
  - "We replaced mention of 'intention to treat' analysis, a widely misused term, by a more explicit request for information about retaining participants in their original assigned groups"

Στρατηγική για ITT ανάλυση σε μη-πλήρη δεδομένα

1. Προσπάθεια **παρακολούθησης** όλων των τυχαιποιημένων ατόμων ακόμα και αν σταματήσουν την θεραπεία
2. Βασική ανάλυση όλων δεδομένων έχοντας **ευλογοφανείς** υποθέσεις για τα δεδομένα που λείπουν
3. Ανάλυση ευαισθησίας (**sensitivity analyses**)
4. **Όλα τα τυχαιοποιημένα άτομα θα πρέπει να συμπεριλαμβάνονται τουλάχιστον στην ανάλυση ευαισθησίας**

(White et al, BMJ, 2011)

Αντιμετώπιση Ελλειπουσών τιμών Διεθνής Εμπειρία

Παρά τη συνεχή ανάπτυξη μεθόδων αντιμετώπισης ελλειπουσών τιμών αυτές συνήθως δεν αντιμετωπίζονται επαρκώς

Wood et al, Clinical Trials 2004

Ανασκόπηση 71 δημοσιευμένων εργασιών σε BMJ, JAMA, Lancet, NEJM

- 89% είχαν ελλείπουσες τιμές στην ανταπόκριση
- Σε 37 μελέτες με επαναλαμβανόμενες μετρήσεις ανταπόκρισης, έγινε ανάλυση ατόμων με πλήρη δεδομένα (complete case analysis)
- Μόνο 21% διεξήγαγε ανάλυση ευαισθησίας (sensitivity analysis)

Chan et al, Lancet, 2005

- Το 65% των μελετών σε PubMed περιοδικά δεν αναφέρουν πως χειρίζονται τις ελλείπουσες τιμές

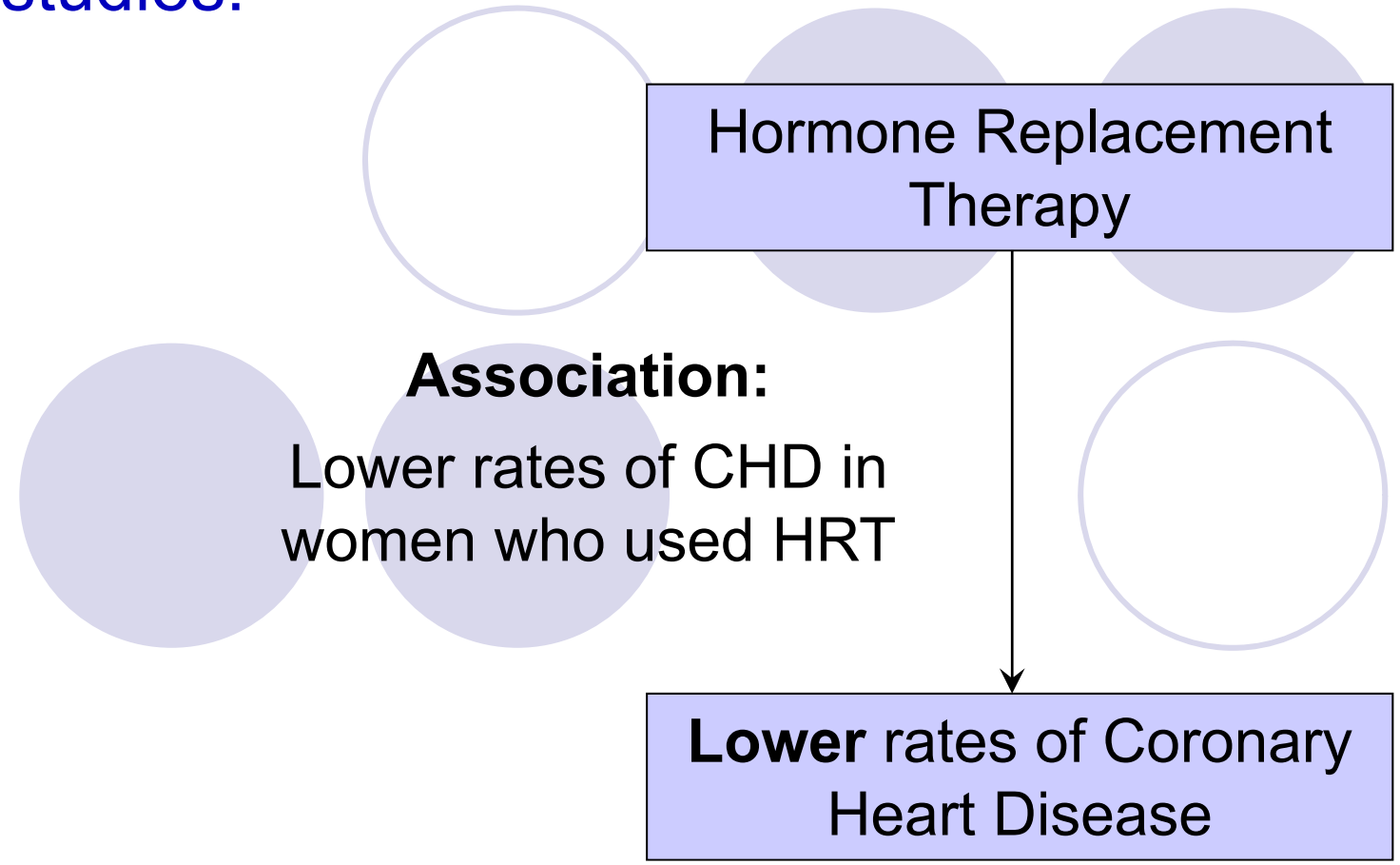
- CONSORT κατευθυντήριες οδηγίες: Αναφορά των ελλειπουσών τιμών ανά ομάδα (θεραπείας ή έκθεσης)

