

# Το συστηματικό σφάλμα (bias) στην επιδημιολογική έρευνα

9 Ιουλίου 2015

Ομιλητής: Θοδωρής Λύτρας  
Κέντρο Ελέγχου και Πρόληψης Νοσημάτων (ΚΕΕΛΠΝΟ)

- 1 Εισαγωγή - Τυχαίο vs συστηματικό σφάλμα
- 2 Τα συστηματικά σφάλματα στις επιδημιολογικές μελέτες
- 3 Συγχυτικό σφάλμα (confounding)
- 4 Συστηματικά σφάλματα στις τυχαιοποιημένες μελέτες
- 5 Επίλογος

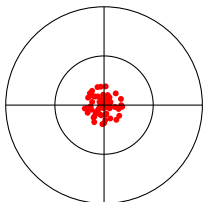
- 1 Εισαγωγή - Τυχαίο vs συστηματικό σφάλμα
- 2 Τα συστηματικά σφάλματα στις επιδημιολογικές μελέτες
- 3 Συγχυτικό σφάλμα (confounding)
- 4 Συστηματικά σφάλματα στις τυχαιοποιημένες μελέτες
- 5 Επίλογος

“...epidemiologic research is an exercise in measurement [...] to obtain an estimate that is generalizable to relevant target populations.”

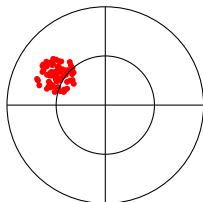
— Rothman KJ, et al. (2008) *Modern Epidemiology*. 3rd ed.

# Παράδειγμα 1

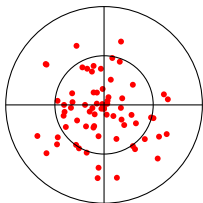
Precision (ακρίβεια) vs Accuracy (πιστότητα)



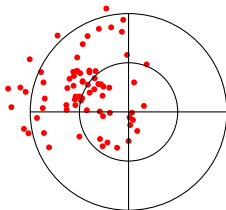
↓ τυχαίο, ↓ συστηματικό σφάλμα  
(↑ ακρίβεια, ↑ πιστότητα)



↓ τυχαίο, ↑ συστηματικό σφάλμα  
(↑ ακρίβεια, ↓ πιστότητα)



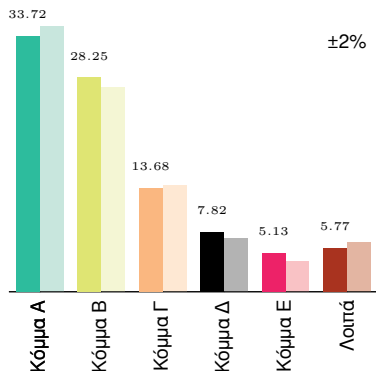
↑ τυχαίο, ↓ συστηματικό σφάλμα  
(↓ ακρίβεια, ↑ πιστότητα)



↑ τυχαίο, ↑ συστηματικό σφάλμα  
(↓ ακρίβεια, ↓ πιστότητα)

Στην επιδημιολογία, το τυχαίο σφάλμα προκύπτει από το γεγονός ότι κάνουμε **δειγματοληψία**.

# Παράδειγμα 2



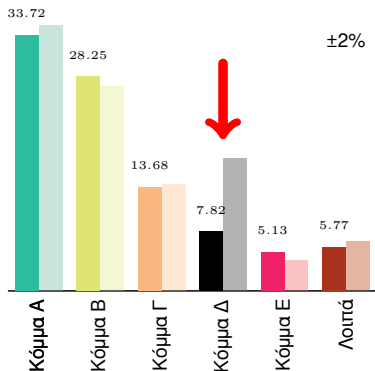
Δείγμα → Πληθυσμός (άγνωστος)

±2%

Το τυχαίο σφάλμα:

- Υπάρχει *πάντα*, όταν γίνεται δειγματοληψία
- **Μειώνεται όσο αυξάνεται το δείγμα της μελέτης**
  - Μπορεί να μηδενιστεί? **ΟΧΙ**
- Δεν “μεροληπτεί” υπέρ κάποιου ενδεχόμενου
  - Μπορεί όμως να διαστρεβλώσει την εικόνα, αν είναι πολύ μεγάλο.

# Παράδειγμα 2



Το συστηματικό σφάλμα:

- “Μεροληπτεί” υπέρ κάποιου ενδεχόμενου
  - Διαστρεβλώνει την εικόνα
- Δεν επηρεάζεται από το μέγεθος του δείγματος
- (Στην επιδημιολογία) Δεν μπορούμε να ελέγξουμε την παρουσία του
  - Μπορούμε μόνο να κάνουμε **ποιοτικές υποθέσεις** γι' αυτό



# Τυχαίο σφάλμα vs Συστηματικά σφάλματα

## Τυχαίο σφάλμα

- Οφείλεται στο γεγονός της δειγματοληψίας
- Μειώνεται όσο μεγαλύτερο δείγμα εξετάζουμε
- Μπορούμε να εκτιμήσουμε το μέγεθός του

## Συστηματικά σφάλματα

- Οφείλονται στον τρόπο που μετράμε, ή στον τρόπο που επιλέγουμε το δείγμα
- Δεν επηρεάζονται από το μέγεθος του δείγματος
- Δε μπορούμε να ελέγξουμε την ύπαρξή τους

## Παράδειγμα: μελέτη κοόρτης

		Διάρροια		Σύνολο
		Ναι	Όχι	
Παγωτό βανίλια	Έφαγε	43	11	54
	Δεν έφαγε	3	18	21

*Σχετικός κίνδυνος (Risk Ratio)*

# Παράδειγμα: μελέτη κοόρτης

		Διάρροια		
		Ναι	Όχι	Σύνολο
Παγωτό βανίλια	Έφαγε	43	11	54
	Δεν έφαγε	3	18	21

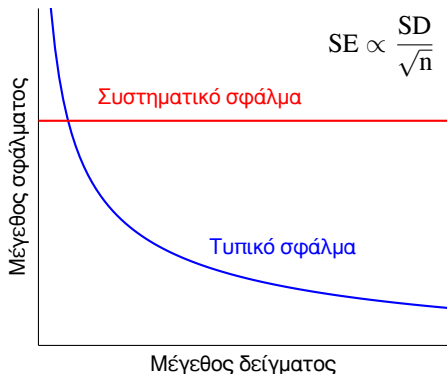
Σχετικός κίνδυνος (*Risk Ratio*)

