

Bioestadística para no estadísticos

1. Principios



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Departament d'Estadística
i Investigació Operativa

1

A u t o r e s



Erik
Cobo



José Antonio
González



Jordi
Cortés



Marta
Vilaró



Nerea
Bielsa



Daniela
Román



Sonia
Navarro

A g r a d e c i m i e n t o s



Enrique Ventuña
eventura@adam.es



Pinceladas

1. Variabilidad
2. Objetivos clínicos
3. Guías según objetivos
4. Investigación frente a Desarrollo
5. Ciencia reproducible y replicable

1.- Determinismo frente a Variabilidad

A) Determinismo: *¿Se hará de día mañana?*

Para decir SÍ solo hace falta

- 1) asumir cierta igualdad entre pasado y futuro

B) Variabilidad: *¿Lloverá mañana?*

Ahora hace falta:

- 1) como antes, asumir igualdad entre pasado y futuro
- 2) hacer un “modelo”: ¿de qué depende la lluvia?
- 3) tratar la variabilidad y cuantificar la incertidumbre

Sin variabilidad... ... ¡podemos cerrar los libros de estadística!



Me basta así: Ángel González

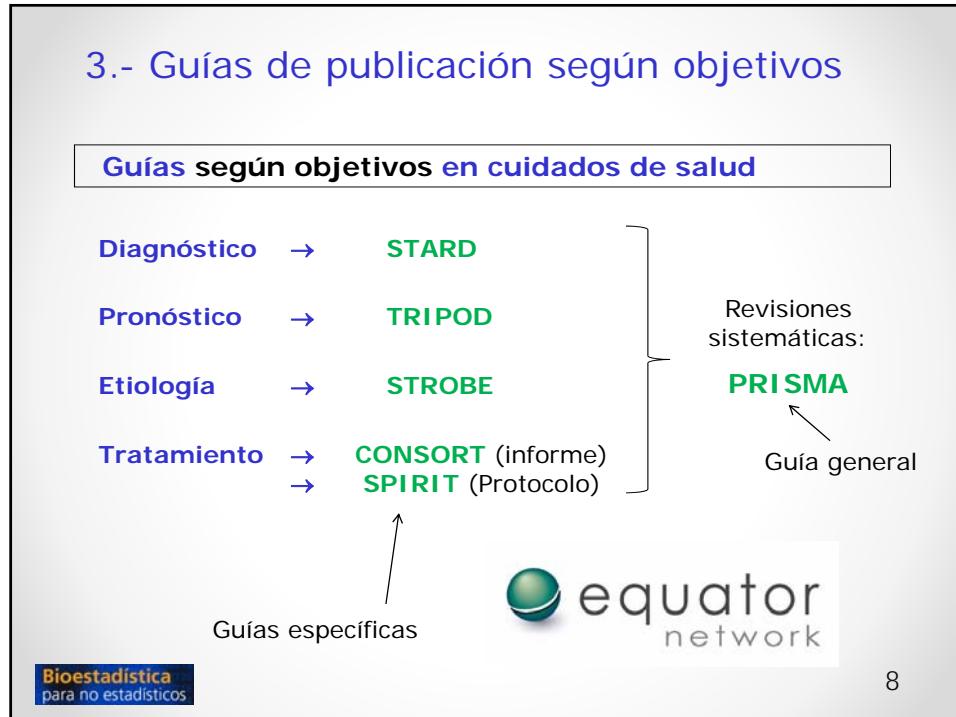
si yo fuese Dios,
 podría repetirte y repetirte,
 siempre la misma y siempre diferente,
 sin cansarme jamás del juego idéntico,
 sin desdeñar tampoco la que fuiste
 por la que ibas a ser dentro de nada;
 ya no sé si me explico, pero quiero
 aclarar que si yo fuese
 Dios, haría
 lo posible por ser Ángel González
 para quererte tal como te quiero,
 para aguardar con calma
 a que te crees tú misma cada día
 a que sorprendas todas las mañanas
 la luz recién nacida con tu propia
 luz, y corras
 la cortina impalpable que separa
 el sueño de la vida,

Sin variabilidad no
hay vida

2.- Objetivos clínicos



3.- Guías de publicación según objetivos



4.- I+D: ¿Investigación o Desarrollo?

Why Most Published Research Findings Are False

John P. A. Ioannidis

Table 4. PPV of Research Findings for Various Combinations of Power ($1 - \beta$), Ratio of True to Not-True Relationships (R), and Bias (u)

$1 - \beta$	R	u	Practical Example	PPV
0.80	1:1	0.10	Adequately powered RCT with little bias and 1:1 pre-study odds	0.85
0.95	2:1	0.30	Confirmatory meta-analysis of good quality RCTs	0.85
0.80	1:3	0.40	Meta-analysis of small inconclusive studies	0.41
0.20	1:5	0.20	Underpowered, but well-performed phase II RCT	0.23
0.20	1:5	0.80	Underpowered, poorly performed phase II RCT	0.17
0.80	1:10	0.30	Adequately powered exploratory epidemiological study	0.20
0.20	1:10	0.30	Underpowered exploratory epidemiological study	0.12
0.20	1:1,000	0.80	Discovery-oriented exploratory research with massive testing	0.0014
0.20	1:1,000	0.20	As in previous example, but with more limited bias (more standardized)	0.0015

Bioestadística para no estadísticos

9



Evidencia: *Información aportada por los datos.*

O bien para construir o bien para ratificar una teoría, pero no para ambas.

I D

I) Exploración: *Genera ideas o hipótesis.*

Se busca el modelo más simple
que mejor ajuste los datos

Todo vale: se torturan los datos

Debe quedar claro: "nuestros resultados sugieren..."

D) Confirmación: *Eleva a teoría una hipótesis previa*

Los datos: ¿ratifican o rechazan el modelo previo?

Sólo sirve el análisis pre-especificado

Permite decir: "hemos establecido que..."

11

5.- Ciencia reproducible y replicable

Pregunta: **la frase** "En general, la literatura médica es sólo una parte de los datos existentes que (...) sobreestima los beneficios, subestima los peligros y tiene un impacto negativo en la atención al paciente y la investigación" **es de:**

1. Un grupo radical antisistema
2. Una asociación anti-ciencia
3. Un grupo de editores y científicos médicos

**SPIRIT 2013 explanation and elaboration:
guidance for protocols of clinical trials**

Cite this as: *BMJ* 2013;346:e7586
doi: 10.1136/bmj.e7586

Overall, the medical literature represents a biased subset of existing data, potentially leading to overestimation of benefits, underestimation of harms, and a detrimental impact on patient care and research. BMJ 372: 377

12

EDITORIAL
The scandal of poor medical research
BMJ 1994; 308 doi: <http://dx.doi.org/10.1136/bmj.308.6924.283> (Published 29 January 1994)
 Cite this as: *BMJ* 1994;308:283

EDITORIAL
The scandal of poor epidemiological research
BMJ 2004; 329 doi: <http://dx.doi.org/10.1136/bmj.329.7471.888> (Published 14 October 2004)
 Cite this as: *BMJ* 2004;329:888

 **The NEW ENGLAND JOURNAL OF MEDICINE**
HOME | ARTICLES & MULTIMEDIA | ISSUES | SPECIALTIES & TOPICS | FOR AUTHORS | CME |

 Perspective
 Improving Patient Safety through Transparency
N Engl J Med 2013; 369:1677-1679 | October 31, 2013 |

Declaración de transparencia para las publicaciones científicas
 Declaration of transparency for scientific publications
 Ferrán Catalá-López, Brian Hutton, David Moher

MEDICINA CLINICA

NIH plans to enhance reproducibility
 Francis S. Collins and Lawrence A. Tabak discuss initiatives that the US National Institutes of Health is exploring to restore the self-correcting nature of preclinical research.

BMJ Helping doctors make better decisions

Is animal research sufficiently evidence based to be a cornerstone of biomedical research?

Research: increasing value, reducing waste 1
How to increase value and reduce waste when research priorities are set
 Ian Chalmers, Michael B Brooks, Ben Djulbegovic, Silvio Garattini, Jonathan Gort, A Metin Gulmezoglu, David I Howell, John P A Ioannidis, Sandy Oliver

Research: increasing value, reducing waste 2
Increasing value and reducing waste in research design, conduct, and analysis

THE LANCET

Research: increasing value, reducing waste 3
Increasing value and reducing waste in biomedical research regulation and management
Research: increasing value, reducing waste 4
Increasing value and reducing waste: addressing inaccessible research
Research: increasing value, reducing waste 5
Reducing waste from incomplete or unusable reports of biomedical research

nature International weekly journal of science

13



Principios

Bioestadística
 para no estadísticos

 UNIVERSITAT POLITÈCNICA DE CATALUNYA
 BARCELONATECH
 Departament d'Estadística
 i Investigació Operativa

14

2. *Intervención* CONSORT (a)



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Departament d'Estadística
i Investigació Operativa

1

1. De la I a la D
 2. Referencia o control
 3. Descriptiva inicial
 4. Comparabilidad
 - Inicial: asignación al azar
 - Final: seguimiento idéntico y completo
 5. Efecto
 - Diferencia de medias
 - Cociente de proporciones
 - Diferencia de proporciones
- 
- A graphic of a paintbrush with a purple handle and a silver ferrule, angled diagonally. It is surrounded by a dynamic, colorful brushstroke in shades of pink, red, and blue, suggesting motion and application.

1. Tipos de ensayos clínicos según la fase de I+D

Fase I+D	Objetivo	Voluntarios	Tamaño *	Diseño	Seguimiento *
I	Seguridad	Sanos	Unidades	Escalado de dosis	Días
II	Exploración Eficacia		Decenas		Semanas
III	Confirmación Eficacia	Pacientes	Centenas	Aleatorizado Enmascarado	Meses
IV	Efectividad y/o Eficiencia		Miles		Años

*Según objetivo concreto

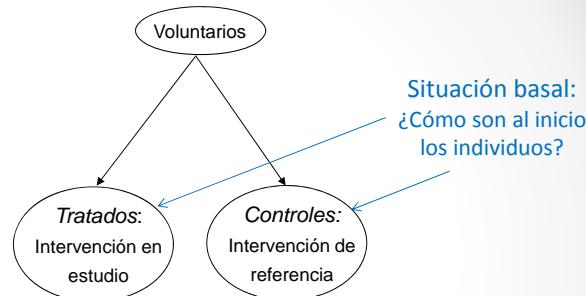
Bioestadística
para no estadísticos

2. Referencia



Bioestadística
para no estadísticos

3. Descriptiva



Bioestadística
para no estadísticos

5

BMJ CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials

Table 4 Example of reporting baseline demographic and clinical characteristics.* (Adapted from table 1 of Yusuf et al., 2005)	
Telmisartan (N = 2954)	Placebo (N = 2972)
Age (years)	66.9 (7.3)
Sex (female)	1280 (43.3%)
Smoking status:	
Current	293 (9.9%)
Past	1273 (43.1%)
Ethnic origin:	
Asian	637 (21.6%)
Arab	37 (1.3%)
African	51 (1.7%)
European	1801 (61.0%)
Native or Aboriginal	390 (13.2%)
Other	38 (1.3%)
Blood pressure (mm Hg)	140.7 (16.8/81.8) (10.1)
Heart rate (beats per min)	68.8 (11.5)
Cholesterol (mmol/l):	
Total	5.09 (1.48)
LDL	3.03 (1.01)
HDL	1.27 (0.37)
Coronary artery disease	2211 (74.8%)
Myocardial infarction	1381 (46.8%)
Angina pectoris	1412 (47.8%)
Peripheral artery disease	349 (11.8%)
Hypertension	2259 (76.5%)
Diabetes	1059 (35.8%)

* Data are mean (SD) or numbers (%).