Ερμηνεία Αποτελεσμάτων

- Σφάλματα
- Συγχυτικοί Παράγοντες
- Γενικευσιμότητα αποτελεσμάτων

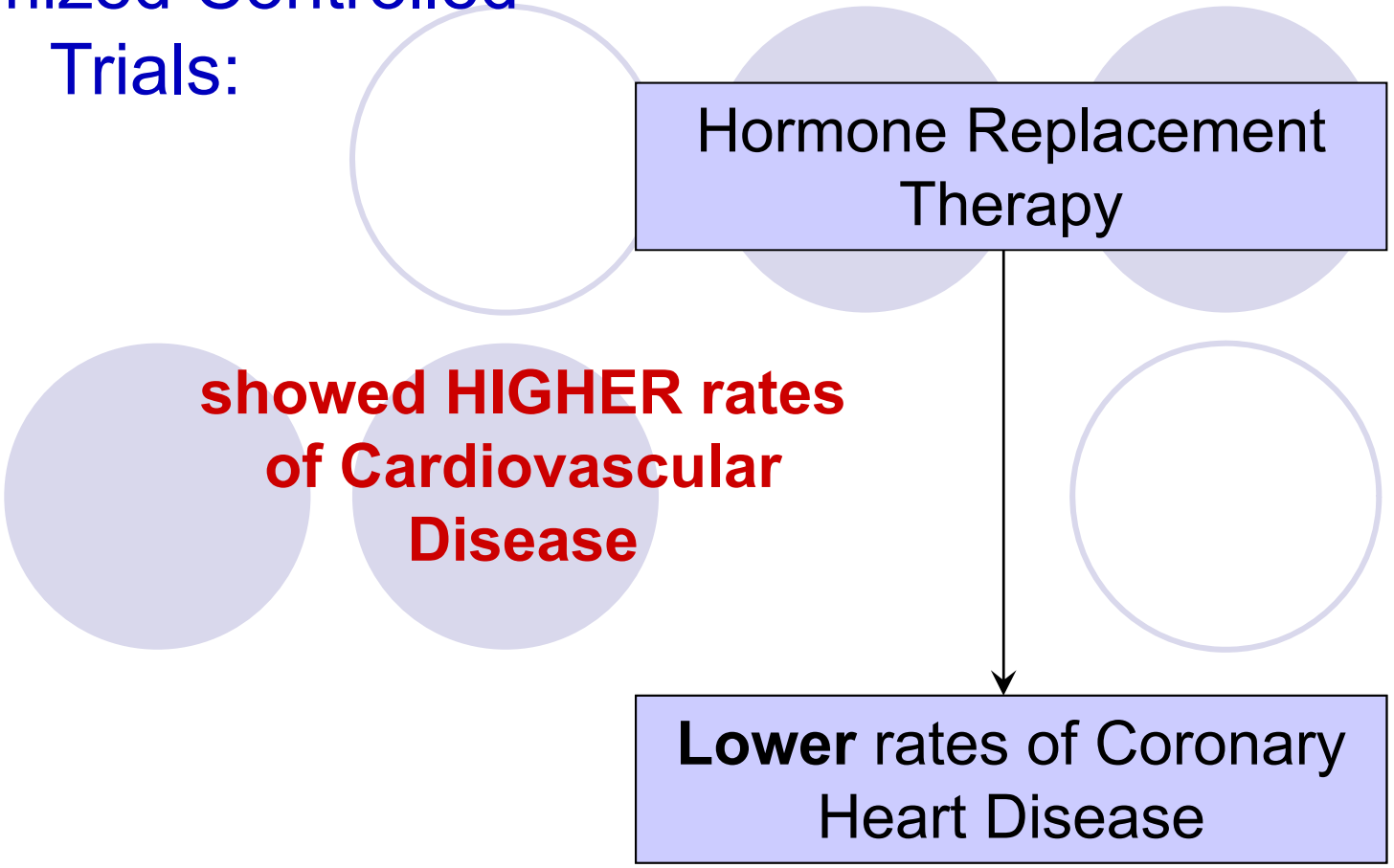
Channeling effect

Cohort studies:



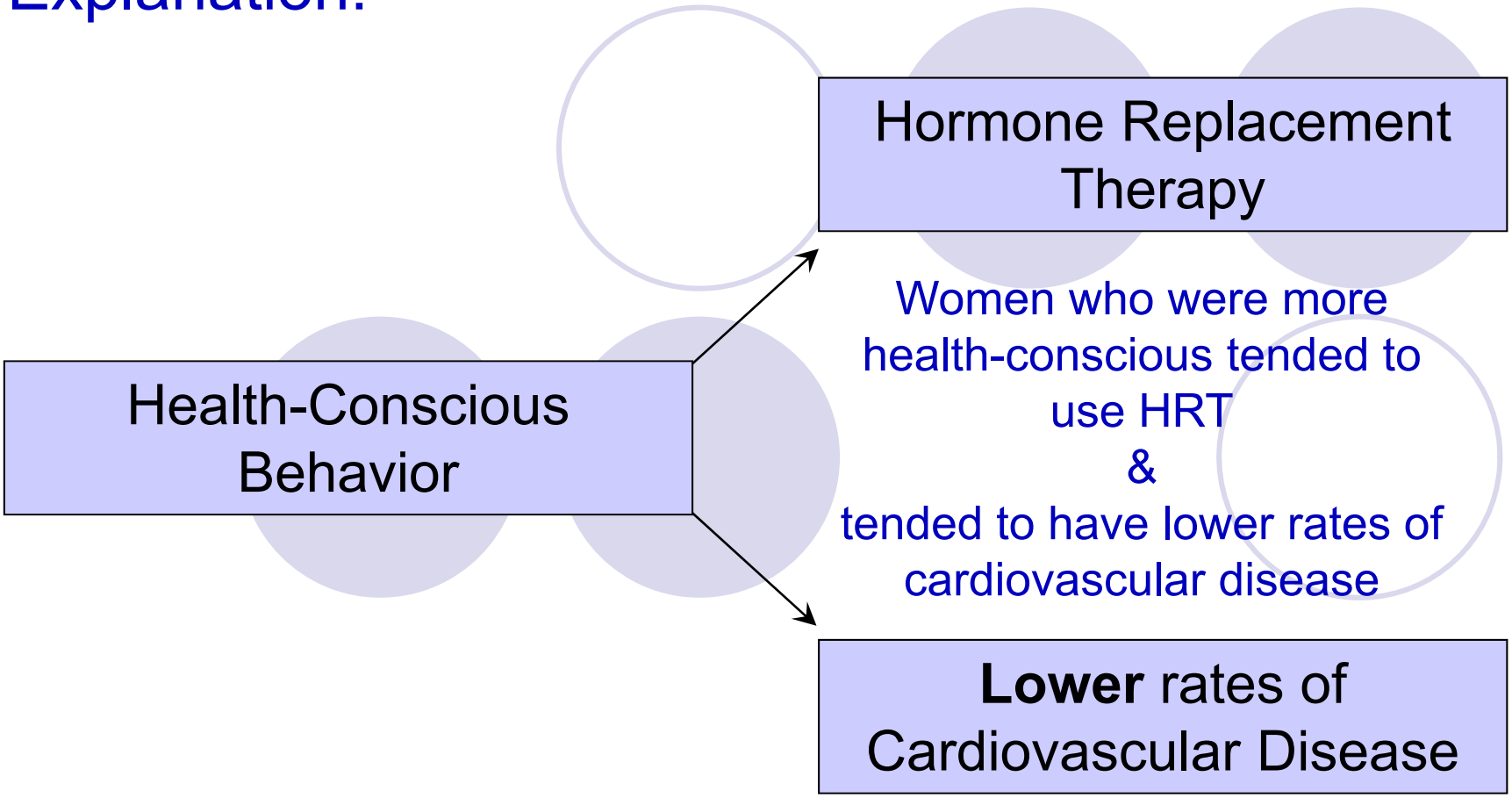
Channeling effect

Randomized Controlled Trials:



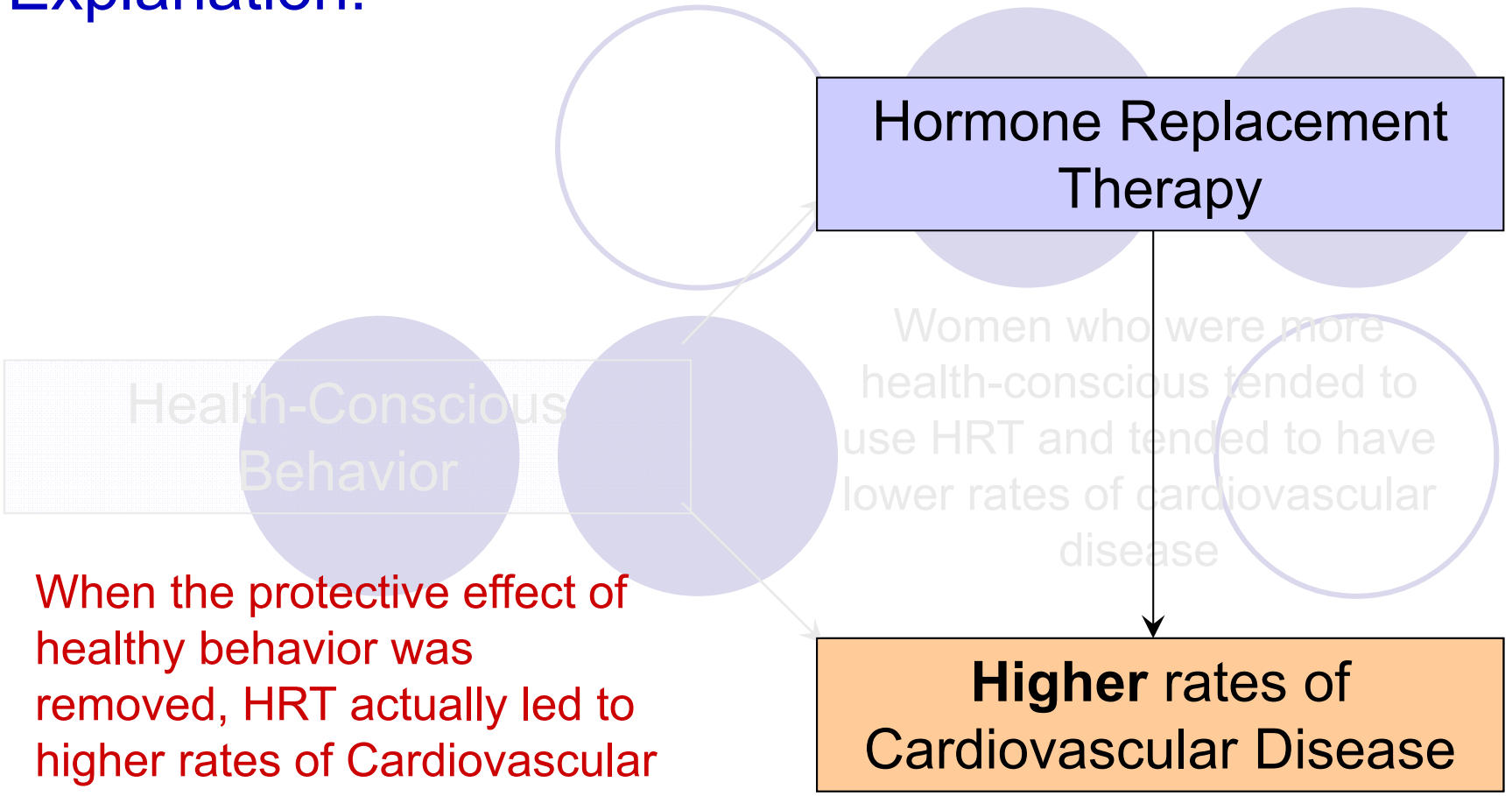
Channeling effect

Explanation:

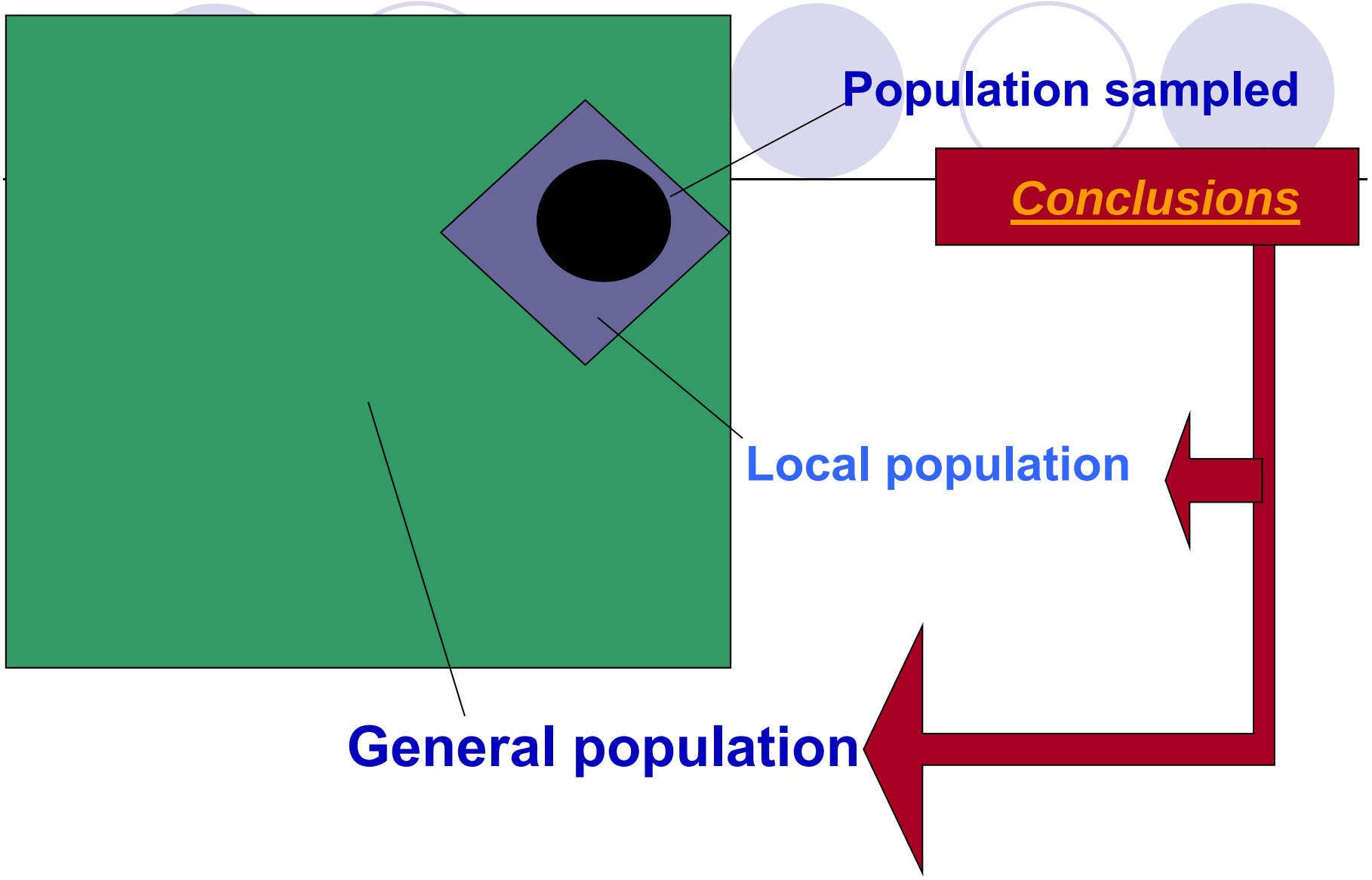


Channeling effect

Explanation:



When the protective effect of healthy behavior was removed, HRT actually led to higher rates of Cardiovascular disease.



External Validity (Generalizability)

Study quality in medicine

- Schulz (1995) evaluated study quality in 250 randomized clinical trials (RCTs) from 33 meta-analyses. Poor quality studies led to positively biased estimates:
  - lack of concealment (30-41%),
  - lack of double-blind (17%),
  - participants excluded after randomization (NS).
- Moher et al. (1998) reanalysed 127 RCTs randomized clinical trials from 11 meta-analyses for study quality.
  - Low quality trials resulted in significantly larger effect sizes, 30-50% exaggeration in estimates of treatment efficacy.
- Wood et al. (2008) evaluated study quality (1,346 RCTs from 146 meta-analyses.
  - subjective outcomes: inadequate/unclear concealment & lack of blinding resulted in substantial biases.
  - objective outcomes: no significant effects.
  - conclusion: systematic reviewers should assess risk of bias.

Διασφάλιση της ορθότητας των δημοσιευόμενων ερευνητικών αποτελεσμάτων

- Τα αποτελέσματα των ιατρικών ερευνών χρησιμοποιούνται στην πρόληψη και στην ιατρική πράξη, με διακύβευμα την υγεία και τη ζωή των ανθρώπων και επομένως είναι απαραίτητο να διασφαλιστεί όσο είναι δυνατόν η ποιότητα τους. Εξάλλου είναι δυνατόν εκ των υστέρων να αξιολογηθεί η ορθότητα τους.