Coefficients:

	Estimate	Std.err	Wald	Pr(> W )	
(Intercept)	-1.946	0.535	13.2	0.00027	***
<code>vanilla.ice.creamY</code>	1.718	0.539	10.2	0.00143	**

Σε κάθε μέτρο: σημειακή εκτίμηση (point estimate) + τυπικό σφάλμα (standard error)

# Υπολογισμός τυπικού σφάλματος



		Νόσος	
		(+)	(-)
Εκθεση	(+)	$\alpha$	$\beta$
	(-)	$\gamma$	$\delta$

$$SE(\log OR) = \sqrt{\frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma} + \frac{1}{\delta}}$$

$$SE(\log RR) = \sqrt{\frac{1}{\alpha} - \frac{1}{\alpha + \beta} + \frac{1}{\gamma} - \frac{1}{\gamma + \delta}}$$

# Εναλλακτικά

(τρόποι ποσοτικής έκφρασης του τυχαίου σφάλματος, με πιο “δαισθητικό” τρόπο)

## ■ 95% Διάστημα Εμπιστοσύνης (95% Confidence interval – 95% CI):

Ένα διάστημα το οποίο, **αν επαναλαμβάναμε** 100 φορές τη δειγματοληψία, **θα αναμέναμε** να περιέχει την πραγματική (πληθυσμιακή) τιμή 95 φορές.\*

\* Εφ'όσον δεν υπάρχουν συστηματικά σφάλματα!

- Οριοθετεί ένα εύρος μέσα στο οποίο είναι *αρκετά πιθανό* να βρίσκεται η πληθυσμιακή τιμή του μέτρου
- Ευθεία μαθηματική σχέση με το **τυπικό σφάλμα**:  $x \pm 1,96 \times SE$
- Μικρότερο SE → στενότερο 95% CI
- **Στατιστικά σημαντικό αποτέλεσμα** 95% CI δεν περιέχει τη μονάδα

## ■ Τιμή p (p-value):

Η πιθανότητα να δούμε ένα αποτέλεσμα το ίδιο ή περισσότερο ακραίο, με αυτό το μέγεθος δείγματος, εάν ισχύει η “μηδενική” υπόθεση (μη ύπαρξη διαφοράς)

- Εκφράζει το πόσο βέβαιοι είμαστε ότι η παρατηρούμενη διαφορά δεν οφείλεται στο τυχαίο σφάλμα
  - ΔΕΝ ισούται με την πιθανότητα να μην υπάρχει διαφορά
- Ευθεία μαθηματική σχέση με το **τυπικό σφάλμα**:  $\Phi \left( \frac{-|x|}{SE} \right)$
- Μικρότερο SE → μικρότερη τιμή p
- **Στατιστικά σημαντικό αποτέλεσμα**  $p < 0,05$

**TABLE**

Relative risk of gastrointestinal disease among questionnaire respondents (n=61), by food exposure at hotel dinner, outbreak of *Clostridium perfringens* infection, Trondheim, Norway, 2012

Food item	Exposed		Unexposed		Relative risk	95% CI
	Number of cases who ate food item/total	Attack rate (%)	Number of cases who did not eat food item/total	Attack rate (%)		
Beef stew	42/47	89	1/14	7	12.51	1.89–82.91
Rice	41/49	84	2/12	17	5.02	1.41–17.90
Green salad	28/40	70	15/21	71	0.98	0.70–1.37
Bread	30/46	65	13/15	87	0.75	0.56–1.01
Cured meat sausage	3/6	50	40/55	73	0.69	0.30–1.55
Ham	5/17	29	38/44	86	0.34	0.16–0.72
Cheese	6/21	29	37/40	93	0.31	0.16–0.61
Liver paste	1/5	20	42/56	75	0.27	0.05–1.55

Wahl E, Rømme S, Granum PE. A *Clostridium perfringens* outbreak traced to temperature-abused beef stew, Norway, 2012. Euro Surveill. 2013;18(9):pii=20408

**Table 1. West Nile virus IgG seroprevalence by gender, age-group, and area of residence**

		WNV IgG-positive		Prevalence ratio (95% CI)*	P-value
		n/N	% (95% CI)*		
<b>Gender</b>	Female	23/344	5.6 (3.5–8.7)	Reference	
	Male	18/300	6.0 (3.3–10.5)	1.07 (0.57–2.00)	0.835
<b>Age (years)</b>	18–39	4/161	2.7 (1.0–7.1)	Reference	
	40–59	14/238	5.7 (3.1–10.1)	2.10 (0.82–5.34)	0.120
	60–69	11/103	10.6 (6.1–17.9)	3.95 (1.33–11.69)	0.013
	70+	12/142	8.7 (4.8–15.3)	3.22 (1.14–9.09)	0.027
<b>Area of residence†</b>	Urban	3/201	1.4 (0.3–6.9)	Reference	
	Semi-urban	12/198	5.3 (2.8–9.8)	3.84 (0.68–21.68)	0.127
	Rural	26/245	10.0 (6.5–15.3)	7.28 (1.37–38.78)	0.020
<b>Total</b>		41/644	5.8 (3.8–8.6)		

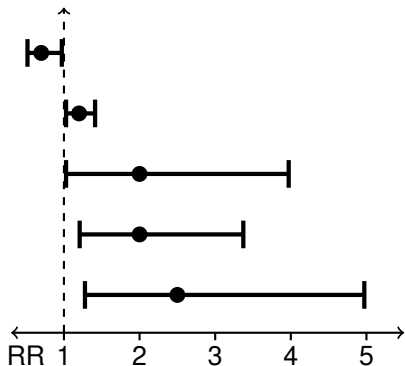
WNND: West Nile neuroinvasive disease

95% CI: 95% confidence interval

\*Proportions and confidence intervals are weighted by age and area of residence, and adjusted for cluster design.

†Area of residence: urban: >20,000 inhabitants; semi-urban: 2,000–20,000 inhabitants; rural: <2,000 inhabitants.