SD = Standard Deviation
= Desviación estándar
= Desviación Típica
~distancia típica al centro

Así empiezan

2954 casos con Telmisartan y
2972 con la referencia

Los promedios iniciales de
LDL eran de 3.02 y 3.03

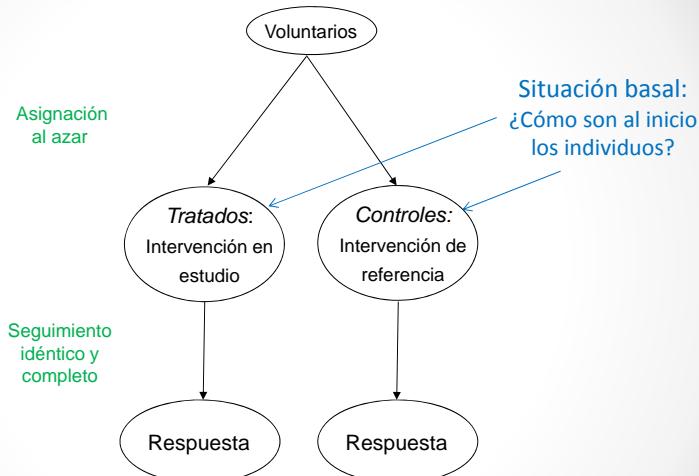
1.01 y 1.02 son la distancia
típica de un caso al centro

En el grupo “tratado”,
1412 de 2854 (47.8%)
tenían historia de angina

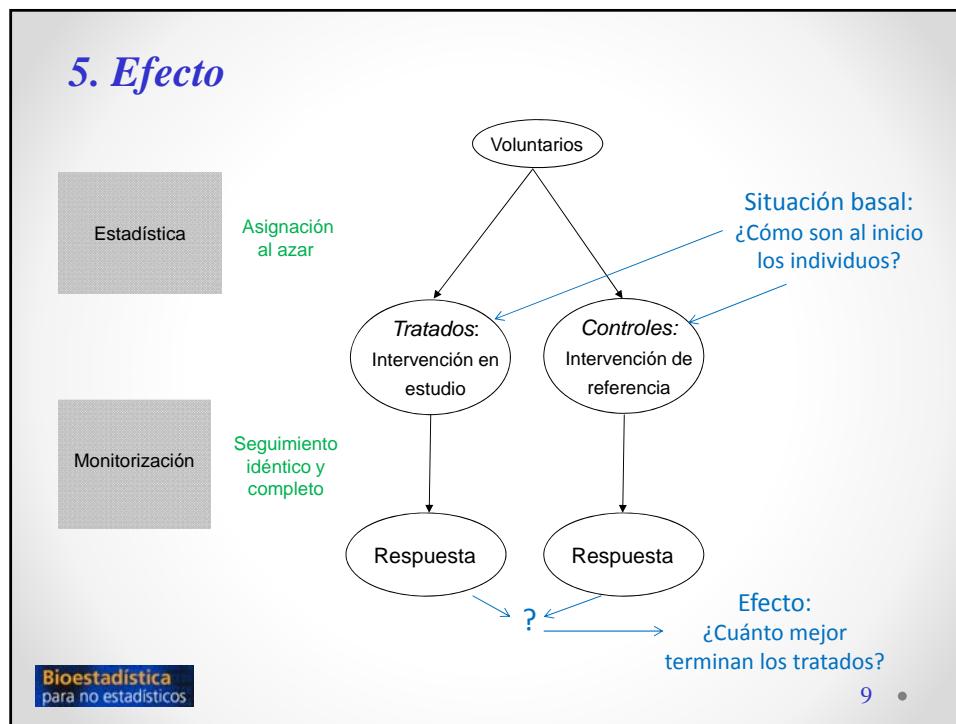
Y en el grupo de
referencia, un 47.5%
(1412 de 2972)



4. Comparabilidad



5. Efecto



BMJ CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials

Así terminan... y así empiezan (baseline)

65 y 66 voluntarios en los grupos

Table 6 | Example of reporting of summary results for each study group (continuous outcomes). (Adapted from table 3 of van Linschoten²³⁴)

	Exercise therapy (n=65)		Control (n=66)		Adjusted difference* (95% CI) at 12 months
	Baseline (mean (SD))	12 months (mean (SD))	Baseline (mean (SD))	12 months (mean (SD))	
Function score (0-100)	64.4 (13.9)	83.2 (14.8)	65.9 (15.2)	79.8 (17.5)	4.52 (-0.73 to 9.76)
Pain at rest (0-100)	4.14 (2.3)	1.43 (2.2)	4.03 (2.3)	2.61 (2.9)	-1.29 (-2.16 to -0.42)
Pain on activity (0-100)	6.32 (2.2)	2.57 (2.9)	5.97 (2.3)	3.54 (3.38)	-1.19 (-2.22 to -0.16)

*Function score adjusted for baseline, age, and duration of symptoms.

Al final, el ejercicio gana por $-1.18 = 1.43 - 2.61$

Como al inicio perdía por $+0.11 = 4.14 - 4.03$, ajustando por los valores iniciales, el efecto es $-1.29 = -1.18 - 0.11$

Ajustando además por edad y duración de los síntomas, también vale -1.29

BMJ CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials

Interpretación de los intervalos de confianza

Table 6 | Example of reporting of summary results for each study group (continuous outcomes).
(Adapted from table 3 of van Linschoten²³⁴)

	Exercise therapy (n=65)		Control (n=66)		Adjusted difference* (95% CI) at 12 months
	Baseline (mean (SD))	12 months (mean (SD))	Baseline (mean (SD))	12 months (mean (SD))	
Function score (0-100)	64.4 (13.9)	83.2 (14.8)	65.9 (15.2)	79.8 (17.5)	4.52 (-0.73 to 9.76)
Pain at rest (0-100)	4.14 (2.3)	1.43 (2.2)	4.03 (2.3)	2.61 (2.9)	-1.29 (-2.16 to -0.42)
Pain on activity (0-100)	6.32 (2.2)	2.57 (2.9)	5.97 (2.3)	3.54 (3.38)	-1.19 (-2.22 to -0.16)

*Function score adjusted for baseline, age, and duration of symptoms.

El efecto puntual del ejercicio en esta puntuación vale 4.52, si bien valores negativos y positivos son compatibles con estos datos, luego no tenemos evidencia a favor de un efecto sólo positivo

Ahora bien, el efecto puntual en el dolor en reposo vale -1.29. Como sólo valores negativos (entre bajar 0.42 y 2,16) son compatibles con estos datos, hay evidencia a favor de efecto en el dolor: el ejercicio baja el dolor, por lo menos, en 0.42

BMJ CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials

30 pacientes observados en cada grupo

Así terminan
26 de los 30 pacientes tuvieron una respuesta positiva, lo que representa un 86.67%≈87%

Table 5 | Example of reporting of summary results for each study group (binary outcomes).*
(Adapted from table 2 of Mease et al¹⁰³)

Endpoint	Number (%)		Risk difference (95% CI)
	Etanercept (n=30)	Placebo (n=30)	
Primary endpoint			
Achieved PsARC at 12 weeks	26 (87)	7 (23)	63% (44 to 83)
Secondary endpoint			
Proportion of patients meeting ACR criteria:			
ACR20	22 (73)	4 (13)	60% (40 to 80)
ACR50	15 (50)	1 (3)	47% (28 to 66)
ACR70	4 (13)	0 (0)	13% (1 to 26)

*See also example for item 6a.
PsARC=psoriatic arthritis response criteria. ACR=American College of Rheumatology.

La diferencia entre 87 y 23% es 63% ($86.67 - 23.33 = 63.33 \approx 63\%$)

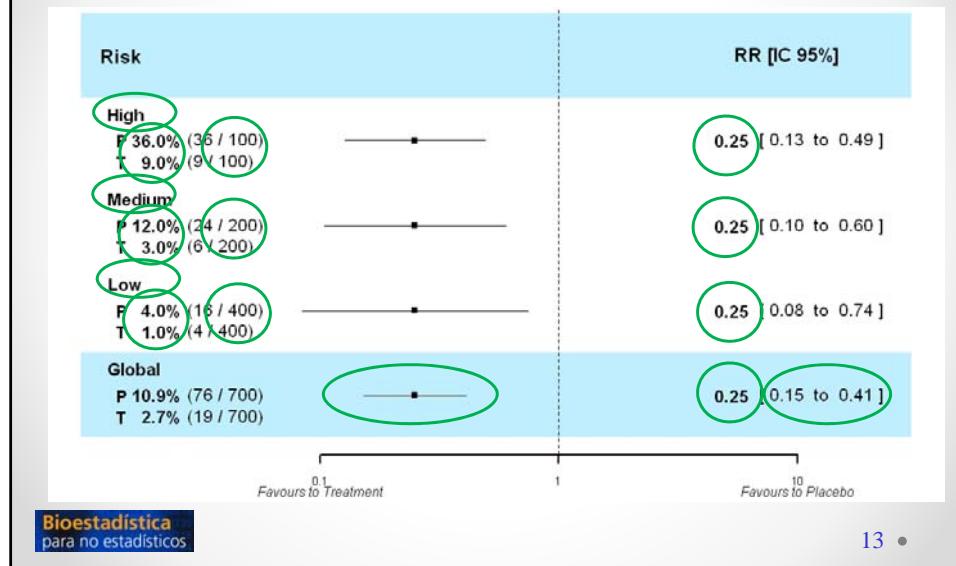
Añadir Etanercept a la práctica clínica aumenta en 0.63 la probabilidad de que un paciente tenga una evolución positiva.

Estos resultados podrían venir de diferencias poblacionales entre 44 y 83%.

Por lo menos, la intervención aumenta la probabilidad de positivo en un 0.44.

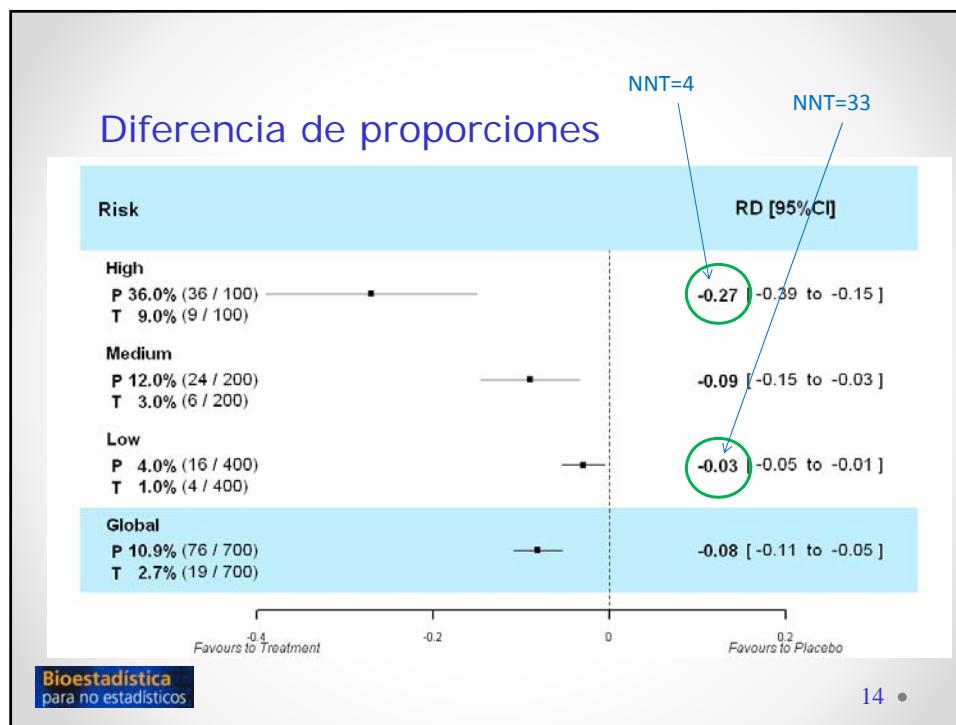
Medidas del efecto en respuestas dicotómicas

Cociente (o razón) de proporciones



13 •

Diferencia de proporciones



14 •

**Interprete ambos:
diferencias y cocientes**

BMJ CONSORT 2010 Explanation and Elaboration: updated
guidelines for reporting parallel group randomised trials

**Table 7 | Example of reporting both absolute and relative effect sizes. (Adapted from table 3 of
The OSIRIS Collaborative Group²⁴²)**

Primary outcome	Percentage (No)		Risk ratio (95% CI)	Risk difference (95% CI)
	Early administration (n=1344)	Delayed selective administration (n=1346)		
Death or oxygen dependence at "expected date of delivery"	31.9 (429) EA	38.2 (514) DSA	0.84 (0.75 to 0.93)	-6.3 (-9.9 to -2.7)

Descriptiva de resultados por grupo: "En los EA , 429 de 1344 (31.9%) tuvieron el evento"

31.9 / 38.2 = 0.84
En EA, la probabilidad de tener un evento fue 0.84 veces la de DSA.