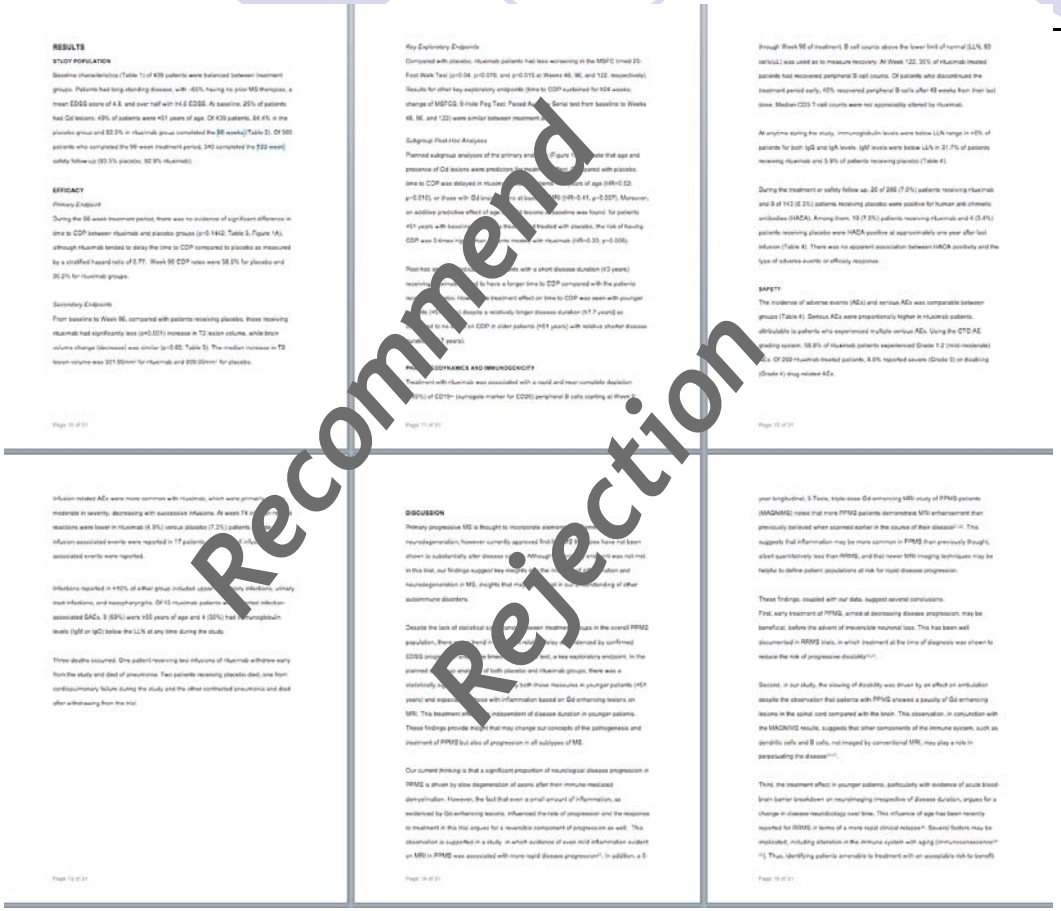
- Η προσπάθεια βελτίωσης της ποιότητας οδήγησε σε εξειδικευμένη και εκτεταμένη βιβλιογραφία στην οποία μπορεί κανείς να αναφερθεί.

Επιστημονικές δημοσιεύσεις: Η πραγματικότητα

- Πολλές μελέτες από αυτές που απορρίπτονται αξίζουν καλύτερης παρουσίασης…

- Εάν μία μελέτη αξίζει συγγραφής, καλό θα ήταν η συγγραφή να γίνει καλά!

Επιστημονικές δημοσιεύσεις: Η πραγματικότητα



«Χαρισματικός» και «έξυπνος» τρόπος γραφής δεν μπορεί να «**διορθώσει**» για κακό σχεδιασμό, πραγμάτωση, και ανάλυση των δεδομένων μίας επιστημονικής μελέτης…

Αλλά μία κακή/ελλιπής συγγραφή μπορεί να κάνει μία καλή μελέτη να α**πορριφθεί**!

Συγγραφή επιστημονικής δημοσίευσης

- Σαφής και ξεκάθαρη παρουσίαση της μελέτης,
- Στον editor του περιοδικού καταρχήν…
- Και στους αναγνώστες…
- Οι συγγραφείς θα πρέπει να θυμούνται σε ποιους (ή μάλλον σε πόσους) απευθύνονται…

Ένα σύνηθες «πρόβλημα»

Η μελέτη

- Καλός σχεδιασμός
- Πρωτοτυπία
- Επιτυχής πραγμάτωση
- Έγκυρη ανάλυση

Η προς υποβολή εργασία

- Όχι τόσο προσεκτικός σχεδιασμός, αδυναμίες στην συγγραφή, απορρίψεις/καθυστερήσεις στη δημοσίευση
- Αναστρέψιμη διαδικασία

Γιατί οι μελέτες αποτυγχάνουν να δημοσιευθούν;

57 articles evaluated to *Emergency Medicine*—28 accepted, 29 rejected
Of these 29:

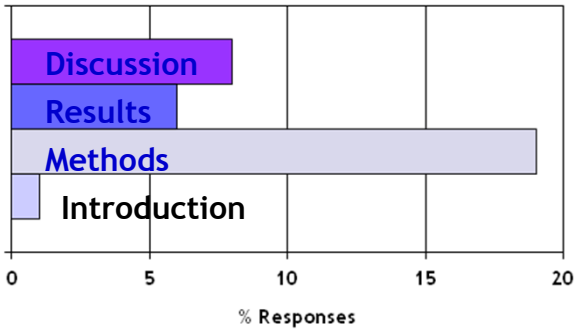
Ambiguous methods	77%
Ambiguous results	70%
Conclusions not warranted by data	72%
Poor referencing	56%
Inadequate study design description	51%
Unclear tables	49%
Overly long discussion	49%
Inadequate definition of terms	49%

"Deficiencies in manuscript preparation are more frequent than mistakes in study design and execution. Specific training…in manuscript preparation is recommended."

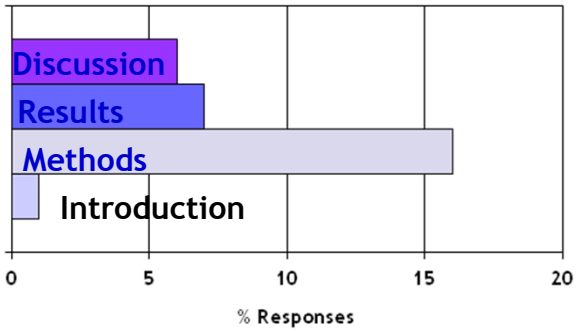
Taylor and Brown, Emergency Medicine 13(4):444-50, 2001

Editors' Responses

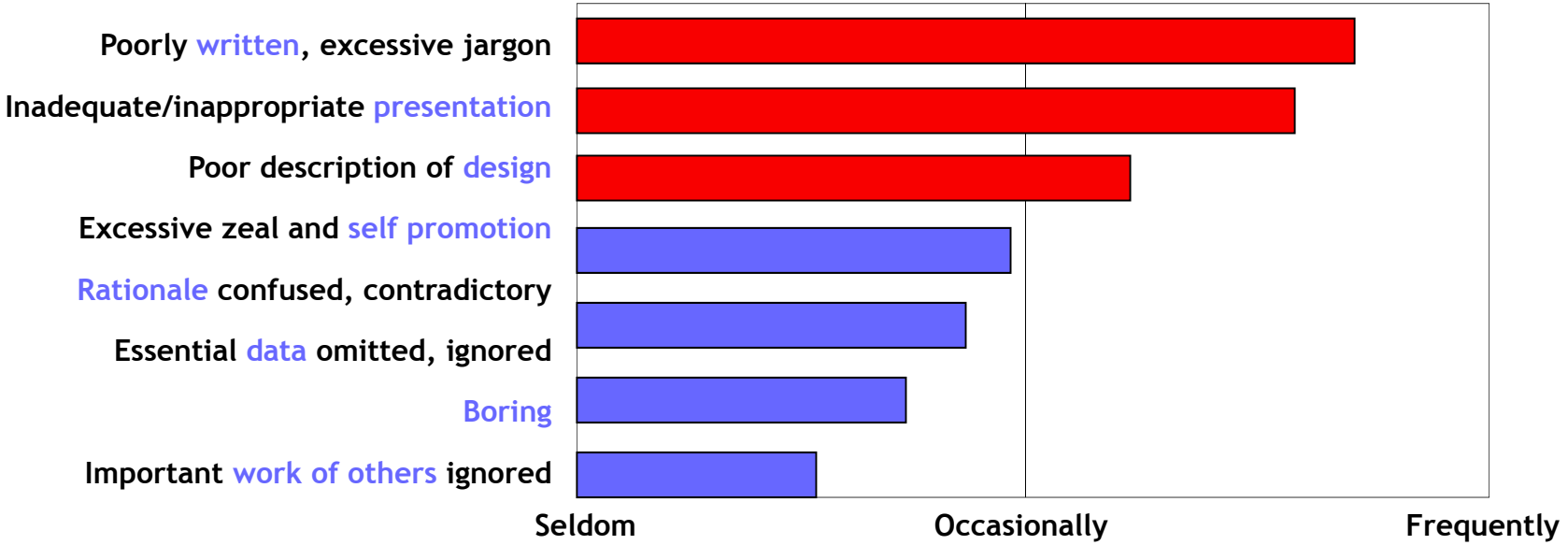
What section contains the most flaws?