## Τιμή $p$ = απώλεια πληροφορίας



Σημειακή εκτίμηση	95% ΔΑ	τιμή $p$
RR = 0,7	0,49 – 1,0	0,05
RR = 1,2	1,0 – 1,44	0,05
RR = 2,0	1,0 – 4,0	0,05
RR = 2,0	1,18 – 3,4	0,01
RR = 2,5	1,25 – 5,0	0,01

Η τιμή  $p$  δε μας λέει την κατεύθυνση της συσχέτισης.

- Αν  $p \leq 0,05$  τότε το 95% ΔΑ **δεν περιέχει** τη μονάδα (στατιστικά σημαντική διαφορά)
- Αν  $p > 0,05$  τότε το 95% ΔΑ **περιέχει** τη μονάδα (μη στατιστικά σημαντική διαφορά)



# Συνοψίζοντας

(για το τυχαίο σφάλμα)

- Οφείλεται στο (αναπόφευκτο) γεγονός της **δειγματοληψίας**
- Μέτρα ποσοτικής έκφρασης του τυχαίου σφάλματος:
  - Τυπικό σφάλμα      **95% Διάστημα Εμπιστοσύνης**      τιμή  $p$
- Ελαττώνεται όσο αυξάνεται το μέγεθος δείγματος
  - Τα πάντα μπορούν να γίνουν “στατιστικά σημαντικά”, αν έχουμε επαρκές μέγεθος δείγματος
- **95% CI** και **p-values**  
ΔΕΝ είναι “μέτρα αξιοπιστίας” μιας μελέτης!

- 1 Εισαγωγή - Τυχαίο vs συστηματικό σφάλμα
- 2 Τα συστηματικά σφάλματα στις επιδημιολογικές μελέτες
- 3 Συγχυτικό σφάλμα (confounding)
- 4 Συστηματικά σφάλματα στις τυχαιοποιημένες μελέτες
- 5 Επίλογος

# Σφάλματα στις επιδημιολογικές μελέτες

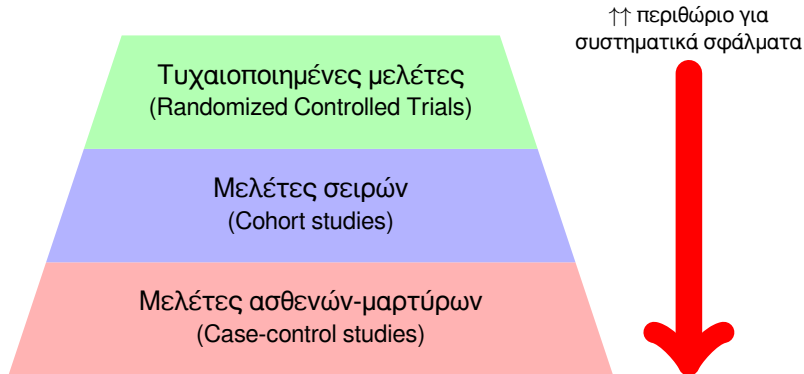
- Τυχαίο σφάλμα (Random error)
- Συστηματικό σφάλμα (Bias)
  - Διαφορά μεταξύ της παρατηρούμενης και της πραγματικής (πληθυσμιακής) τιμής, από οποιαδήποτε αιτία ΕΚΤΟΣ της δειγματοληψίας.
  - ΔΕΝ μπορεί να διορθωθεί στην ανάλυση
- Σύγχυση ή Συγχυτική επίδραση ή Συγχυτικό σφάλμα (Confounding)
  - “Ειδική περίπτωση” συστηματικού σφάλματος
  - Μπορεί (μερικώς) να διορθωθεί στην ανάλυση
    - Υπολειπόμενη σύγχυση (Residual confounding)

# Συστηματικά σφάλματα

- 1 Οφείλονται στον τρόπο που μετράμε, ή στον τρόπο που επιλέγουμε το δείγμα
  - 2 Δεν επηρεάζονται από το μέγεθος του δείγματος
  - 3 Δε μπορούμε να ελέγξουμε την ύπαρξη ή απουσία τους
  - 4 Μπορούμε να κάνουμε μόνο ποιοτικές υποθέσεις γι' αυτά
- 
- Συστηματικό σφάλμα **επιλογής**  
(selection bias)
  - Συστηματικό σφάλμα **πληροφόρησης**  
(information bias)

# Ιεραρχία των τεκμηρίων

Hierarchy of evidence



## Alphabetical list of biases

Allocation of intervention bias	Incidence-prevalence bias (synonym of Neyman bias)	Publication bias
Apprehension bias	Inclusion bias	Purity diagnostic bias
Ascertainment bias	Lack of intention to treat analysis	Recall bias
Berkson's bias	Language bias	Referral filter bias
Centripetal bias	Lead-time bias	Regression dilution bias
Citation bias	Length biased sampling	Regression to the mean
Competing risks	Losses/withdrawals to follow up	Relative control bias
Compliance bias	Mimicry bias	Reporting bias
Confounding by group	Misclassification bias	Rumination bias
Confounding by indication	Missing information in multivariable analysis	Selective survival bias (synonym of Neyman bias)
Contamination bias	Mode for mean bias	Sick quitter bias
Detection bias	Neyman bias	Spectrum bias
Detection bias	Non-differential misclassification bias	Survivor treatment selection bias
Diagnostic/treatment access bias	Non-random sampling bias	Susceptibility bias (synonym of confounding)
Diagnostic suspicion bias	Non-response bias	Telephone random sampling bias
Differential maturing	Obsequiousness bias	Temporal ambiguity
Differential misclassification bias	Observer expectation bias	Unacceptable disease/exposure
Dissemination bias	Observer/interviewer bias	Underreporting bias
Ecological fallacy	Overmatching	Unmasking—detection signal—bias
Exclusion bias	Participant expectation bias	Verification bias (synonym of work up bias)
Exposure suspicion bias	Popularity bias	Will Rogers phenomenon
Family aggregation bias	Post hoc analysis	Work up bias
Friend control bias	Protopathic bias	
Hawthorne effect		
Healthcare access bias		
Healthy volunteer bias		
Healthy worker effect		