31.9 - 38.2 = -6.3
En EA, la proporción de eventos fue un 6.3% inferior respecto a DSA.

Bioestadística para no estadísticos

Intervención CONSORT (a)

**Bioestadística
para no estadísticos**

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Departament d'Estadística
i Investigació Operativa

16

Bioestadística
para no estadísticos

3. *Intervención* CONSORT (b)



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Departament d'Estadística
i Investigació Operativa

1



1. Objetivo frente a hipótesis
2. Control (o ajuste)
3. Bendito azar
4. Control del seguimiento
5. Riesgos de sesgo... y su prevención

1. Objetivo frente a Hipótesis

A) Objetivo: Es la finalidad *subjetiva* del estudio.

- Ejemplos:
 - demostrar que A es mejor que B
 - permitir el acceso a un nuevo dispositivo

B) Hipótesis: que deseamos confirmar

Debe ser *reproducible* (en los mismos datos) y
replicable (en otros datos)

Lo que requiere pre-especificar sin ambigüedades:

- la respuesta (método y momento de obtención)
- el análisis estadístico concreto

Para garantizar que sólo llegan a los pacientes intervenciones eficaces.

Bioestadística
para no estadísticos

3 ●

2. Opciones para el control (o ajuste)

Opción	Fase*	Nombre	<i>Bendito control</i>	<i>Maldito control</i>
Criterios de selección	Diseño	Criterios de selección	Control Completo Barato Simple de diseñar Simple de analizar	Reduce generalidad Limitado número de variables Posible confusión residual (si restricciones amplias)
	Análisis	Análisis de 1 subgrupo		
Análisis de subgrupos	Diseño	Bloques (Apareamiento)	Potencia Eficiencia Sin premisas Directa Cálculo simple	Pierde flexibilidad Coste Dispersión de casos en estratos Diferentes estratificaciones Difícil 'sumarización'
	Análisis	Estratificación (Apareamiento)		
Modelado	Diseño	Covarianza	Factible con pocos casos Redondea efectos menores Permite predicciones Permite continuas Permite varias	Muchas premisas Elección del modelo Elección de variables Interpretación Parametrización del software
	Análisis	Regresión		
Global	Diseño	Minimización	Ajusta varias a la vez No reduce generalidad	Logística sofisticada
	Análisis	Par óptimo		

*Confirmar requiere pre-especificar el método de control. Ajustar en el análisis no corrige errores del diseño.

Adaptada de Kleinbaum et al. Epidemiologic Research

4 ●

3. Azar



Avicena (980-1037)



Lind (1747)



Hill (1948)

BRITISH MEDICAL JOURNAL

London Saturday October 30 1948

STREPTOMYCIN TREATMENT OF PULMONARY TUBERCULOSIS
A MEDICAL RESEARCH COUNCIL INVESTIGATION

Bioestadística
para no estadísticos

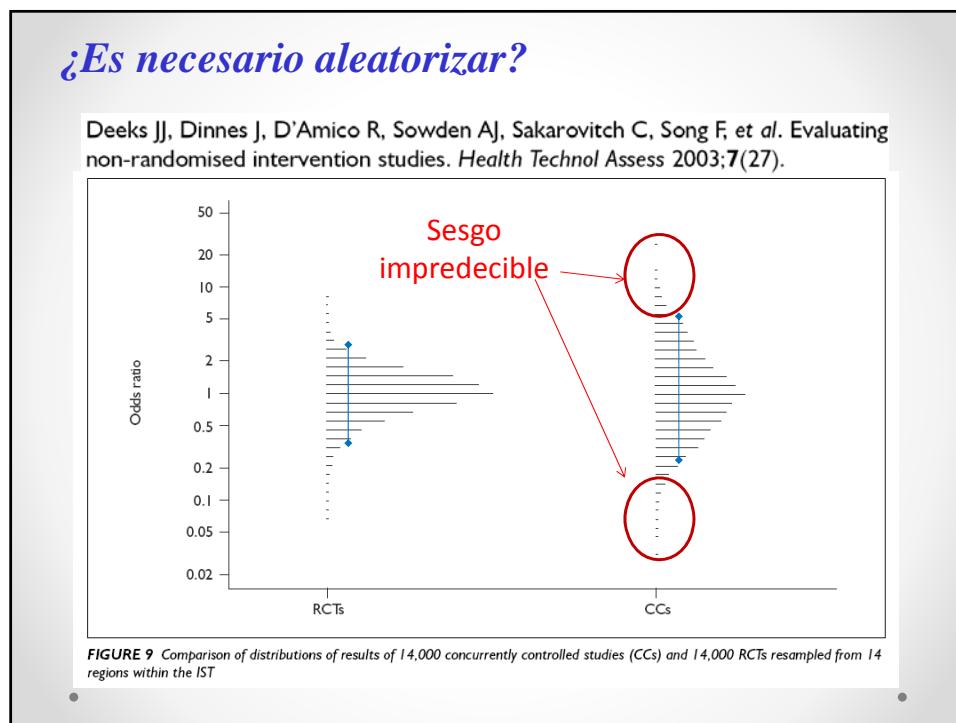
5 •

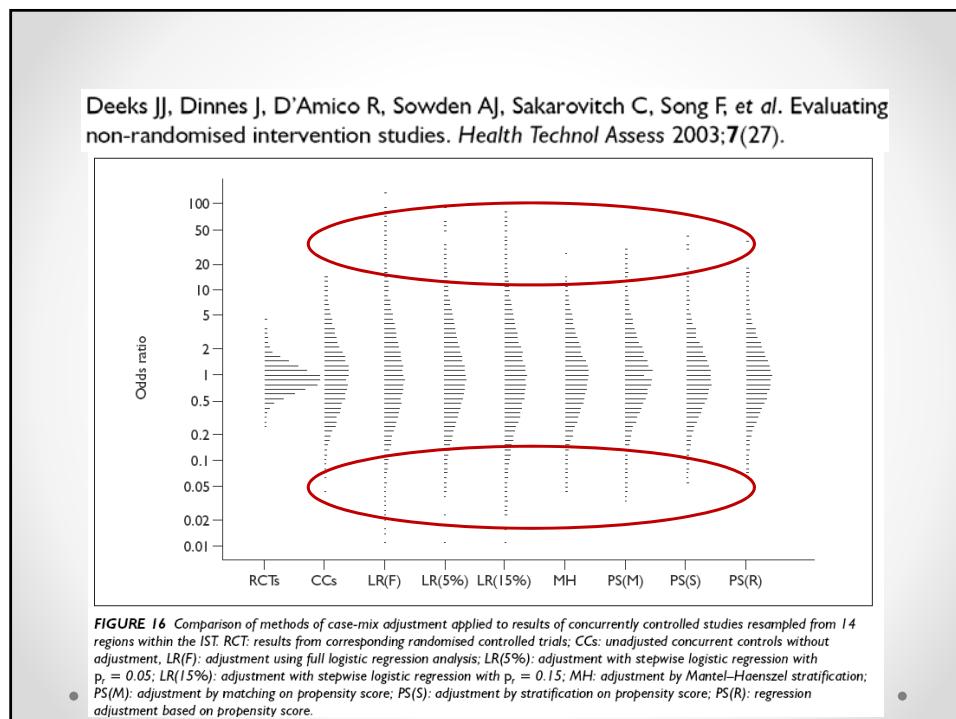
Azar, bendito azar



Bioestadística
para no estadísticos

6 •





*Maldito control**Maldito azar*

Bioestadística
para no estadísticos

11 •

Estadística

Asignación
al azar

Monitorización

Seguimiento
idéntico y
completo

Voluntarios

Tratados:
Intervención en
estudio

Controles:
Intervención de
referencia

Respuesta

Respuesta

Bioestadística
para no estadísticos

12 •

4. Control del seguimiento

The NEW ENGLAND JOURNAL of MEDICINE

SPECIAL REPORT

The Prevention and Treatment of Missing Data in Clinical Trials

Roderick J. Little, Ph.D., Ralph D'Agostino, Ph.D., Michael L. Cohen, Ph.D., Kay Dickersin, Ph.D., Scott S. Emerson, M.D., Ph.D., John T. Farrar, M.D., Ph.D., Constantine Frangakis, Ph.D., Joseph W. Hogan, Sc.D., Geert Molenberghs, Ph.D., Susan A. Murphy, Ph.D., James D. Neaton, Ph.D., Andrea Rotnitzky, Ph.D., Daniel Scharfstein, Sc.D., Weichung J. Shih, Ph.D., Jay P. Siegel, M.D., and Hal Stern, Ph.D.

N ENGL J MED 367;14 NEJM.ORG OCTOBER 4, 2012

Bioestadística para no estadísticos 13 •

Eficacia frente a Efectividad

A) Eficacia del método:

Con adherencia absoluta al protocolo de la intervención

Ejemplo: menos del 1% (límite superior del IC_{95%}) de mujeres que toman regularmente anticonceptivos se quedan embarazadas.

De mayor interés para teóricos que desarrollan intervenciones.

B) Eficacia de uso o Efectividad:

En condiciones normales del entorno considerado.

De mayor interés para clínicos y gestores.

Otras: Balance beneficios/riesgos

Balance eficacia/coste o eficiencia

Usual population definitions (ICH-E9)

Per protocol set (o “datos evaluables” o “casos válidos”): The set of data generated by the subset of subjects who complied with the protocol sufficiently to ensure that these data would be likely to exhibit the effects of treatment, according to the underlying scientific model. Compliance covers such considerations as exposure to treatment, availability of measurements and absence of major protocol violations.

Intention-to-Treat principle (o “datos completos” o “tal como se aleatorizaron”): The principle that asserts that the effect of a treatment policy can be best assessed by evaluating on the basis of the intention to treat a subject (i.e. the planned treatment regimen) rather than the actual treatment given. It has the consequence that subjects allocated to a treatment group should be followed up, assessed and analysed as members of that group irrespective of their compliance to the planned course of treatment.

5. Riesgos de sesgo... y su prevención

1) Sesgo de selección

Se previene con la aleatorización y la asignación oculta

2) Sesgo de actuación o rendimiento (performance)

Se previene con el uso del placebo y el enmascaramiento

3) Sesgo por desgaste (attrition) o pérdida de casos.

Se previene con un seguimiento completo

4) Sesgo de evaluación

Se previene con el enmascaramiento o el uso de variables “automáticas”

5) Sesgo de análisis (selective outcome reporting)

Se previene siguiendo el plan de análisis previo (protocolo)

No confunda ‘dejar el tratamiento en estudio’
con ‘dejar el estudio’

A cartoon illustration within a grey frame. On the left, a man in a lab coat stands holding a clipboard, saying: "EN MIS INVESTIGACIONES, YO SÓLO USO INVESTIGADORES Y VOLUNTARIOS CIEGOS". In the center, another man in a lab coat sits at a desk with a computer monitor displaying a grid of letters (L P E D / P E C F D / F E L O P Z D / D E F O T E C / Z L E P O T E C / P E T). He is looking at the monitor and says: "BUENO, A MÍ ME BASTA CON OCULTARLES LA INTERVENCIÓN". The desk has a nameplate that reads "OFTALMOLOGO". To the right, the word "Intervención" is written vertically, followed by "CONSORT (b)" in a large blue font.