What section responsible for outright rejection?



How frequently do Editors encounter manuscript problems?



Byrne DW, Publishing Medical Research Papers, Williams and Wilkins, 1998

Τα χαρακτηριστικά μιας καλής δημοσίευσης

Τίτλος: Περιεκτικός και σαφής

Περίληψη: Σαφής, περιεκτική και σύμφωνα με τα κριτήρια του περιοδικού (δομή/ αριθμός λέξεων)

Εισαγωγή: Πλαίσιο - σύντομο και περιεκτικό

Ερευνητική υπόθεση: Ξεκάθαρη και απλή διατύπωση

Βιβλιογραφία: Σχετική και περιεκτική

Μέθοδοι/Υλικό: Λεπτομερής περιγραφή των μεθόδων που ακολουθήθηκαν καθώς και των στατιστικών μεθόδων που ακολουθήθηκαν

Πίνακες/γραφήματα: Αυτόνομοι, να υποστηρίζουν τα συμπεράσματα, καλή εμφάνιση

Αποτελέσματα: **Ξεκάθαρη παρουσίαση αποτελεσμάτων – σαφής και περιεκτική**

Συζήτηση: Πώς τα ευρήματα συνδέονται (απαντούν και με ποιόν τρόπο) με την αρχική υπόθεση

Τρόπος γραφής: Ξεκάθαρος και λογικός

Κείμενο, Πίνακες, Βιβλιογραφία: Σύμφωνα με τους κανόνες του περιοδικού!

Reporting guidelines

1996	CONSORT	RCTs
1999	QUOROM	Meta-analyses of RCTs *new: PRISMA*
2001	CONSORT II	RCTs
2003	STARD	Diagnostic studies
2005	REMARK	Tumour marker studies
2007	STROBE	Case-control studies Cross-sectional studies Cohort studies

The Revised CONSORT Statement for Reporting Randomized Trials: Explanation and Elaboration

Douglas G. Altman, DSc; Kenneth F. Schulz, PhD; David Moher, MSc; Matthias Egger, MD; Frank Davidoff, MD; Diana Elbourne, PhD; Peter C. Gøtzsche, MD; and Thomas Lang, MA, for the CONSORT Group

17 April 2001 Annals of Internal Medicine Volume 134 Number 8

CONSORT (Consolidated Standars of Reporting)

- "Intent is to make experimental process more clear, flawed or not, so that users of the data can more appropriately evaluate its validity for their purposes"
  - **checklist**
  - **figure**
  - **available at www.consort-statement.org**
- **Χρησιμοποιείται από αρκετά περιοδικά υψηλού συντελεστή απήχησης**
- **Διαθέσιμη λίστα σε 6 γλώσσες**

CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	
	2b	Specific objectives or hypotheses	
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	
Participants	4a	Eligibility criteria for participants	
	4b	Settings and locations where the data were collected	
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	
	6b	Any changes to trial outcomes after the trial commenced, with reasons	
Sample size	7a	How sample size was determined	
	7b	When applicable, explanation of any interim analyses and stopping guidelines	
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those	

		assessing outcomes) and how	
	11b	If relevant, description of the similarity of interventions	
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	
	13b	For each group, losses and exclusions after randomisation, together with reasons	
Recruitment	14a	Dates defining the periods of recruitment and follow-up	
	14b	Why the trial ended or was stopped	
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	
Other information			
Registration	23	Registration number and name of trial registry	
Protocol	24	Where the full trial protocol can be accessed, if available	
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.

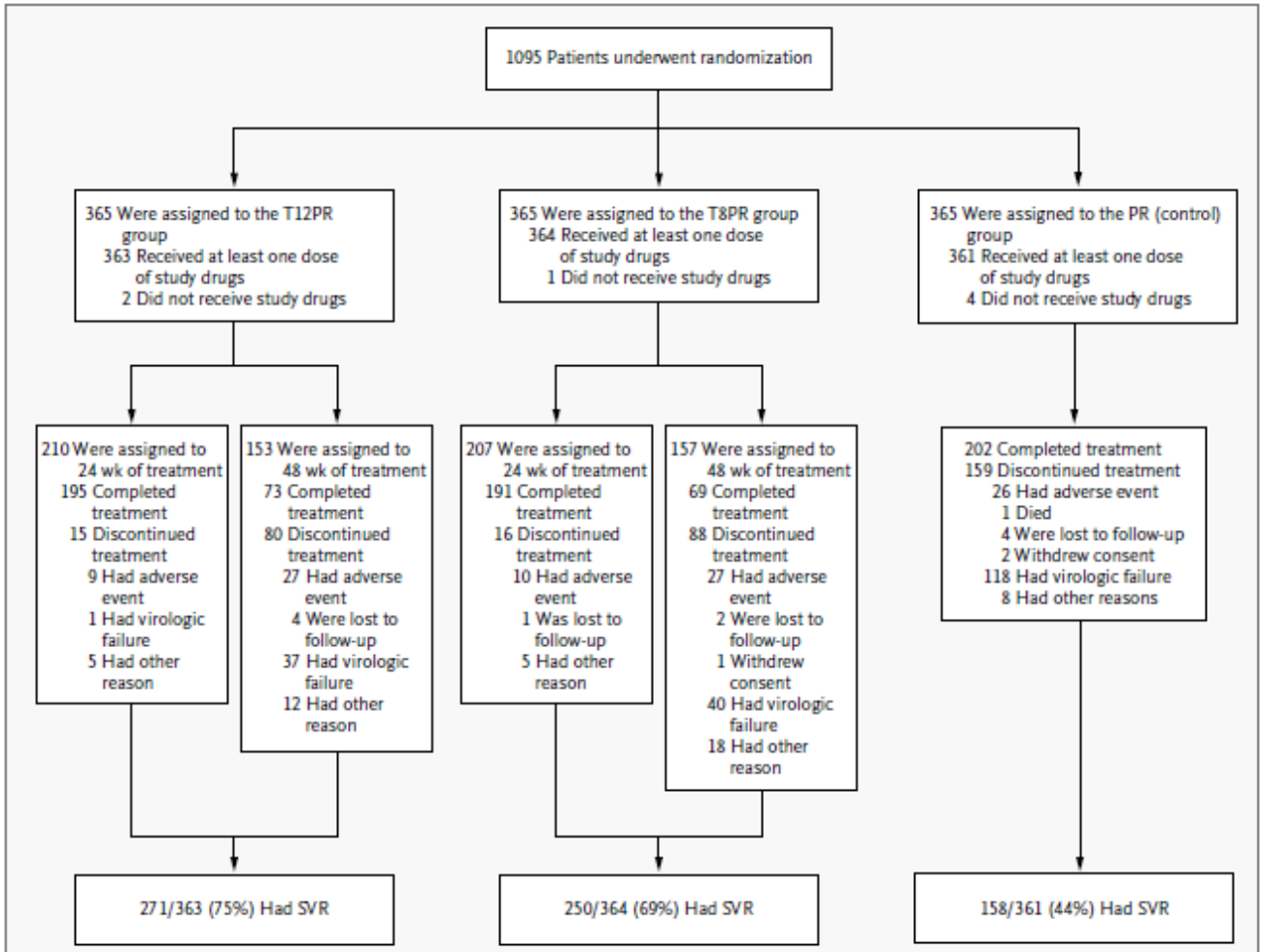


Figure 1. Randomization and Sustained Virologic Response (SVR) in Study Patients.

The T12PR group was assigned to receive telaprevir combined with peginterferon alfa-2a and ribavirin for 12 weeks, followed by peginterferon–ribavirin alone for 12 weeks if hepatitis C virus (HCV) RNA was undetectable at weeks 4 and 12 or for 36 weeks if HCV RNA was detectable at either time point; the T8PR group was assigned to receive telaprevir with peginterferon–ribavirin for 8 weeks and telaprevir-matched placebo plus peginterferon–ribavirin for 4 weeks, followed by peginterferon–ribavirin alone for 12 weeks if HCV RNA was undetectable at weeks 4 and 12 or by peginterferon–ribavirin for 36 weeks if HCV RNA was detectable at either time point; the PR (control) group was assigned to receive telaprevir-matched placebo plus peginterferon–ribavirin for 12 weeks, followed by peginterferon–ribavirin alone for 36 weeks. Of the 1095 patients who underwent randomization, 1088 received at least one dose of the study drugs and 7 did not receive any study drugs. In the T12PR group, 3 patients had an extended rapid virologic response but were assigned to the 48-week treatment group, and 1 patient who did not have an extended rapid virologic response was assigned to receive 24 weeks of treatment; however, this patient also met the week-12 stopping rule and discontinued treatment after the week-12 visit.