# Συστηματικά σφάλματα

## ■ Συστηματικό σφάλμα επιλογής

- Προκύπτει από τον τρόπο επιλογής του δείγματος της μελέτης, που δεν είναι αντιπροσωπευτικός του πληθυσμού ενδιαφέροντος
- (εναλλακτικά) Πιθανότητα επιλογής στο δείγμα όχι ίδια για όλο τον πληθυσμό = μη τυχαίο δείγμα (Εξαρτώμενη από την έκθεση ή την έκβαση)

## ■ Συστηματικό σφάλμα πληροφόρησης

- Προκύπτει από συστηματικά λάθη στον τρόπο μέτρησης της έκθεσης ή της έκβασης ⇒ δυσταξινόμενη

## Σφάλμα επιλογής: παράδειγμα #1

	Ασθενείς με κίρρωση ήπατος	Μάρτυρες (ορθοπεδική)	Μάρτυρες (ΩΡΛ)
Υψηλή κατανάλωση αλκοόλ	80	60	30
Μικρή/καθόλου κατανάλωση αλκοόλ	20	40	70
		OR = 2,66	OR = 9,33
		OR = $(80 \cdot 40) / (20 \cdot 60) = 2,66$	

- Οι μάρτυρες (νοσηλεύόμενοι ορθοπεδικής) είναι πιθανότερο να πίνουν αλκοόλ σε σχέση με το γενικό πληθυσμό
- Πιθανότητα νοσηλείας επηρεάζεται από την έκθεση  
⇒ σφάλμα συμμετοχής (participation bias)



## Σφάλμα επιλογής: παράδειγμα #2

	Ασθενείς με κίρρωση ήπατος	Μάρτυρες (ΩΡΛ)	Μάρτυρες (εθελοντές γεν.πληθ.)
Υψηλή κατανάλωση αλκοόλ	80	30	20
Μικρή/καθόλου κατανάλωση αλκοόλ	20	70	80
		OR = 9,33	OR = 16

- Οι εθελοντές συχνά έχουν διαφορετικά χαρακτηριστικά από το γενικό πληθυσμό

- Μάρτυρες που επιλέγονται μόνοι τους

⇒ σφάλμα εθελοντή (volunteer bias)

ή σφάλμα αυτοεπιλογής (self-selection bias)

# Σφάλμα επιλογής: παράδειγμα #3

## ■ Καφές και Ca παγκρέατος

■ **MacMahon B, Yen S, Trichopoulos D, et al. NEJM 1981;304(11):630–3**

■ 369 ασθενείς, 644 μάρτυρες

■ OR = 1,8 (95% ΔΑ 1,0 – 3,0)

έως 2 κούπες/ημέρα

■ OR = 2,7 (95% ΔΑ 1,6 – 4,7)

≥ 3 κούπες

## ■ Πως επελέγησαν οι μάρτυρες;

■ Από τους ίδιους ιατρούς που διέγνωσαν τους ασθενείς  
(=γαστρεντερολόγοι)

■ Άτομα με γαστρεντερικές διαταραχές -> συμβουλή: διακοπή καφέ

■ Άρα: οι μάρτυρες λιγότερο πιθανό (σε σχέση με το γενικό πληθυσμό)  
να καταναλώνουν καφέ.

⇒ **συστηματικό σφάλμα αποκλεισμού (exclusion bias)**

## Σφάλμα επιλογής: παράδειγμα #4

- Θνησιμότητα μετά από έμφραγμα
  - Vaccarino V, et al. NEJM 1999;341(4):217–25
  - 384.878 εμφραγματίες. Ενδονοσοκομειακή θνησιμότητα.
  - Σύνολο: 11,5% άνδρες, 16,7% γυναίκες
  - <50 ετών: 2,9% άνδρες, 6,1% γυναίκες
  
- Αιφνίδιοι καρδιακοί θάνατοι, **άνδρες/γυναίκες = 3 : 1**
  - Cupples LA, et al. Circulation 1992;85(1S):I11-8
  - Ακόμα και μετά από στάθμιση παραγόντων κινδύνου
  - Η έκθεση (φύλο) επηρεάζει την επιβίωση των περιστατικών

⇒ **συστηματικό σφάλμα επιβίωσης (exclusion bias)**  
ή Neyman bias, incidence-prevalence bias

## Σφάλμα επιλογής: παράδειγμα #5

- Άσθμα σε χοιροτρόφους (βιομηχανικών μονάδων)
  - **Vogelzang PF, et al. Eur Respir J 1999;13(1):187–9**
  - 239 χοιροτρόφοι, 311 αγρότες
  - Υπόθεση: αυξημένη επίπτωση άσθματος (επαγγ. άσθμα)
  - Χρήση ενώσεων 4-ταγούς αμμωνίου: OR = 9,4
  - Όμως επίπτωση: 5,9% χοιροτρόφοι, 5,5% αγρότες
  - Οι χοιροτρόφοι λιγότερο πιθανό να καπνίζουν ή να έχουν ιστορικό ατοπίας
  - Το να εργάζεται κανείς, είναι δείκτης υγείας!  
(οι ασθενείς αποφεύγουν τις επιβλαβείς εργασίες ή αποσύρονται από αυτές όταν νοσήσουν)

⇒ φαινόμενο υγιούς εργάτη (healthy worker effect)

# Σφάλμα πληροφόρησης: παράδειγμα #1

- Εμβολιασμοί και σωματικά συμπτώματα
  - **Murphy D, et al. BMJ 2008;337:a220**
  - 4882 στρατιώτες, ερωτήθηκαν για συμπτώματα χρόνιας κόπωσης σε σχέση με προηγούμενους εμβολιασμούς ( $\geq 2$ )
  - RR = 1,17 (95% ΔΑ 1,05 – 1,30) από συνέντευξη
  - RR = 0,92 (95% ΔΑ 0,66 – 1,28) από ιατρικό φάκελο
  - Διαφορά στην ανάκληση της έκθεσης ανάλογα με την έκβαση
  - Πολύ συνηθισμένη σε εμβολιασμούς, χρόνια νοσήματα, κλπ. (χρονικά απομακρυσμένες εκθέσεις)
  - Συνήθως οι πάσχοντες θυμούνται καλύτερα την επίμαχη έκθεση