Bioestadística
para no estadísticos

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Departament d'Estadística
i Investigació Operativa

17

Bioestadística
para no estadísticos

4. *Protocolos* SPIRIT



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Departament d'Estadística
i Investigació Operativa



1. Ciencia y Técnica
2. Tamaño muestral
3. Variables
4. Responsabilidades
5. Diseminación

Bioestadística
para no estadísticos

2

1.- Ciencia y Técnica

A) Ciencia → ¿Qué sé? → Conocimiento → Evidencia

'Evidencia' de Fisher: proceso inductivo → Inferencia

P-valor o significación: evidencia en contra de H_0

Reportamos P : menor P , mayor evidencia

Solo tiene H_0 : "ausencia de pruebas no es prueba de ausencia"

Absuelto no implica inocente

A VECES ME SIENTO COMO UNA HIPÓTESIS, PUEDO SER ABANDONADA, PERO NUNCA ACEPTADA



1.- Ciencia y Técnica

A) Ciencia → ¿Qué sé? → Conocimiento → Evidencia

'Evidencia' de Fisher: proceso inductivo → Inferencia

P-valor o significación: evidencia en contra de H_0

Reportamos P : menor P , mayor evidencia

Solo tiene H_0 : "ausencia de pruebas no es prueba de ausencia"

B) Técnica → ¿Qué hago? → Decisión

Newman-Pearson minimizan los errores a largo plazo.

Es "frecuentista": fijamos un % máximo α d errores tipo I
y minimizamos el % β de errores tipo II

Se reportan α y β (en lugar de P)

Permite 'adoptar' ambas decisiones

Tabaco: ¿Saber o hacer?

A) Ciencia → ¿Qué sé sobre los efectos del tabaco?

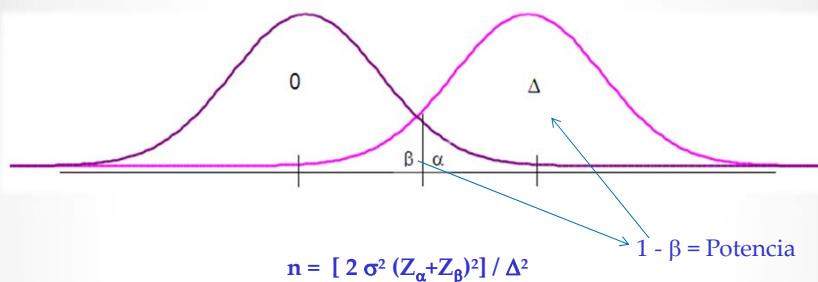
Fisher: No experimentos en personas → No evidencia

Hill: Criterios de Koch → Causalidad interpretación más verosímil

B) Técnica → ¿Qué hago?

Greenland: Humo en la sierra → ¿enviar ecólogos o bomberos?

2. Tamaño de un ensayo decisorio



El tamaño muestral depende de:

Δ : magnitud del **efecto** de la intervención

$\downarrow \Delta \rightarrow \uparrow n$

α : **riesgo** de autorizar cuando **NO** hay efecto

$\downarrow \alpha, \beta \rightarrow \uparrow n$

β : **riesgo** de no autorizar cuando el efecto es Δ

$\uparrow \sigma \rightarrow \uparrow n$

σ : variabilidad de la respuesta

Example

σ → To detect a mean difference in SFQ score of 2 point ($SD = 4.53$) at 72 weeks with a two-sided significance level of 1% and power of 80% with equal allocation to two arms would require 120 patients in each arm of the trial.

α

$1 - \beta$

n



ALTERNATIVAS AL INCREMENTO DE CASOS		
↓ Dispersión	inter-individual	a) subgrupo mas homogéneo b) estratificación c) apareamiento d) evolución respecto basal e) modelado que suprime fuentes de variación
	intra-individual	a) promedios de varias mediciones c) búsqueda de variables mas fiables
↑ efecto a estimar	a) concentrarse en casos mas 'sensibles' b) usar variable mas relacionada con la intervención c) usar variable con mejores propiedades estadísticas	

3. Variable respuesta

BMJ SPIRIT 2013 explanation and elaboration:
guidance for protocols of clinical trials

Outcome = Endpoint = Output =
= Variable respuesta = desenlace = criterio de valoración

Examples

“Primary outcome

Primary outcome
Delirium recognition: In accordance with national guidelines [reference], the study will identify delirium by using the RASS [Richmond Agitation-Sedation Scale] and the CAM-ICU [Confusion Assessment Method for the intensive care unit] on all patients who are admitted directly from the emergency room or transferred from other services to the

Secondary outcomes

Secondary outcomes
The study will collect demographic and baseline functional information from the patient's legally authorized representative and/or caregivers. Cognitive function status will be obtained by interviewing the patient's legally authorized representative using the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE). IQCODE is

Singular (para decidir)

Plural (para saber más)

Bioestadística
para no estadísticos

10

The World Health Organization is the United Nations specialized agency for health. It was established on 7 April 1948. WHO's objective, as set out in its Constitution, is the attainment by all peoples of the highest possible level of health. Health is defined in WHO's Constitution as a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity.

11

Variables subrogadas*

* "intermedias", "sucedáneas", "anticipadas", "prestadas"

Criterio de Prentice:

Para ser válida, una variable intermedia Y'
debe "conducir" **toda** la relación de la intervención X
con la auténtica respuesta de interés Y

12

Fiabilidad de la respuesta

Cuento menos repetibles son los valores en un mismo caso (variabilidad intra-caso), menor es la fiabilidad.

Menor fiabilidad, requiere mayor 'n'.

La fiabilidad puede aumentarse:

- Estandarizando la determinación
- Calibrando los dispositivos de medida
- Controlando las condiciones del paciente
- Acreditando los evaluadores
- Promediando repeticiones (excluyendo o no valores extremos)

El diseño de experimentos y el análisis estadístico tienen métodos para ajustar y controlar tanto la variabilidad intra-casos como la entre-casos.

Bioestadística
para no estadísticos

13

BMJ SPIRIT 2013 explanation and elaboration:
guidance for protocols of clinical trials

4. Responsabilidades en un ensayo

Example

“Principal investigator and research physician

Design and conduct of RITUXVAS

Preparation of protocol and revisions

Preparation of investigators brochure (IB) and CRFs [case report forms]

Organising steering committee meetings

Managing CTO [clinical trials office]

Publication of study reports

Members of TMC [Trial Management Committee]

Steering committee (SC)

(see title page for members)

Agreement of final protocol

All lead investigators will be steering committee members. One lead investigator per country will be nominated as national coordinator.

Recruitment of patients and liaising with principle [sic] investigator

Reviewing progress of study and if necessary agreeing changes to the protocol and/or investigators brochure to facilitate the smooth running of the study.

Trial management committee (TMC)
 (Principle [sic] investigator, research physician, administrator)

Study planning
 Organisation of steering committee meetings
 Provide annual risk report MHRA [Medicines and Healthcare Products Regulatory Agency] and ethics committee
 SUSAR [Serious unexpected suspected adverse events] reporting to MHRA and Roche
 Responsible for trial master file
 Budget administration and contractual issues with individual centres
 Advice for lead investigators
 Audit of 6 monthly feedback forms and decide when site visit to occur.
 Assistance with international review, board/independent ethics committee applications
 Data verification
 Randomisation
 Organisation of central serum sample collection

Data manager
 Maintenance of trial IT system and data entry

Data verification

Lead investigators
 In each participating centre a lead investigator (senior nephrologist/rheumatologist/ immunologist) will be identified, to be responsible for identification, recruitment, data collection and completion of CRFs, along with follow up of study patients and adherence to study protocol and investigators brochure. ... Lead investigators will be steering committee members, with one investigator per country being nominated as national coordinator.⁶²

Example
 "Trial Sponsor: University of Nottingham
 Sponsor's Reference: RIS 8024...
 Contact name: Mr PC
 Address: King's Meadow Campus...
 Telephone: ...
 Email: ..."⁵⁰

Example
 "This funding source had no role in the design of this study and will not have any role during its execution, analyses, interpretation of the data, or decision to submit results."⁵⁴

Bioestadística para no estadísticos

5. Diseminación

BMJ SPIRIT 2013 explanation and elaboration:
guidance for protocols of clinical trials

Dissemination policy—trial results

Item 31a: Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions

Dissemination policy—reproducible research

Item 31c: Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code

Bioestadística
para no estadísticos

17

YO PAGO POR NO SEGUIR LAS LEYES DE MENDEL

Y YO POR NO SEGUIR LAS LEYES DE NEWTON

Essentially, all models are wrong, but some are useful

Protocolos
SPIRIT

Bioestadística
para no estadísticos

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Departament d'Estadística
i Investigació Operativa

18

Bioestadística
para no estadísticos

5. *Revisiones* PRISMA



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Departament d'Estadística
i Investigació Operativa



1. La pregunta
2. Evidencia y decisión
3. Efecto constante en un ensayo
4. Heterogeneidad en un meta-análisis
5. Recomendaciones

Bioestadística
para no estadísticos

2

PLoS Medicine The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration

1. La pregunta

Box 2. Helping To Develop the Research Question(s): The PICOS Approach

Formulating relevant and precise questions that can be answered in a systematic review can be complex and time consuming. A structured approach for framing questions that uses five components may help facilitate the process. This approach is commonly known by the acronym "PICOS" where each letter refers to a component: the patient population or the disease being addressed (P), the interventions or exposure (I), the comparator group (C), the outcome or endpoint (O), and the study design chosen (S) [186]. Issues relating to PICOS impact several PRISMA items (i.e., Items 6, 8, 9, 10, 11, and 18).

Population
Intervention
Comparator
Outcome
Study design

Bioestadística para no estadísticos

3

2. Evidencia y Decisión

I.- Ciencia y Técnica

A) **Ciencia** → ¿Qué sé? → Conocimiento → **Evidencia**
 'Evidencia' de Fisher: proceso induktivo → Inferencia
 P-valor o significación: evidencia en contra de H_0
 Reportamos P ; menor P , mayor evidencia
 Solo tiene H_0 : "ausencia de pruebas no es prueba de ausencia"

B) **Técnica** → ¿Qué hago? → **Decisión**
 Newman-Pearson minimizan los errores a largo plazo.
 Es "frecuentista": fijamos un % máximo α de errores tipo I y minimizamos el % β de errores tipo II
 Se reportan α y β (en lugar de P)
 Permite 'adoptar' ambas decisiones

3

Deming propuso el ensayo confirmatorio de fase III como pivote para la **decisión** de autorizar una intervención

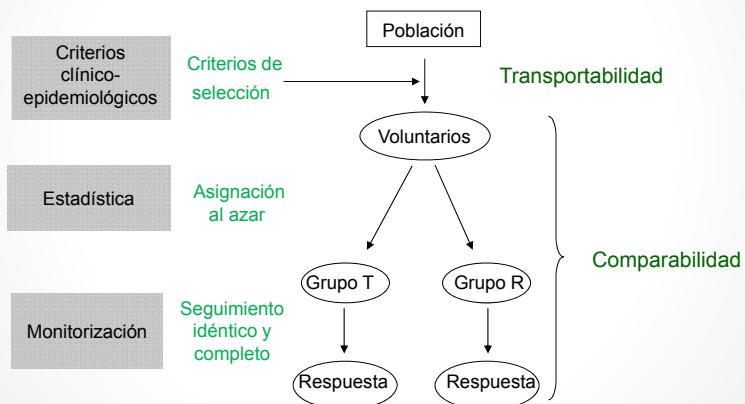
Chalmers impulsó la acumulación de toda la **evidencia** disponible sobre una intervención