Jacobson et al, NEJM, 2011; 364:2405-16

OPEN ACCESS Freely available online

PLOS MEDICINE

Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration

Jan P. Vandenbroucke[1], Erik von Elm[2,3], Douglas G. Altman[4], Peter C. Gøtzsche[5], Cynthia D. Mulrow[6], Stuart J. Pocock[7], Charles Poole[8], James J. Schlesselman[9], Matthias Egger[2,10*] for the STROBE Initiative

1 Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands, **2** Institute of Social & Preventive Medicine (ISPM), University of Bern, Bern, Switzerland, **3** Department of Medical Biometry and Medical Informatics, University Medical Centre, Freiburg, Germany, **4** Cancer Research UK/NHS Centre for Statistics in Medicine, Oxford, United Kingdom, **5** Nordic Cochrane Centre, Rigshospitalet, Copenhagen, Denmark, **6** University of Texas Health Science Center, San Antonio, United States of America, **7** Medical Statistics Unit, London School of Hygiene and Tropical Medicine, London, United Kingdom, **8** Department of Epidemiology, University of North Carolina School of Public Health, Chapel Hill, United States of America, **9** Department of Biostatistics, University of Pittsburgh Graduate School of Public Health, and University of Pittsburgh Cancer Institute, Pittsburgh, United States of America, **10** Department of Social Medicine, University of Bristol, Bristol, United Kingdom

Funding: The initial STROBE workshop was funded by the European Science Foundation (ESF). Additional funding was received from the Medical Research Council Health Services Research Collaboration and the National Health Services Research & Development Methodology Programme. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Citation: Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, et al. (2007) Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration. PLoS Med 4(10): e297. doi:10.1371/journal.pmed.0040297

Received: July 20, 2007
Accepted: August 30, 2007
Published: October 16, 2007

ABSTRACT

Much medical research is observational. The reporting of observational studies is often of insufficient quality. Poor reporting hampers the assessment of the strengths and weaknesses of a study and the generalisability of its results. Taking into account empirical evidence and theoretical considerations, a group of methodologists, researchers, and editors developed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) recommendations to improve the quality of reporting of observational studies. The STROBE Statement consists of a checklist of 22 items, which relate to the title, abstract, introduction, methods, results and discussion sections of articles. Eighteen items are common to cohort studies, case-control studies and cross-sectional studies and four are specific to each of the three study designs. The STROBE Statement provides guidance to authors about how to improve the reporting of observational studies and facilitates critical appraisal and interpretation of studies by reviewers, journal editors and readers. This explanatory and elaboration document is intended to enhance the use, understanding, and dissemination of the STROBE Statement. The meaning and rationale for each checklist item are presented. For each item, one or several published examples and, where possible, references to relevant empirical studies and methodological literature are provided. Examples of useful flow diagrams are also included. The STROBE Statement, this document, and the associated Web site (http://www.strobe-statement.org/) should be helpful resources to improve reporting of observational research.

Table 1. The STROBE Statement—Checklist of Items That Should Be Addressed in Reports of Observational Studies

	Item number	Recommendation
TITLE and ABSTRACT	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found
INTRODUCTION		
Background/ rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any prespecified hypotheses
METHODS		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	6	(a) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants (b) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses
RESULTS		
Participants	13*	(a) Report the numbers of individuals at each stage of the study—e.g., numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed (b) Give reasons for non-participation at each stage (c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) *Cohort study*—Summarise follow-up time (e.g., average and total amount)
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time *Case-control study*—Report numbers in each exposure category, or summary measures of exposure *Cross-sectional study*—Report numbers of outcome events or summary measures
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses
DISCUSSION		
Key results	18	Summarise key results with reference to study objectives
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	21	Discuss the generalisability (external validity) of the study results

Οδηγίες συγγραφής επιστημονικής δημοσίευσης σε κάθε επιστημονικό περιοδικό

1. Instructions to Authors

2. Περιέχει την διαμόρφωση που πρέπει να ακολουθεί η εργασία υπό-υποβολή προκειμένου να γίνει αποδεκτή ως υποβαλλόμενη!!!!

3. Περιλαμβάνει όλες τις ενότητας μίας εργασίας

4. Οι συγγραφείς υποχρεούνται να διαμορφώσουν την εργασία τους σύμφωνα με τους κανόνες του εκάστοτε περιοδικού

Πίνακες/Γραφήματα
Αυτόνομοι, Επαρκείς, Λιτοί

Πίνακες και Γραφήματα

- Κρίσιμοι—για τους συντάκτες και τους αναγνώστες – συνήθως αποτελούν την πρώτη ανάγνωση από την δημοσίευση
- Οι συντάκτες «κρίνουν» καταρχήν την υποβαλλόμενη δημοσίευση από το πόσο καλά οι πίνακες έχουν διαμορφωθεί όσον αφορά περιεχόμενο και εμφάνιση
- Αυτόνομοι: παρουσίαση της δημοσίευσης χωρίς να χρειάζεται να διαβαστεί η δημοσίευση!
- Σαφής παρουσίαση των αποτελεσμάτων με μια ματιά!

Πίνακας ή κείμενο;

Table 2. Response during and after the Treatment Period, According to Treatment Group.

Response	T12PR (N=363)	T8PR (N=364)	PR (N=361)
Undetectable HCV RNA during treatment period — no. (%)*			
At week 4	246 (68)	242 (66)	34 (9)
At weeks 4 and 12	212 (58)	207 (57)	29 (8)
Undetectable HCV RNA at end of treatment period — no. (%)	314 (87)	295 (81)	229 (63)
Undetectable HCV RNA 24 wk after end of treatment: sustained virologic response — no./total no. (%)†			
All patients‡	271/363 (75)	250/364(69)	158/361 (44)
Patients with undetectable HCV RNA at weeks 4 and 12	189/212 (89)	171/207 (83)	28/29 (97)
Patients with detectable HCV RNA at weeks 4 or week 12	82/151 (54)	79/157 (50)	130/332 (39)
Patients with undetectable HCV RNA at week 4	206/246 (84)	188/242 (78)	32/34 (94)
Patients with detectable HCV RNA at week 4	65/117 (56)	62/122 (51)	126/327 (39)
Undetectable HCV RNA at 72 wk — no. (%)§	265 (73)	243 (67)	158 (44)
Relapse among patients with undetectable HCV RNA at end of treatment period — no./total no. (%)			
All patients	27/314 (9)	28/295 (9)	64/229 (28)
Patients who completed treatment	17/264 (6)	18/247 (7)	51/189 (27)

* Patients with undetectable HCV RNA at week 4 met the criterion for a rapid virologic response, and patients with undetectable HCV RNA at weeks 4 and 12 met the criterion for an extended rapid virologic response.
† Sustained virologic response (undetectable HCV RNA 24 weeks after the end of treatment) was the primary end point.
‡ All patients who received at least one dose of study drug were included in the analysis. The difference in response rates was 31 percentage points (95% confidence interval [CI], 24 to 38) between the T12PR and PR groups and 25 percentage points (95% CI, 18 to 32) between the T8PR and PR groups.
§ The 72-week assessment was performed 24 weeks after the end of treatment in patients who received 48 weeks of treatment and 48 weeks after end of treatment in patients who received 24 weeks of treatment.