⇒ **συστηματικό σφάλμα ανάκλησης (recall bias)**

## Σφάλμα πληροφόρησης: παράδειγμα #2



- Χρήση ναρκωτικών
  - **Turner CF, et al. Science 1998;280(5365):867–73**
  - Επιπολασμός χρήσης in ναρκωτικών, σε 1690 άτομα  
*0% με ερωτηματολόγιο, 0,8% με αυτόματη τηλ. συνέντευξη*
  - Χρήση κρακ ή κοκαΐνης  
*3,3% ερωτηματολόγιο, 6,0% με αυτόματη τηλ. συνέντευξη*
  - Οι συμμετέχοντες δεν απαντούν ειλικρινώς
  - Ο τρόπος που ρωτά ο ερευνητής επηρεάζει τις απαντήσεις
  
- ⇒ **συστηματικό σφάλμα υπεκφυγής (prevarication bias)**
- ⇒ **συστηματικό σφάλμα ερευνητή (interviewer bias)**

# Σφάλμα πληροφόρησης: παράδειγμα #3

## ■ Δυσταξινόμενη



- Bakke PS, et al. Am J Epidemiol 2001;154(5):477–83
- Έκθεση σε σκόνη χαλαζία και πρωϊνός βήχας

### Αυτοαναφερόμενη έκθεση

	Έκθεση(+)	Έκθεση(-)	Ποσοστό
Πρωϊνός βήχας(+) (n=180)	17 	163	9,4%
Πρωϊνός βήχας(-) (n=534)	11 	523	2,1%

OR = 4,96 (95% ΔΑ 2,28 – 10,80)

### Βάσει δομημένης συνέντευξης

	Έκθεση(+)	Έκθεση(-)	Ποσοστό
Πρωϊνός βήχας(+) (n=180)	22 	158	12,2%
Πρωϊνός βήχας(-) (n=534)	40 	494	7,5%

OR = 1,72 (95% ΔΑ 0,99 – 2,98)

# Αρχές αποφυγής συστηματικού σφάλματος

(στις επιδημιολογικές μελέτες)

## ■ Σφάλμα επιλογής

- Σαφείς ορισμός πληθυσμού μελέτης (ποιοί μπαίνουν, ποιοί εξαιρούνται)
- (μελέτες σειρών) Ομάδες με όμοια χαρακτηριστικά πλην έκθεσης
- (μελέτες ασθενών-μαρτύρων) Μάρτυρες να αντιπροσωπεύουν τον πληθυσμό προέλευσης των ασθενών
- Επιλογή μαρτύρων ανεξάρτητη από την έκθεση
- **Τυχαιοποίηση!**

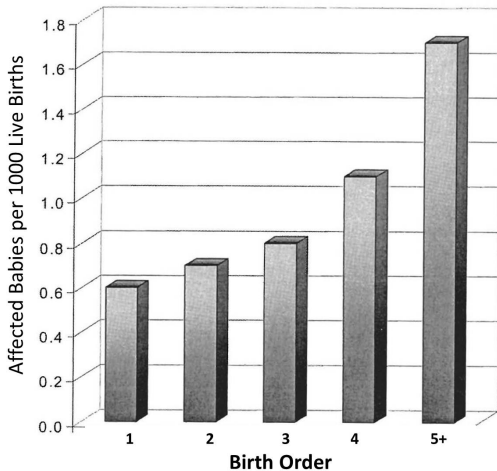
## ■ Σφάλμα πληροφόρησης

- Σαφής ορισμός εκθέσεων και εκβάσεων
- Μετρήσεις κατά το δυνατόν αντικειμενικές, με πρότυπες μεθόδους (π.χ. σταθμισμένα ερωτηματολόγια)
- Δεύτερη πηγή μετρήσεων

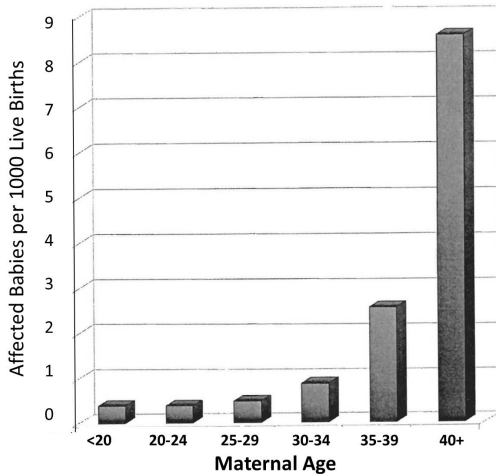


- 1 Εισαγωγή - Τυχαίο vs συστηματικό σφάλμα
- 2 Τα συστηματικά σφάλματα στις επιδημιολογικές μελέτες
- 3 Συγχυτικό σφάλμα (confounding)
- 4 Συστηματικά σφάλματα στις τυχαιοποιημένες μελέτες
- 5 Επίλογος

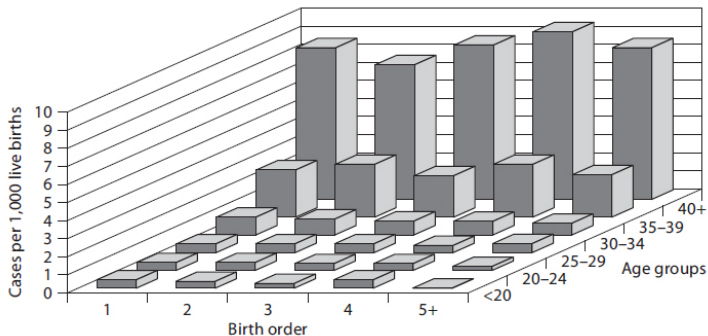
# Παράδειγμα #1: Σύνδρομο Down και σειρά γέννησης



# Παράδειγμα #1: Σύνδρομο Down και σειρά γέννησης



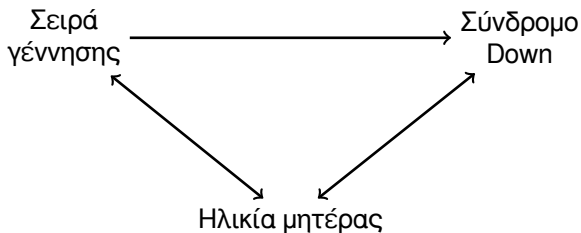
# Παράδειγμα #1: Σύνδρομο Down και σειρά γέννησης