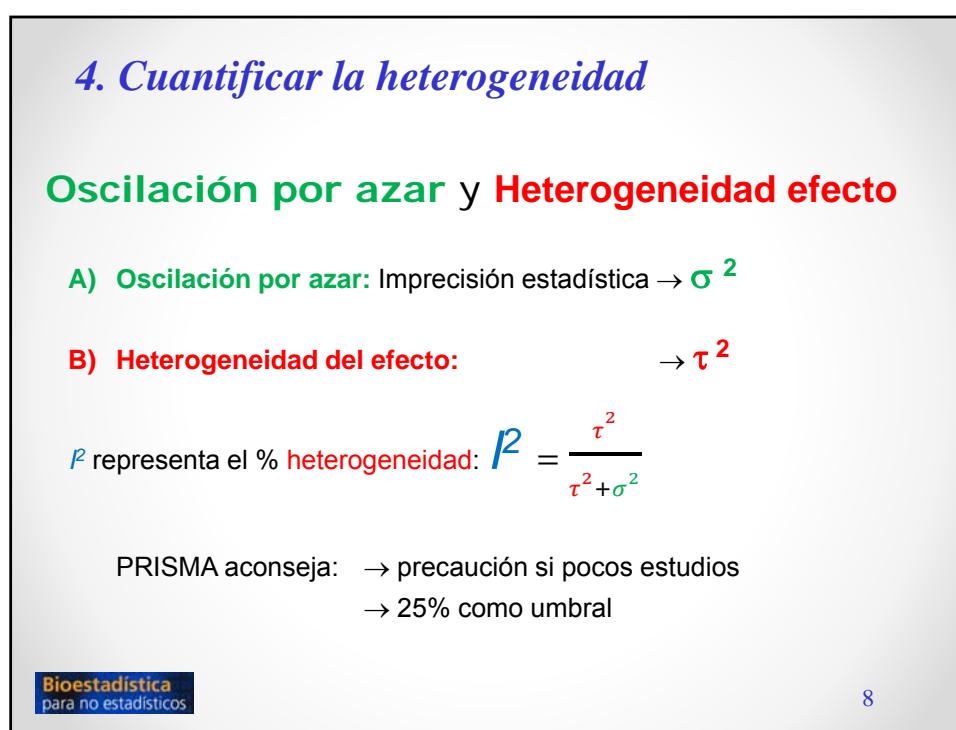
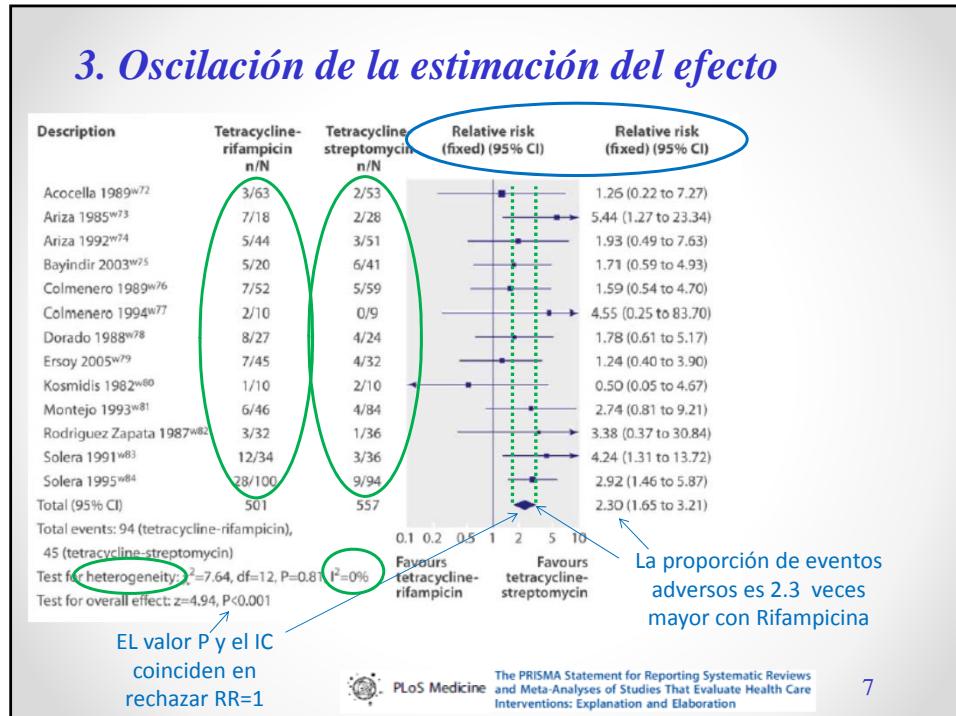
Bioestadística para no estadísticos

4



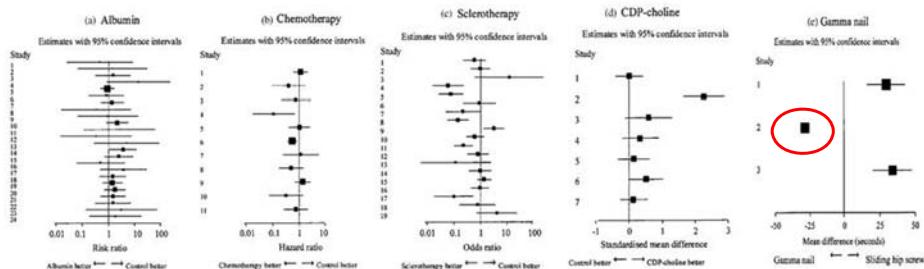
2. Efecto homogéneo





4. Cuantificar la heterogeneidad

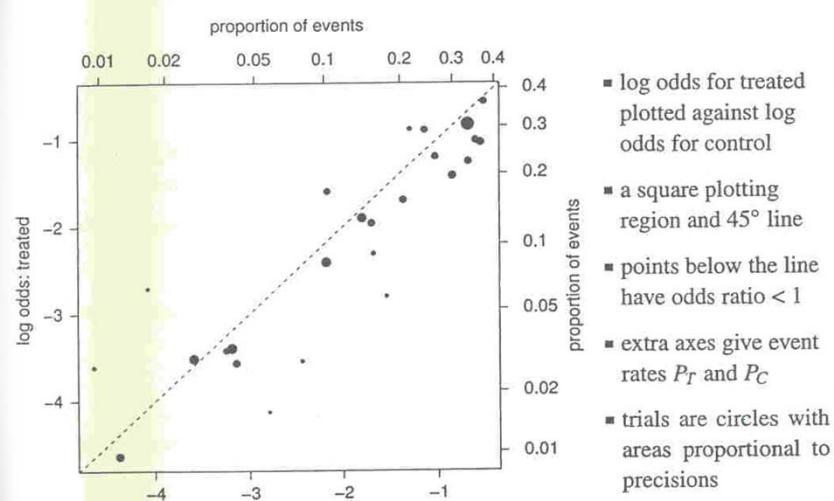
Quantifying heterogeneity in a meta-analysis
 Julian P. T. Higgins*,† and Simon G. Thompson *Statist. Med.* 2002; **21**:1539–1558 (DOI: 10.1002/sim.1186)

 $\hat{\tau}^2 = 0\%$ $\hat{\tau}^2 = 29\%$ $\hat{\tau}^2 = 78\%$ $\hat{\tau}^2 = 85\%$ $\hat{\tau}^2 = 98\%$

Bioestadística
para no estadísticos

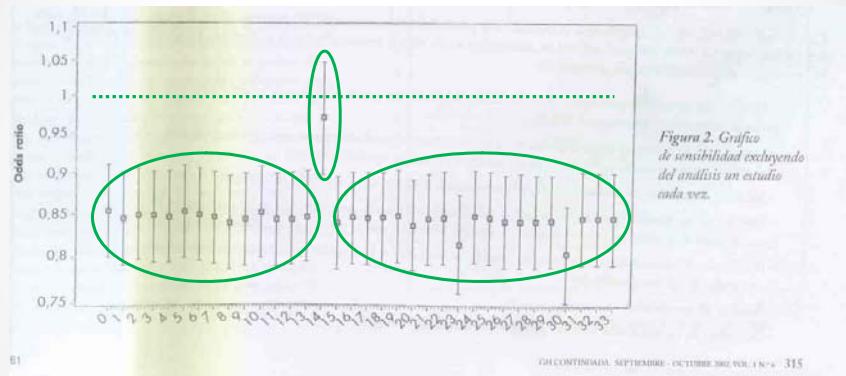
9

L'Abbe plot for the IHD data



10

Sensibilidad



El estudio 14 es necesario para que el conjunto sea significativo

Fuentes de heterogeneidad

(1) Metodológicas:

- Medida estadística del efecto
- Sesgos (baja calidad estudio)
- Otras: P.e., Determinación respuesta



Riesgo de sesgo en los estudios incluidos

The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration

The Cochrane risk of bias tool consists of five items for which there is empirical evidence for their biasing influence on the estimates of an intervention's effectiveness in randomized trials (sequence generation, allocation concealment, blinding, incomplete outcome data, and selective outcome reporting) and a catch-all item called "other sources of bias" [11].

- Sequence generation
- Allocation concealment
- Blinding (*masking*)
- Incomplete Outcome data
- Selective outcome reporting

S
A
B
I
O
S

13

Riesgo de sesgo en el meta-análisis

The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration

Item 15: RISK OF BIAS ACROSS STUDIES. Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).

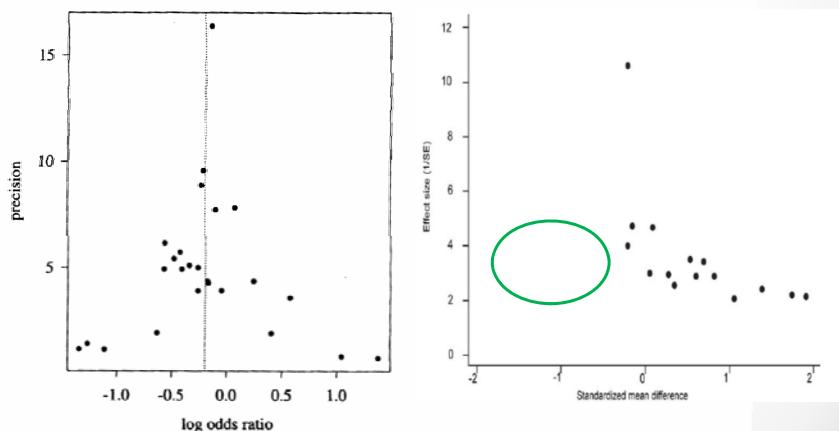
Missing Studies Several empirical investigations have shown that the findings from clinical trials are more likely to be published if the results are statistically significant ($p < 0.05$) than if they are not [125,206,207].

Detection of Missing Information Missing studies may increasingly be identified from trials registries. Evidence of missing outcomes may come from comparison with the study protocol, if available, or by careful examination of published articles [11]. Study publication bias and selective outcome reporting are difficult to exclude or verify from the available results, especially when few studies are available.

14

Gráfico en embudo

Funnel plot of the IHD data



15

Fuentes de heterogeneidad

(2) Fuentes clínicas

El efecto podría variar entre:

- Pacientes
- Clínicos
- Intervencionistas
- Hospitales
- Gravedad previa
- ¿Otras condiciones?
- ¿Edad?
- ¿Género?

Anónimos:

Sugiere incluir nueva variabilidad τ^2

Conocidos:

Sugiere restringir población objetivo

16

Variabilidad τ^2 entre intervencionista.
Sin estudios sobre *mi* intervencionista, el modelo de efectos aleatorios debe considerar ambas incertidumbres σ^2 y τ^2

Outcomes	Number of studies† (comparisons)	Mean effect size‡ at 6 wk to 24 mo (95% CI)	P value
Behavioral and psychological symptoms of dementia	17 (25)	0.34 (0.20 to 0.48)	< 0.01
Caregiver adverse reactions to behavioral and psychological symptoms of dementia	13 (16)	0.15 (0.04 to 0.26)	0.006

*CI defined in Glossary. Number of comparisons exceeds number of studies because some studies had multiple comparisons.

†7 of 23 studies were classified as pseudorandomized.

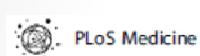
‡Positive effect size indicates positive overall effect of caregiver interventions on outcomes. Effect sizes (difference between treatment and control groups or between pre-and posttreatment assessments) were standardized mean differences (Cohen d) calculated using a random effects model.

17

5. Recomendaciones: (1) Meta-análisis

NO: Borrado secuencial sistemático de estudios

1) Defina a priori una calidad mínima de los estudios.



The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration



Manual Cochrane de revisiones sistemáticas de intervenciones



The GRADE system provides a framework for evaluating how risk of bias, publication bias, imprecision, inconsistency, and indirectness may reduce confidence in estimates of relative effects of interventions on outcomes.