Jacobson et al, NEJM 2011; 364:2405-16

Αποφυγή επανάληψης των δεδομένων των πινάκων - γραφημάτων

Table 2.—Selected Pregnancy Outcomes and Neonatal Measurements in the Zinc Supplement and Placebo Subgroups by Body Mass Index (BMI) Categories

	BMI ≥26			BMI <26		
	Zinc Supplement (n=155)	Placebo (n=145)	*P*	Zinc Supplement (n=134)	Placebo (n=134)	*P*
Maternal Characteristics						
Age, y	24.8	24.2	.32	22.9	21.2	.01
BMI, kg/m²	33.4	33.0	.64	22.3	22.2	.57
Current smoker, %	7.7	5.5	.44	3.0	3.0	.98
Pregnancy Outcome						
Birth weight, g	3240	3241	.99	3190	2942	.005
Gestational age, wk	39.0	38.7	.47	38.6	37.9	.08
Preterm birth <32 wk, %	3.2	5.5	.33	3.0	6.8	.15
Birth weight <1500 g, %	3.9	3.5	.84	2.3	6.0	.12
Anthropometric Measurements						
Crown-heel length, cm	50.2	49.8	.41	50.3	49.7	.20
Head circumference, cm	34.3	34.0	.50	34.1	33.4	.005
Abdominal circumference, cm	33.3	33.1	.64	32.8	32.6	.58
Arm length, cm	9.9	9.7	.27	9.9	9.6	.03
Subscapular skinfold, mm	4.2	3.9	.05	3.9	3.6	.06
Neonatal Outcome						
Neonatal hospital stay, d	3.9	4.5	.47	3.1	4.9	.10
Neonatal sepsis, %	0.7	1.4	.52	0	2.2	.08

As shown In Table 2 the mean age of women in the Zinc group with BMI ≥ 26 was 24.8 and for women in the placebo group with BMI ≥26 kg/m^2 mean age was 24.2 and the difference was not statistically significant (p=0.32)...

ΚΕΙΜΕΝΟ = ΠΙΝΑΚΑΣ

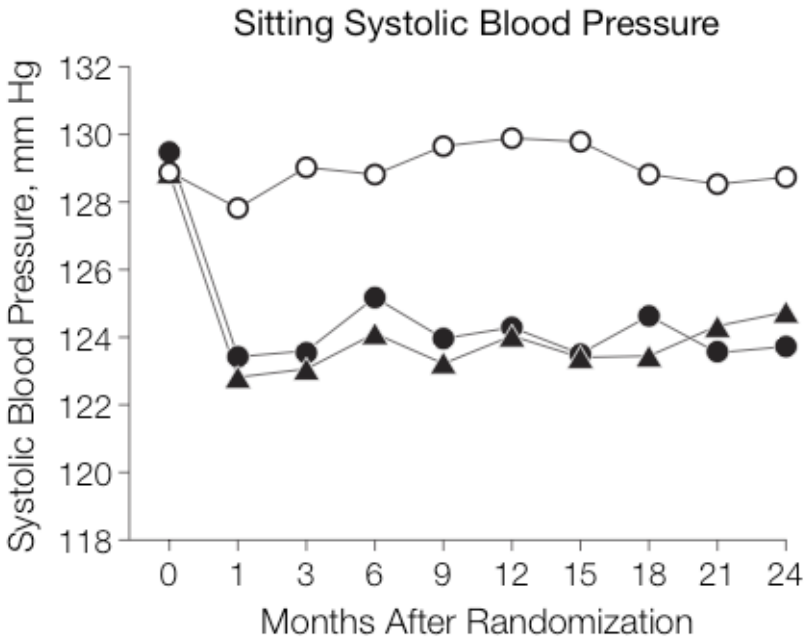
Goldenberg et al, JAMA, 1995

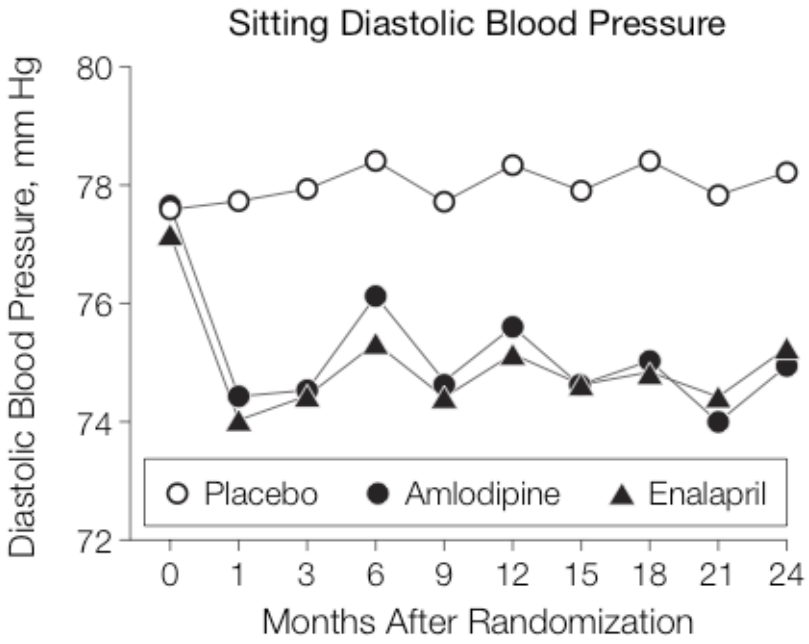
ΓΡΑΦΗΜΑΤΑ

- Εύκολη αποτύπωση του μηνύματος
- "Μία εικόνα αξίζει όσο χίλιες λέξεις"
- Στόχος δεν είναι να «ομορφύνουν» το κείμενο αλλά να μεταφέρουν την πληροφορία με άμεσο τρόπο
- Όσον αφορά το περιεχόμενο ισχύει ό,τι και στους πίνακες

Γραφήματα: Απλά

Figure 2. Mean Patient Blood Pressure at Baseline and During Treatment

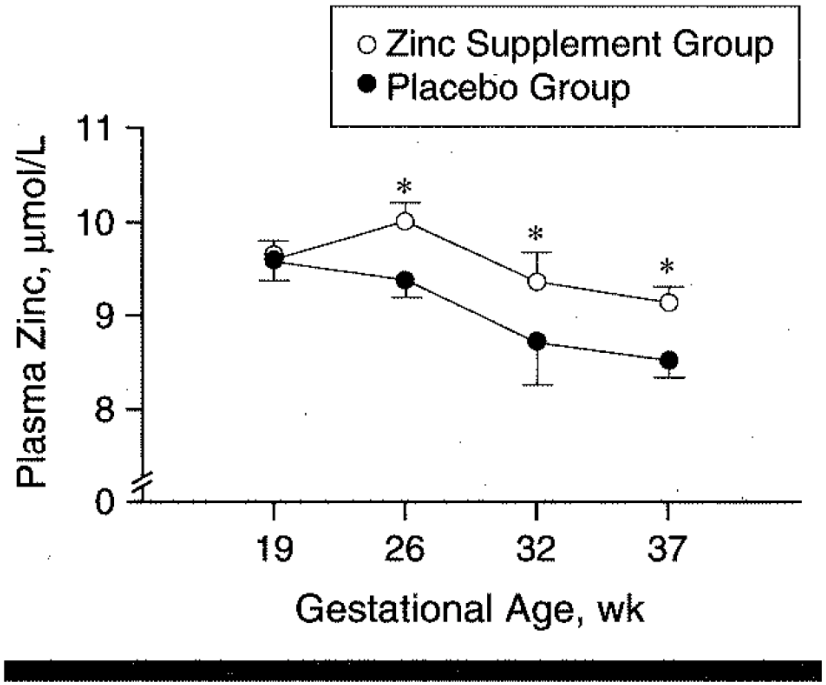




- >3-4 ομάδες – Πολύπλοκα
- Μικρός αριθμός συμβόλων
- Τυπική απόκλιση, P-values εάν χρειάζονται