Chen CL. Am J Epidemiol 1999;149(5):442

Για ορισμένη ηλικία μητέρας, η σειρά γέννησης δεν επηρεάζει τον κίνδυνο εμφάνισης συνδρόμου Down

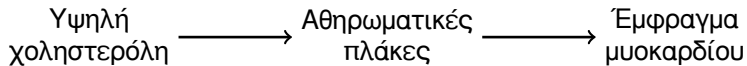
# Συγχυτικός παράγοντας



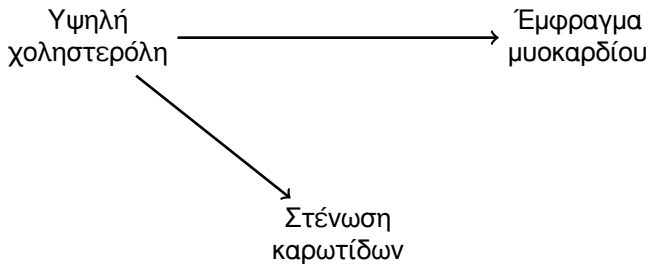
Συγχυτικός παράγοντας = Εξωγενής μεταβλητή που:

- 1 Σχετίζεται με την έκθεση, χωρίς να είναι *συνέπεια* της έκθεσης
- 2 Σχετίζεται με την έκβαση, ανεξάρτητα από την έκθεση

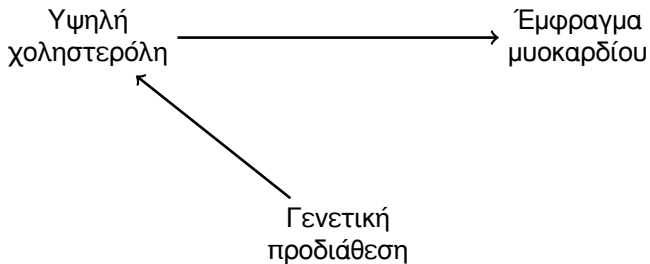
# Όχι συγχυτικός παράγοντας



# Όχι συγχυτικός παράγοντας



# Όχι συγχυτικός παράγοντας





## Παράδειγμα #2

Είναι τα Smart πιο επικίνδυνες από τις Mercedes?

	Σύνολο	Ατυχήματα	Rate(%)	RR (95%CI)
Mercedes	1000	200	20	Ref
Smart	1000	300	30	1.5 (1.3–1.8)

Τι μπορεί να σκεφτούμε? Ηλικία οδηγού

## Παράδειγμα #2

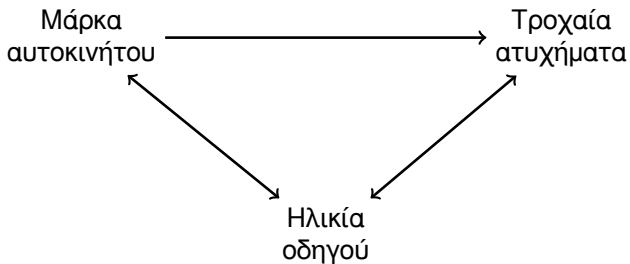
<25 ετών

	Σύνολο	Ατυχήματα	Rate(%)	RR
Mercedes	300	120	40.0	Ref
Smart	550	250	45.5	1.14 (0.9–1.3)

>=25 ετών

	Σύνολο	Ατυχήματα	Rate(%)	RR
Mercedes	700	90	11.4	Ref
Smart	450	50	11.1	0.97 (0.7–1.4)

# Συγχυτικός παράγοντας



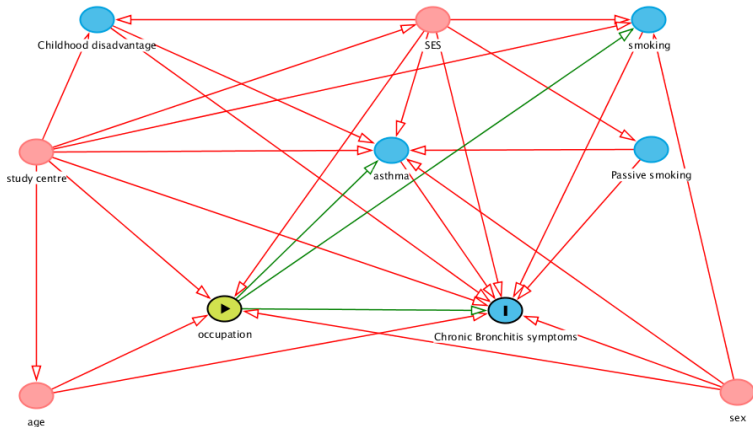
# Πως ελέγχουμε τους συγχυτικούς παράγοντες?

- 1 Περιορισμός (restriction) ή εξομοίωση (matching)
- 2 Διαστρωμάτωση (stratification)
- 3 **Μοντέλα πολλαπλής παλινδρόμησης**  
(π.χ. λογιστική παλινδρόμηση)
  - Μας δίνουν ένα “προσαρμοσμένο” (adjusted) μέτρο συσχέτισης, δηλαδή συσχέτιση με “σταθερές” τις υπόλοιπες μεταβλητές.

Το συγχυτικό σφάλμα **δεν διορθώνεται πλήρως**

- Υπολοιπόμενο συγχυτικό σφάλμα (**residual confounding**)
  - Συγχυτικοί παράγοντες που δεν ελέγξαμε  
(γιατί δε τους σκεφτήκαμε ή/και δε μαζέψαμε πληροφορίες γι' αυτούς)
  - Συγχυτικοί παράγοντες που ελέγξαμε “χάνοντας” πληροφορία  
(γιατί τους “συρρικνώσαμε” σε κατηγορικές ή διχοτόμες μεταβλητές)

# Directed Acyclic Graphs – DAG



Τρόπος οργάνωσης της σκέψης μας.  
(Ποιοί είναι οι πιθανοί συγχυτικοί παράγοντες τους οποίους πρέπει να ελέγξουμε?)