2) Cuantifique heterogeneidad I^2 en su modelado

ya que no requiere exigente premisa $I^2 = 0$

y si $I^2 = 0 \rightarrow$ modelo efectos aleatorios = efectos fijos



18

5. Recomendaciones: (2) Ensayo Clínico

NO: estudiar si un subgrupo es significativo

1) Estudiar interacción sólo para gravedad inicial

Otras variables sólo con hipótesis y potencia previas

2) Seguir CONSORT y guía de la revista destino

www.thelancet.com Vol 365 January 8, 2005

Power & Rules of subgroup analysis: a proposed guideline for design, analysis, interpretation, and reporting
Trial designs
• Subgroup analyses should be defined before starting the trial and should be limited to a small number of clinically important questions.
• Expert input is required in the design of subgroup analyses, and it is important that other members of the team are involved.
• The subgroup analysis should be part of anticipated subgroup effects.
• Effects should be tested at the level of the study.
• The subgroup analysis should be described explicitly as the result of either a hypothesis or a pre-specified analysis plan.
• Definitions for continuous or binary variables should be clearly stated.
• Definitions for categorical variables should be clearly stated.
• The magnitude of effect size should be reported.
• The direction of effect size should be reported.
• Interpretant subgroup interactions are interactions between a subgroup variable and a treatment variable.
• Interactions involving subgroup variables are not simple interactions.
• Adjusting subgroup interactions can account for anticipated subgroup treatment effect interactions, and not simply the overall treatment effect.
• Evaluating treatment effect is likely to be related to randomization, and not to subgroup analysis. The risk associated with randomization is therefore more relevant than the relevance of the treatment effect.
Analyses and reporting
• The subgroup analysis should be reported in the methods section along with details of how and why subgroup analyses were performed.
• Significance of the effect of treatment is determined by the primary outcome, and not by subgroup analysis. This includes subgroup analyses of secondary outcomes.
• Adjusting subgroup interactions should be reported.
• A subgroup analysis should be reported if it is necessary to test only one interaction of subgroup variables for the subgroup (different lengths of follow-up etc).

Clinical Review & Education
How to Use a Subgroup Analysis
Users' Guides to the Medical Literature
JANUARY 8, 2005 • VOLUME 365 • NUMBER 9451 • THE LANCET • www.thelancet.com

Clinical trials, when trying to apply their results to patient care, usually include either patients who probably have the disease or those who do not. In addition, some trials include patients who have different degrees of severity of the disease. These different groups are called subgroups. In subgroup analysis, the investigator tries to determine whether the treatment is effective in all subgroups. If the treatment is effective in all subgroups, it is considered to be effective in all patients. If the treatment is effective in some subgroups but not in others, it is considered to be effective in some patients but not in others. This is called subgroup heterogeneity.

Interpretation
• Report of the significance of the effect of treatment in subgroup analysis is not recommended. The significance of the effect of treatment is determined by the primary outcome, and not by subgroup analysis. This includes subgroup analyses of secondary outcomes.
• Adjusting subgroup interactions should be reported.
• The risk of subgroup analysis is relatively low.
• The first subgroup analysis is not statistically valid and may lead to false-positive results. This is because the power of the test of subgroup treatment effect is often very low. The power of the test of subgroup treatment effect increases as more patients are included in the study.
• The risk of subgroup treatment effect interactions is more acceptable in other studies.
• The risk of subgroup treatment effect interactions is higher in studies where a true interaction exists and results are high.

IN NEW ENGLAND JOURNAL OF MEDICINE

SPECIAL REPORT

Statistics in Medicine — Reporting of Subgroup Analyses in Clinical Trials

Rui Wang, M.S., Stephen W. Lagakos, Ph.D., Jerome H. Wake, Ph.D., David J. Thomas, M.B., B.S., and Jeffrey M. Drazen, M.D.

19



Revisiones PRISMA

Bioestadística
para no estadísticos



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Departament d'Estadística
i Investigació Operativa

Bioestadística
para no estadísticos

6. *Etiología* STROBE (a)



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Departament d'Estadística
i Investigació Operativa



1. Asociación frente a causalidad
2. El odds ratio
3. Efectos confundidos

1. Asociación frente a causalidad

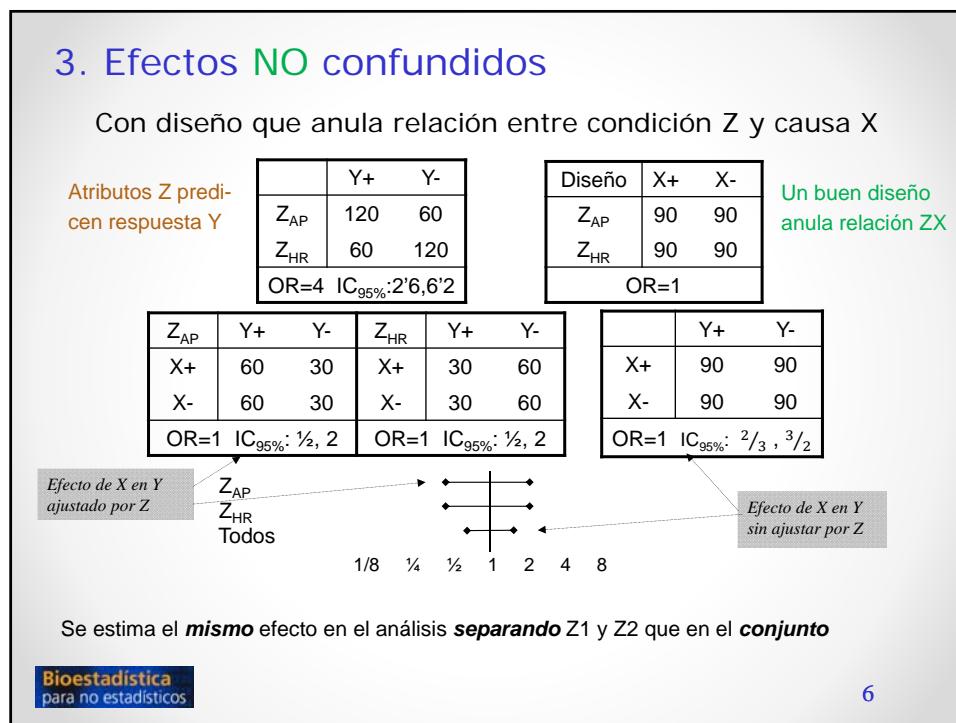
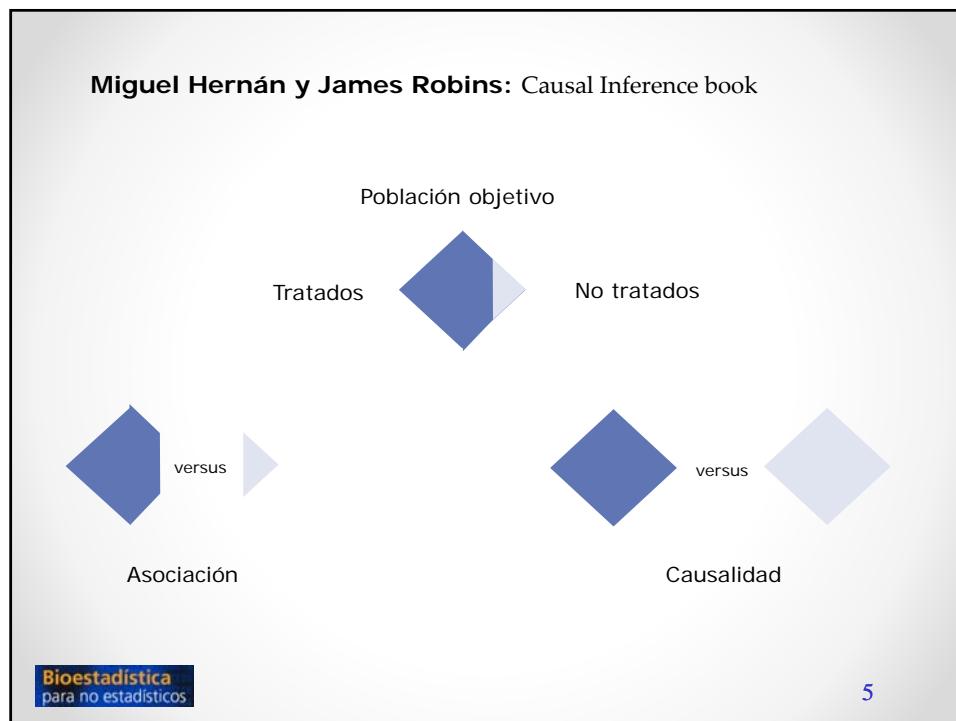
"That correlation is not causation is perhaps the first thing that must be said" (Barnard)

Ejemplos de relaciones entre variables:

Interruptor	→	Luz
Horas estudio	→	Notas
Altura	→	Capacidad mental
Internet	→	Cáncer
X	→	Y
supuesta causa	→	respuesta de interés

Sabiendo que existe relación estadística entre X e Y
en todas puedo predecir Y a partir de X, pero
¿podré cambiar Y, interviniendo sobre X?





El reto de la relación ZX

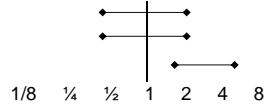
Atributos Z predicen respuesta Y	<table border="1"> <tr><th></th><th>Y+</th><th>Y-</th></tr> <tr><td>Z_{AP}</td><td>120</td><td>60</td></tr> <tr><td>Z_{HR}</td><td>60</td><td>120</td></tr> <tr><td>OR=4 IC_{95%}:2'6,6'2</td><td></td><td></td></tr> </table> <table border="1"> <tr><th></th><th>X+</th><th>X-</th></tr> <tr><td>Z_{AP}</td><td>150</td><td>30</td></tr> <tr><td>Z_{HR}</td><td>30</td><td>150</td></tr> <tr><td>OR=25</td><td></td><td></td></tr> </table>		Y+	Y-	Z _{AP}	120	60	Z _{HR}	60	120	OR=4 IC _{95%} :2'6,6'2				X+	X-	Z _{AP}	150	30	Z _{HR}	30	150	OR=25			Sin diseño, hay relación ZX
	Y+	Y-																								
Z _{AP}	120	60																								
Z _{HR}	60	120																								
OR=4 IC _{95%} :2'6,6'2																										
	X+	X-																								
Z _{AP}	150	30																								
Z _{HR}	30	150																								
OR=25																										

Z _{AP}	Y+	Y-	Z _{HR}	Y+	Y-
X+	100	50	X+	10	20
X-	20	10	X-	50	100
OR=1 IC _{95%} :0'4,2'3			OR=1 IC _{95%} :0'4,2'3		

	Y+	Y-
X+	110	70
X-	70	110
OR=2'5 IC _{95%} :1'6,3'8		

¿Efecto?
NO

Z_{AP}
Z_{HR}
Todos



¿Efecto?
QUIZÁS

Diferentes efectos condicionando por Z que en el global
Si hay relación ZX: Z y X tienen sus efectos confundidos

Bioestadística para no estadísticos

7

Confusión

Para que exista confusión Z debe tener relación con:

X: Z y X están “confundidas”

Y: Z predice Y

Al olvidar Z, “aparece” una relación espúrea

La relación ZY es fija: son las cosas de la vida

La relación ZX puede anularse en el entorno experimental con un buen diseño
pretende observacional ajuste

Bioestadística para no estadísticos

8

OPEN  ACCESS Freely available online

PLOS MEDICINE

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

Jan P. Vandenbroucke¹, Erik von Elm^{2,3}, Douglas G. Altman⁴, Peter C. Gotzsche⁵, Cynthia D. Mulrow⁶, Stuart J. Pocock⁷, Charles Poole⁸, James J. Schlesselman⁹, Matthias Egger^{2,10*} for the STROBE Initiative

Box 5. Confounding

Taking confounders into account is crucial in observational studies, but readers should not assume that analyses adjusted for confounders establish the 'causal part' of an association. Results may still be distorted by residual confounding (the confounding that remains after unsuccessful attempts to control for it [102]), random sampling error, selection bias and information bias (see Box 3).