Charles G Durbin Jr MD FAARC, Respiratory Care, 2004

Αποφυγή διατύπωσης του προφανούς

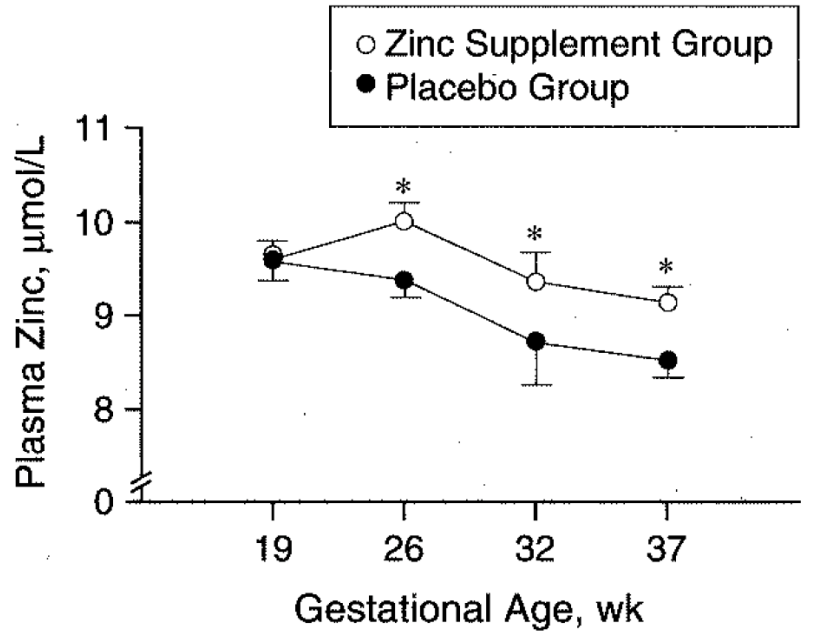


Changes in plasma zinc concentrations. Asterisk indicates significant difference between the values of the zinc supplement and placebo groups ($P \leq .05$). Vertical bars indicate SEMs.

Figure 1 is a line graph illustrating the plasma zinc levels (µmoL/L) over the 37 weeks versus gestational age in both the zinc supplement group and placebo group. The placebo and the zinc group both decreased over the 37 weeks of the study, but the differences were significant for the zinc group.

Goldenberg et al, JAMA, 1995

Αλλά διατύπωση του σημαντικού!



Changes in plasma zinc concentrations. Asterisk indicates significant difference between the values of the zinc supplement and placebo groups ($P \leq .05$). Vertical bars indicate SEMs.

Goldenberg et al, JAMA, 1995

We measured mothers' plasma zinc levels before randomization (week 19) and at 26, 32, and 37 weeks' gestational age (Fig 1).

Beginning as early as 26 weeks and at each timepoint, differences in plasma zinc levels between placebo and zinc supplement groups were statistically significant after randomization.

Γραφήματα: Κάποια κοινά λάθη

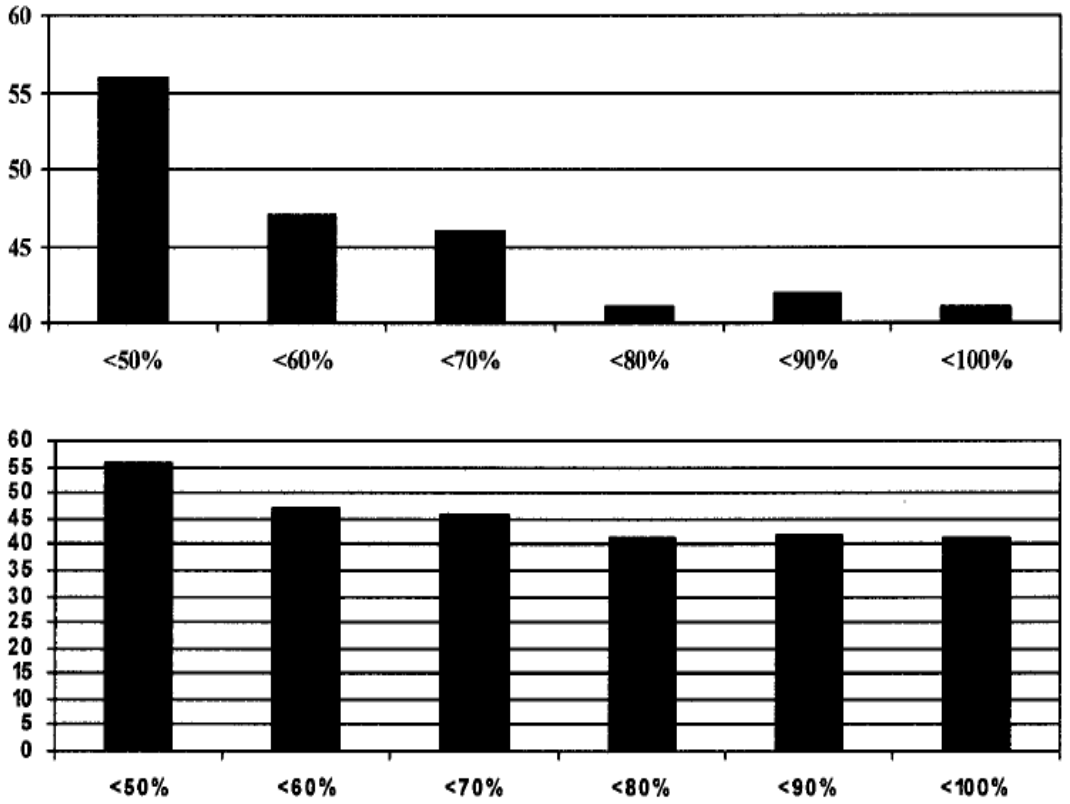


Fig. 7. Improper axis starting points and/or scale can give a misleading appearance that differences are statistically significant. The truncated ordinal axis (ie, zero–40 is missing) in the upper graph makes the differences appear large, whereas in the lower panel the ordinate values begin at zero and the graph is vertically compressed, which visually suggests that the differences are smaller.

Charles G Durbin Jr MD FAARC, Respiratory Care, 2004

ΣΥΖΗΤΗΣΗ

ΕΡΜΗΝΕΙΑ ΤΩΝ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

ΔΟΜΗ

1. Κύρια αποτελέσματα και τι αυτά σημαίνουν
2. Πώς συνδέονται με την βιβλιογραφία
3. Προτερήματα-Αδυναμίες της μελέτης – Καινούριες υποθέσεις που πιθανά αναδύονται
4. Μελλοντικές μελέτες
5. Συμπέρασμα – η τελευταία παράγραφος!!!

Αποφυγή πλατειασμών

- Συνηθισμένα λάθη:
  - Υπερεκτίμηση της σπουδαιότητας των ευρημάτων
  - Μεγαλόσχημα συμπεράσματα που δε υποστηρίζονται από τα ευρήματα
  - Υπερεκτίμηση στατιστικά σημαντικών αποτελεσμάτων χωρίς κριτική αντιμετώπιση

- Συνήθως οδηγούν σε Συζητήσεις μεγάλου μήκους (κουραστικές, επαναλαμβανόμενα ευρήματα) – Περιοδικά υψηλού συντελεστή απήχησης έχουν όριο για το μέγεθος του άρθρου – 3500 – 4000 λέξεις
- **Επιτυχία:** Τα δεδομένα μιλούν από μόνα τους για την μελέτη μας και όχι εμείς για αυτά

ΣΥΜΠΕΡΑΣΜΑ ΤΗΣ ΕΡΓΑΣΙΑΣ

- Διατυπώσεις οι οποίες γίνονται κατανοητές αφού διαβαστεί όλη η εργασία
- Ένα επιτυχημένο συμπέρασμα αναδεικνύει πλευρές για το συγκεκριμένο θέμα οι οποίες δεν έχουν διατυπωθεί στην Εισαγωγή
- Το συμπέρασμα έχει να κάνει με τις επιπτώσεις της συγκεκριμένης έρευνας στο πλαίσιο στο οποίο κινείται

Το «ιστορικό» των δημοσιεύσεων

- 1976: ~ 5,000 biomedical journals, only in libraries
- 2009: >25,000 biomedical journals, electronic on internet
- The number of subspecialties and new vocabularies has increased 5-10-fold over past 20 years

- Το αναγνωστικό κοινό χρειάζεται δημοσιεύσεις που να χαρακτηρίζονται από
  - Σαφήνεια
  - Απλότητα

"We must strive to make our articles increasingly 'reader friendly' and cross-discipline in language…"

The NEW ENGLAND JOURNAL of MEDICINE

Jerome Kassier, M.D.
Former Editor, NEJM

We need less research, better research, and research done for the right reasons……

DG Altman, BMJ 1994; 308:283