- 1 Εισαγωγή - Τυχαίο vs συστηματικό σφάλμα
- 2 Τα συστηματικά σφάλματα στις επιδημιολογικές μελέτες
- 3 Συγχυτικό σφάλμα (confounding)
- 4 Συστηματικά σφάλματα στις τυχαιοποιημένες μελέτες
- 5 Επίλογος

# Συστηματικό σφάλμα στις RCTs

(Randomized Controlled Trials – Τυχαιοποιημένες κλινικές δοκιμές)

Μπορεί να εισέλθει από την *πίσω πόρτα*, π.χ.:

- Ατελές blinding (ή open-label trial)
- Ατελής τυχαιοποίηση => baseline imbalances
- Σφάλμα απόσυρσης (attrition bias)
- Υποκατάστατες εκβάσεις (surrogate outcomes)
- ...

Άλλα σφάλματα:

- Πολλαπλές συγκρίσεις, αναλύσεις υπο-ομάδων (multiple testing, subgroup analyses)

# Παράδειγμα #1

- Βιταμίνη C και κοινό κρυολόγημα
  - **Karlowski TR, et al. JAMA 1975;231(10):1038-42**
  - 311 άτομα. Παρέμβαση: vit.C ή placebo
  - Ίδια επίπτωση κρυολογήματος, όμως όσοι πήραν vit.C είχαν μικρότερη διάρκεια κρυολογήματος: 5.9 ημέρες vs 7.1 ημέρες
  - Σήμερα ξέρουμε ότι η vit.C δεν έχει επίδραση στο κρυολόγημα.
  - Το placebo δεν είχε την ίδια γεύση με τη vit.C
  - Το 40% των συμμετεχόντων μπορούσε να μαντέψει ορθά το φάρμακο που έλαβε
  - Αυτοί που έλαβαν vit.C “ενιωθαν” πιο καλά!
  - Σε πολλές περιπτώσεις η παρέμβαση **δε γίνεται** να αποκρυβεί (π.χ. χειρουργείο) ⇒ open-label trials

⇒ ατελής blinding



# Ατελής τυχαιοποίηση

**Table 1.** Baseline Demographic and Clinical Characteristics of Participants\*

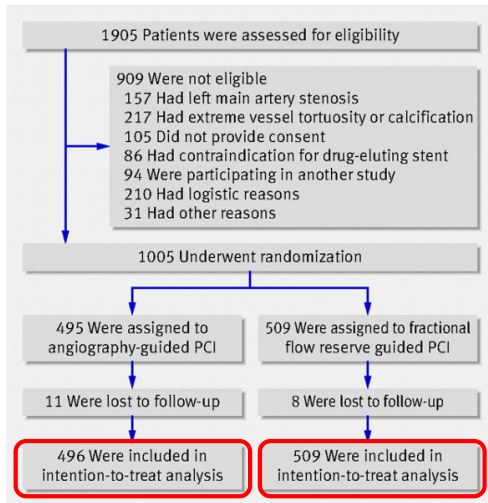
Variable	Surgery (n = 87)	PT (n = 82)
Mean age at enrollment (SD), y†	66.6 (10.5)	69.8 (9.0)
Men, n (%)	44 (51)	44 (54)
Mean height (SD), cm	168.0 (12.5)	170.7 (13.1)
Mean weight (SD), kg	89.1 (19.9)	91.0 (19.3)
Mean BMI (SD), kg/m <sup>2</sup>	31.5 (6.8)	31.1 (5.4)
Race, n (%)		
White	83 (95)	77 (94)
African American	4 (5)	5 (6)
Hispanic ethnicity	0 (0)	0 (0)
Education level, n (%)		
High school or less	38 (44)	40 (49)
Some college	23 (26)	21 (26)
College degree	26 (30)	21 (26)
Marital status, n (%)		
Single	1 (1)	3 (4)
Married	63 (72)	59 (72)
Divorced/separated	14 (16)	8 (10)
Widowed/other	9 (10)	12 (15)
Employment situation, n (%)		
Currently working	22 (25)	17 (21)
Unemployed	4 (5)	1 (1)
Retired	50 (57)	52 (63)
Disabled and/or retired due to health problems	11 (13)	12 (15)



- Allocation concealment
- Λανθασμένη μέθοδος τυχαιοποίησης
- Τυχαία διακύμανση!
  - Ειδικά σε cluster RCTs με λίγα cluster

# Συστηματικό σφάλμα απόσυρσης

## Attrition bias



- Απώλειες κατά την παρακολούθηση
  - Αν όμοιες στις δύο ομάδες → imprecision (↑ τυχαίο σφάλμα)
  - Αν διαφέρουν → bias (συστηματικό σφάλμα)
  - Αυτοί που χάνονται ίσως διαφέρουν από όσους μένουν στη μελέτη.
- Λύση: ανάλυση τύπου *intention-to-treat*
  - Παρονομαστής: όσοι τυχαιοποιήθηκαν

## Παράδειγμα #2

- **Rosiglitazone** και σακχαρώδης διαβήτης
    - Πολλές μικρού και μεσαίου μεγέθους RCTs, που έδειχναν σημαντική επίδραση στη γλυκόζη νηστείας και στην HbA1c (γλυκοζυλιωμένη αιμοσφαιρίνη)
    - Τι θα μπορούσε να πάει στραβά?
    - Ετήσιες πωλήσεις: \$2.5 δις
    - Οι μελέτες δεν έδειχναν όφελος στην καρδιαγγειακή θνησιμότητα.
    - Λίγα χρόνια μετά: κάποιος μάζεψε όλες τις μελέτες (η πλειονότητα *μη δημοσιευμένες* σε επιστημονικά περιοδικά) και έκανε **μετα-ανάλυση**
- Nissen SE. N Eng J Med 2007;356(24):2457
- RR = 1.43 (95%CI 1.03 - 1.98) για έμφραγμα του μυοκαρδίου σε όσους λάμβαναν ροσιγλιταζόνη

⇒ προσοχή στις **υποκατάστατες εκβάσεις**  
(surrogate outcomes)

# Παράδειγμα #3

- Ασπιρίνη μετά από έμφραγμα
  - **ISIS-2 trial, Lancet 1988;2(8607):349-60**
  - 17.187 εμφραγματίες, ασπιρίνη ή placebo
  - 23% χαμηλότερη θνησιμότητα,  $p < 0.00001$
  