9



*Etiología
STROBE (a)*

Bioestadística para no estadísticos

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Departament d'Estadística i Investigació Operativa

7. Etiología STROBE (b)



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Departament d'Estadística
i Investigació Operativa



1. Causas frente a efectos
2. Experimentar frente a observar
3. Tipos de diseños
 - Pro y retrospectivo son ambiguos
4. Sesgo de selección
5. Diagramas dirigidos no-cíclicos

Para calentar: **según STROBE**

- 1. El carácter retro o prospectivo de un estudio debe aclararse al inicio**
- 2. Puede utilizar prospectivo y retrospectivo, pero no es obligatorio**
- 3. Evite definir su estudio tan sólo como prospectivo o retrospectivo**

We recommend that authors refrain from simply calling a study 'prospective' or 'retrospective' because these terms are ill defined [29].

3

Significados de prospectivo y retrospectivo

	Pro	Retro
Pregunta sobre	Efectos	Causas
Variable para incluir	Inicial	Final
	Futuros	Pasados
Hipótesis	Previa	Sugerida

1. ¿ Causas o Efectos ?

a) Etiología:

¿Cuál es la causa de este efecto?

Etiology (alternatively aetiology, aetiology) /iːtɪˈɒlədʒi/ is the study of causation, or origination.

The word is derived from the Greek αἰτιολογία, *aītiología*, "giving a reason for" (αἴτια, *aītia*, "cause" /*logia*).^[1]

R: retrospectivo, mirar atrás



WIKIPEDIA
The Free Encyclopedia

b) Farmacología clínica:

¿Cuál es el efecto de esta causa?

Pharmacology (from Greek φάρμακον, *pharmakon*, "poison" in classic Greek; "drug" in modern Greek; and -λογία, *-logia* "study of", "knowledge of") is the branch of medicine and biology concerned with the study of drug action,^[2] where a drug can be broadly defined

P: prospectivo, mirar adelante

Bioestadística
para no estadísticos

5

1. ¿ Causas o Efectos ?

Ejemplo: Asma pre-olímpica en Barcelona:

brotes epidémicos en días y barrios concretos

1^a pregunta: ¿Causas del asma?

Respuesta: interacción entre viento y descarga de soja

R

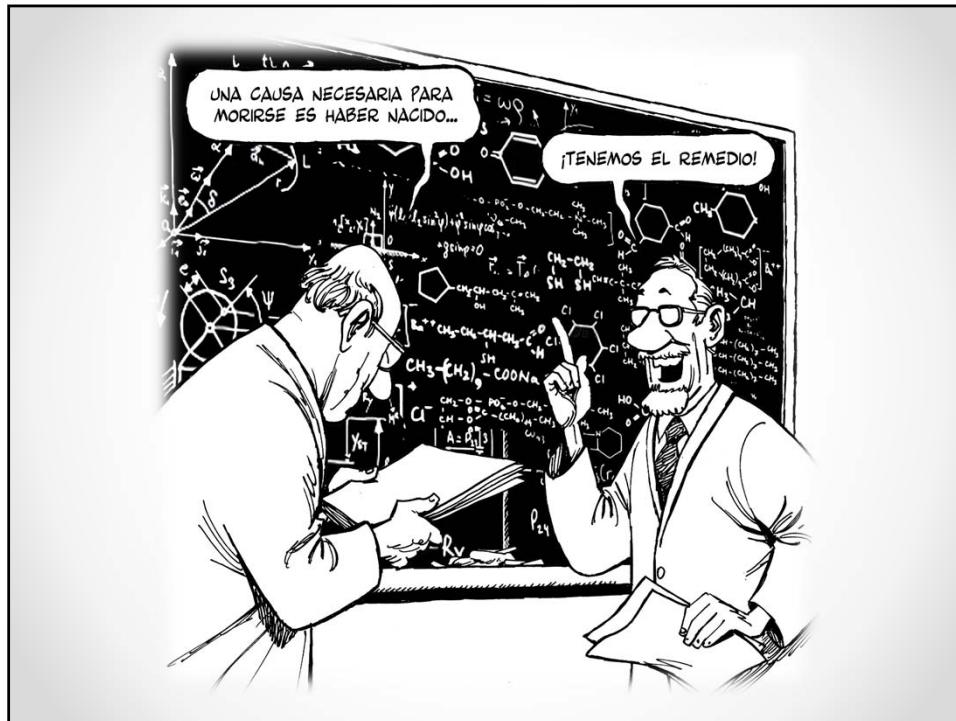
2^a pregunta: ¿Efecto de descargar la soja con protección?

Respuesta: el efecto de lona protectora es no-asma

P

Bioestadística
para no estadísticos

6



2. Tipos de diseños según variables "fijas"

P.e., Ensayo Clínico

		Efecto		Total
Nº total		Y+	Y-	
Causa	X+			
	X-			
Total				1000

P.e., cohorte y transversales

P.e., Caso-Control

		Efecto		Total
Nº total en cada efecto		Y+	Y-	
Causa	X+			500
	X-			500
Total				

		Efecto		Total
Nº total en cada causa		Y+	Y-	
Causa	X+			500
	X-			500
Total				

Mayor precisión para comparar 2 grupos de igual tamaño

3. Experimentar frente a Observar

A) Experimento:

Hacemos (“do”): Asignamos X a las unidades

Ejemplo: **Asignamos 1 o 4** visitas/año a los hipertensos y les seguimos en el tiempo para medir la respuesta.

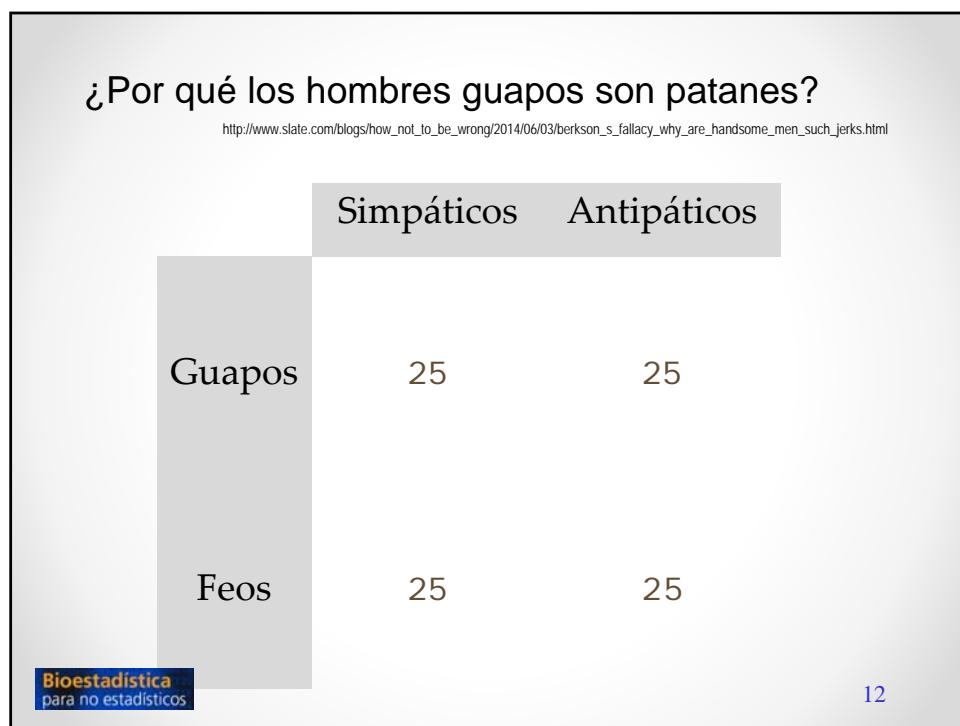
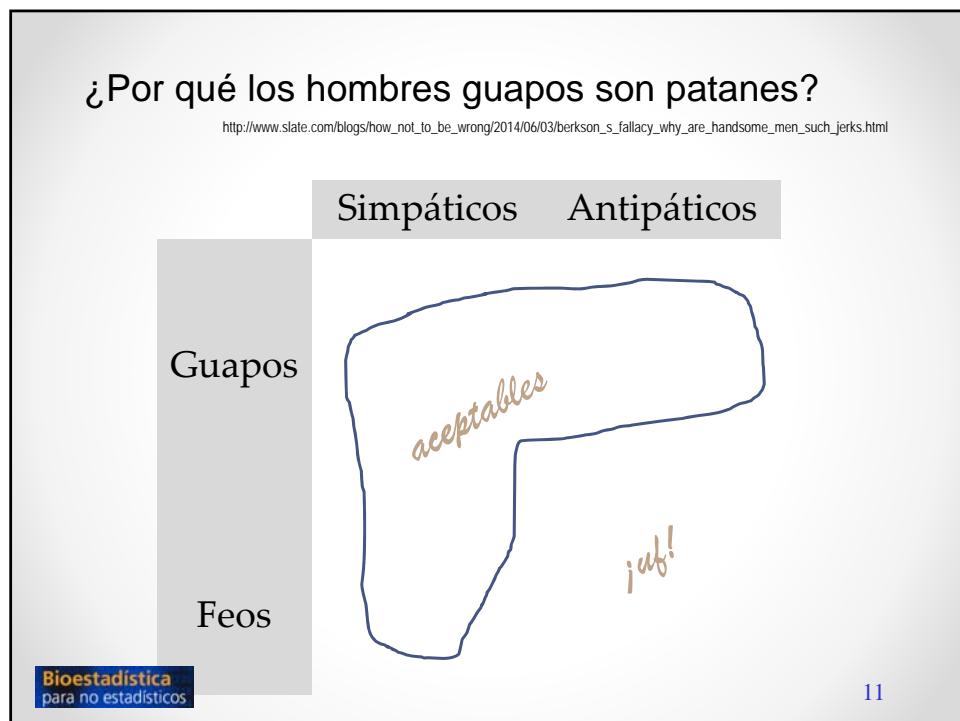
B) Observación:

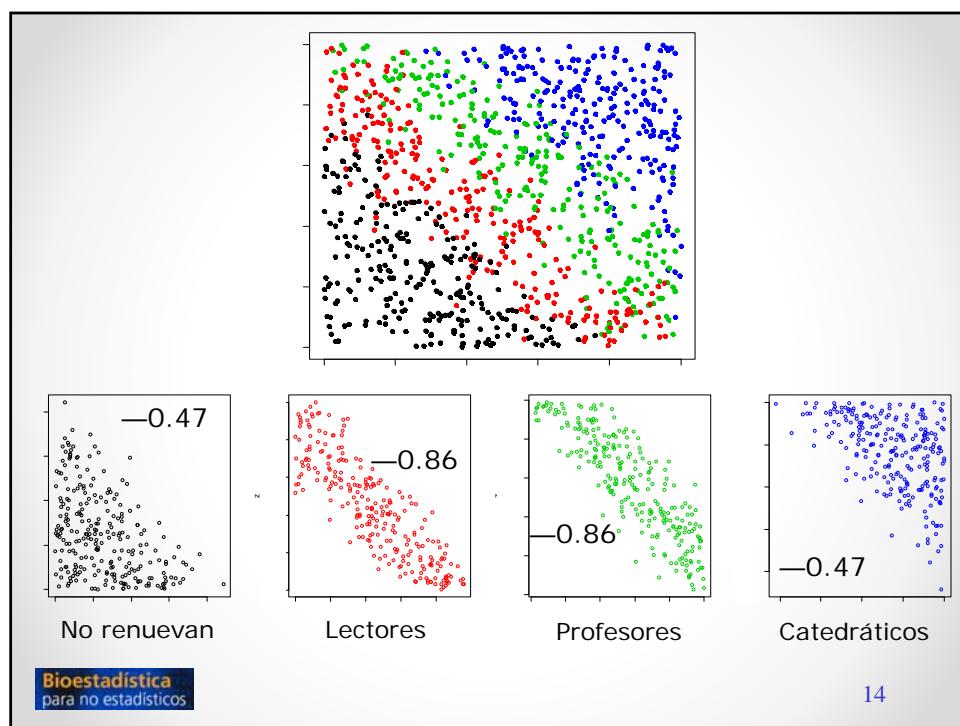
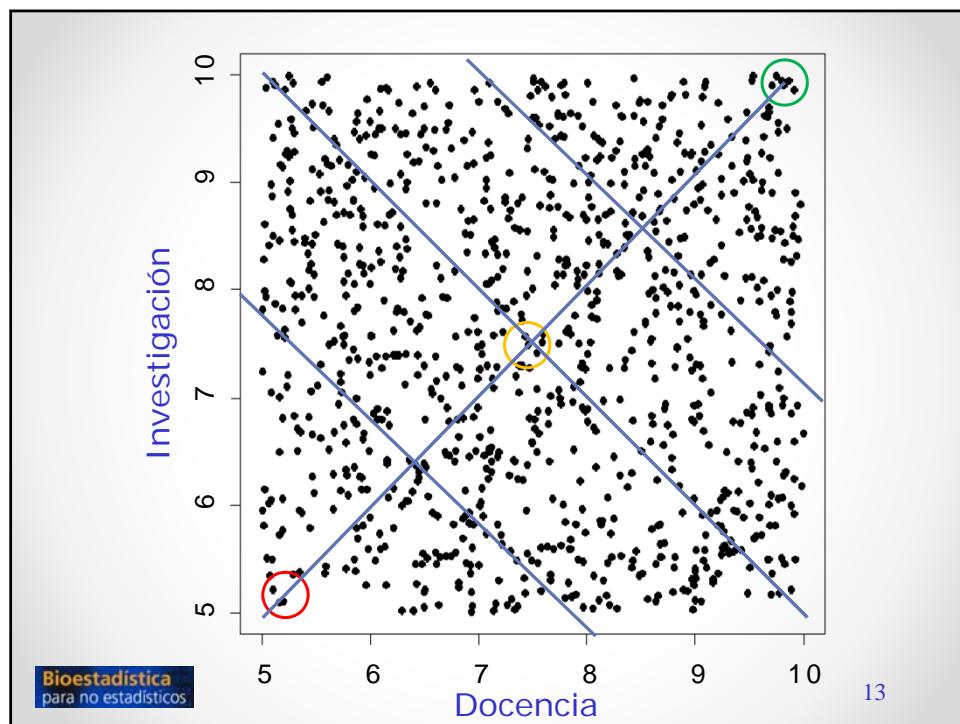
Vemos (“see”): Z en las unidades