  - Ανάλυση 12 υποομάδων, κατά... ζώδιο!
  - Οι δίδυμοι και οι καρκίνοι δεν είχαν όφελος από την ασπιρίνη (μάλιστα είχαν και 9% υψηλότερη θνησιμότητα,  $p = NS$ )

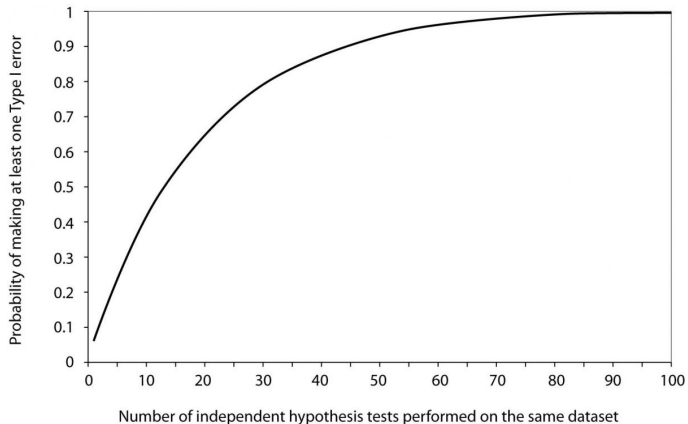
⇒ προσοχή στις **αναλύσεις υποομάδων**  
(subgroup analyses)

# Παράδειγμα #4

Event and Age Group	Treatment Group		Relative Risk (95% CI)*
	Placebo	Pravastatin	
<i>n (%)</i>			
Cancer			
31–64 y	294 (10.8)	264 (9.5)	0.87 (0.73–1.02)
65–75 y	324 (18.3)	367 (21.1)	1.14 (0.98–1.32)
Traumat			
31–64 y	135 (4.9)	125 (4.5)	0.89 (0.69–1.13)
65–75 y	86 (4.9)	88 (5.1)	1.01 (0.75–1.36)
Hepatic system or liver			
31–64 y	106 (3.9)	90 (3.2)	0.81 (0.61–1.07)
65–75 y	67 (3.8)	62 (3.6)	0.92 (0.65–1.30)
Muscular system			
31–64 y	250 (9.2)	278 (10.0)	1.06 (0.89–1.26)
65–75 y	247 (13.9)	231 (13.3)	0.92 (0.77–1.10)
Skin			
31–64 y	90 (3.3)	88 (3.2)	0.94 (0.70–1.27)
65–75 y	68 (3.8)	69 (4.0)	1.02 (0.73–1.42)
Respiratory system‡			
31–64 y	324 (11.9)	384 (13.9)	1.15 (0.99–1.33)
65–75 y	274 (15.5)	272 (15.6)	0.99 (0.83–1.17)
Gastrointestinal system§			
31–64 y	494 (18.1)	495 (17.9)	0.97 (0.86–1.10)
65–75 y	442 (24.9)	405 (23.3)	0.89 (0.77–1.01)
Renal or genital system			
31–64 y	276 (10.1)	307 (11.1)	1.09 (0.93–1.29)
65–75 y	288 (16.2)	310 (17.8)	1.09 (0.93–1.28)

- Είναι το αποτέλεσμα πραγματικό?

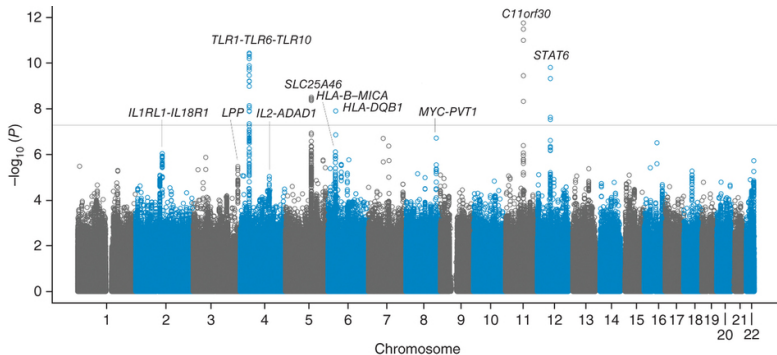
# Πολλαπλοί έλεγχοι και πιθανότητα λάθους



Όσο περισσότερα σκαλίζουμε, τόσο περισσότερα λάθη κάνουμε!

# Παράδειγμα #5: GWAS

(Genome-wide association studies)



Bønnelykke et al. Nat Genet 2013;45:902

- Βάζουμε ψηλότερα τον πήχη
- Επαναλαμβάνουμε τη μελέτη και σε άλλους πληθυσμούς (replication)

- 1 Εισαγωγή - Τυχαίο vs συστηματικό σφάλμα
- 2 Τα συστηματικά σφάλματα στις επιδημιολογικές μελέτες
- 3 Συγχυτικό σφάλμα (confounding)
- 4 Συστηματικά σφάλματα στις τυχαιοποιημένες μελέτες
- 5 Επίλογος**



# Τυχαίο σφάλμα vs Συστηματικά σφάλματα

## Τυχαίο σφάλμα

- Οφείλεται στο γεγονός της δειγματοληψίας
- Μειώνεται όσο μεγαλύτερο δείγμα εξετάζουμε
- Μπορούμε να εκτιμήσουμε το μέγεθός του
- 95% Διάστημα Αξιοπιστίας, τιμή  $p$

## Συστηματικά σφάλματα

- Οφείλονται στον τρόπο που μετράμε, ή στον τρόπο που επιλέγουμε το δείγμα
- Δεν επηρεάζονται από το μέγεθος του δείγματος
- Δεν μπορούμε να ελέγξουμε την ύπαρξή τους
- Συστ. σφάλμα επιλογής, συστ. σφάλμα πληροφόρησης
- Συγχυτικό σφάλμα (confounding)
- Σφάλματα στις RCTs

Συστηματικό σφάλμα = Μια αέναη πνευματική “γυμναστική”



Essay

## Why Most Published Research Findings Are False

John P. A. Ioannidis

PLoS Med 2005;2(8):e124

*“Scientific investigation is the noblest pursuit. I think we can improve the respect of the public for researchers by showing **how difficult success is.**”*

— John P.A. Ioannidis (2007)

# Ευχαριστώ

Το email μου: (για τυχόν απορίες)  
thlytras@gmail.com