Ejemplo: **Miramos** el número de visitas/año y les seguimos en el tiempo para medir la respuesta.

4. Sesgo de selección







Ejemplo: Genética (G), Lípidos (L) y Eventos (E) cardiovasculares

$$L \rightarrow E$$

$$G \rightarrow E$$

	L+	L-
G+	90	90
G-	90	90
OR=1 CI _{95%} =2/3 to 3/2		

En global, Gen G y Lípidos L son independientes.

(E+)	L+	L-	(E-)	L+	L-
G+	80	45	G+	10	45
G-	45	10	G-	45	80
OR=0,4 CI _{95%} =0,18 to 0,86			OR=0,4 CI _{95%} =0,18 to 0,86		

Dentro (E+) y fuera (E-) del hospital, gen (G) previene lípidos (L)

Bioestadística para no estadísticos

15

5. Gráficos dirigidos no-cíclicos

DAGs = Directed Acyclic Graphs

Ejemplo: Presión Arterial

```

graph LR
    Basal --> Tratamiento
    Tratamiento --> Respuesta
  
```

Bioestadística para no estadísticos

16

Causa común implica asociación



Terminología:

Causal: Llevarlo no tiene efecto en Cáncer

Descriptiva: Llevarlo y Cáncer están asociadas

Fumar y Llevarlo tienen **confundidos** sus efectos en cáncer

Condicionar:



Terminología:

Causal: Llevarlo no tiene efecto en Cáncer

Descriptiva: Llevarlo y Cáncer son independientes condicionado por Z

Condicionar **soluciona** efectos **confundidos**

Efecto común



Terminología:

Causal:

Gen no tiene efecto en **Entorno**

Descriptiva:

Gen y **Entorno** son independientes

Condicionar por efecto común



Terminología:

Causal:

Gen no tiene efecto en **Entorno**

Descriptiva:

Gen y **Entorno** están relacionadas
condicionado al efecto común **Cáncer**

Sesgo de selección

EL MÉDICO ME DIJO QUE
LLEVAR TABACO ENCIMA
NO TIENE BUEN
PRONÓSTICO

Y DESDE ENTONCES
SOLO FUMO
DE GORRA

Etiología
STROBE (b)

Bioestadística para no estadísticos

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Departament d'Estadística
i Investigació Operativa

Bioestadística
para no estadísticos

¿situación **actual** del paciente? ¿situación **futura** del paciente?

8. Diagnóstico y pronóstico

STARD

Indicador=1 variable

UPC UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH
Departament d'Estadística i Investigació Operativa



1. Variables
2. Diseños
3. Sensibilidad y Especificidad
4. Valores predictivos
5. Riesgos de sesgo

Bioestadística
para no estadísticos

2

1. Variables

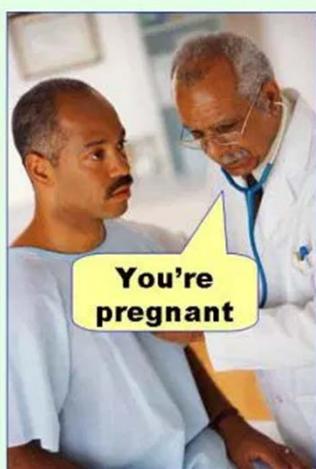
Situación **real** del paciente (*gold standard*)

- Enfermo
- Sano

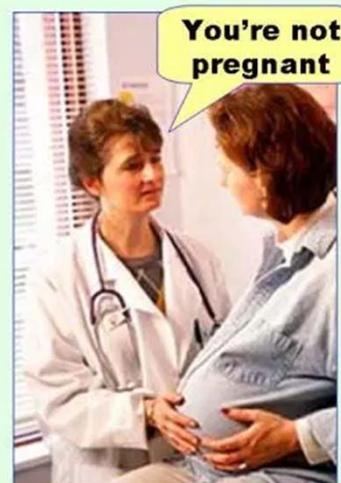
Resultado del **índicador**

- Positivo
- Negativo

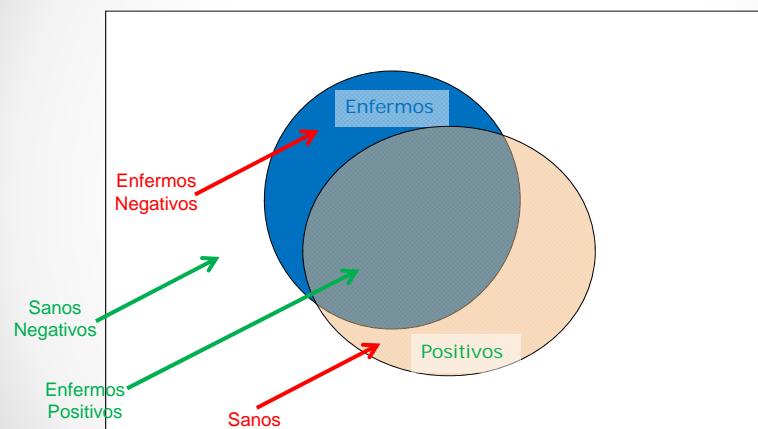
(false positive)



(false negative)



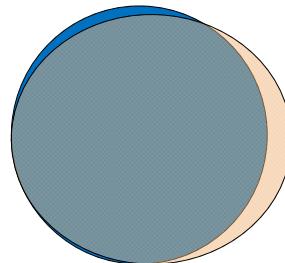
Resultados



Bioestadística
para no estadísticos

5

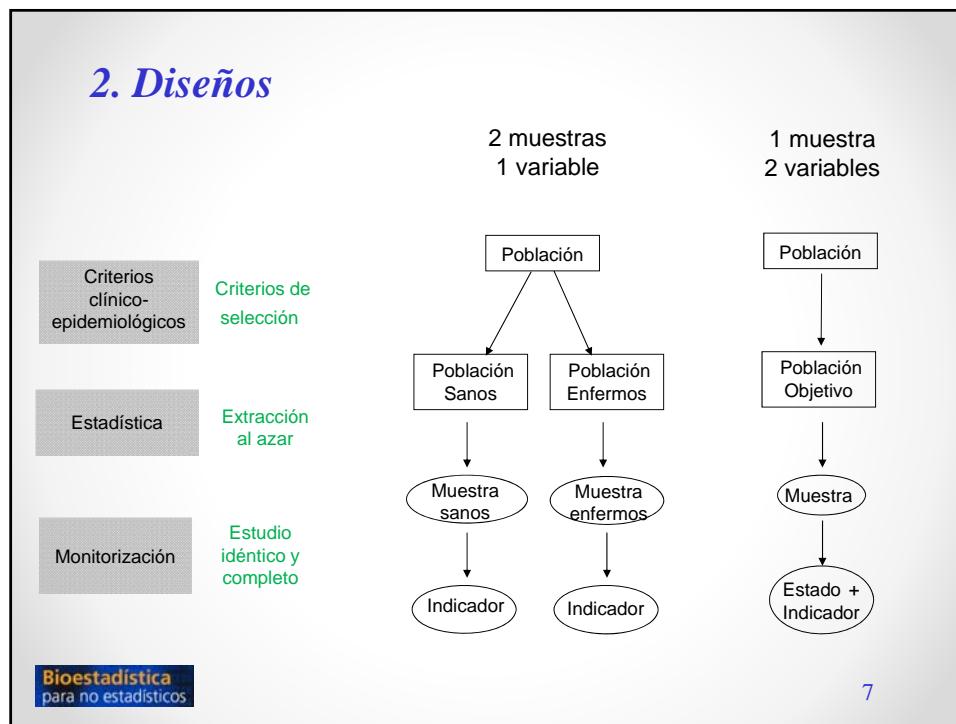
Buena clasificación



Bioestadística
para no estadísticos

6

2. Diseños



		+	-	
2 muestras 1 variable	Enfermos	90	10	100
	Sanos	30	70	100
	120	80	200	

		+	-	
1 muestra 2 variables	Enfermos	9	1	10
	Sanos	27	63	90
		36	64	100

8

Probabilidad condicionada

1 muestra
2 variables

Pelo	Largo	Corto	
Chicas	30	30	60
Chicos	10	30	40
	40	60	100

Sabiendo que es chica,
¿lleva el pelo largo?

Pelo	Largo	Corto	
Chicas	30	30	60

$$P(Largo|Chica) = \frac{30}{60} = 0.5$$

Sabiendo que lleva el
pelo largo, ¿es chica?

Pelo	Largo		
Chicas	30		
Chicos	10		
	40		

$$P(Chica|Largo) = \frac{30}{40} = 0.75$$

9

3. Sensibilidad y Especificidad

Sensibilidad: "Tendencia" de los enfermos a dar positivo

Especificidad: "Tendencia" de los sanos a dar negativo

1 muestra
2 variables

	+	-	
E=Enfermos	9	1	10
S=Sanos	27	63	90
	36	64	100

$$\text{Sensibilidad} = P(+|E) = \frac{9}{10} = 0.9 = 90\%$$

$$\text{Especificidad} = P(-|S) = \frac{63}{90} = 0.7 = 70\%$$

4. Valores Predictivos = VP

VP de un +: Proporción de positivos que están enfermos

VP de un —: Proporción de negativos que están sanos

	+	-	
E=Enfermos	9	1	10
S=Sanos	27	63	90
	36	64	100

1 muestra
2 variables

$$VP+ = P(E|+) = \frac{9}{36} = 0.25 = 25\%$$

$$VP- = P(S|-) = \frac{63}{64} \approx 0.98 = 98\%$$

Diferencias según diseños

1 muestra y 2 variables

	+	-	
E	9	1	10
S	27	63	90
	36	64	100

2 muestras y 1 variable

	+	-	
E	90	10	?
S	30	70	100
	120	80	200

$$Sens = P(+|E) = \frac{9}{10} = 90\%$$

$$Esp = P(-|S) = \frac{63}{90} = 70\%$$

$$VP+ = P(E|+) = \frac{9}{36} = 25\%$$

$$VP- = P(S|-) = \frac{63}{64} \approx 98\%$$

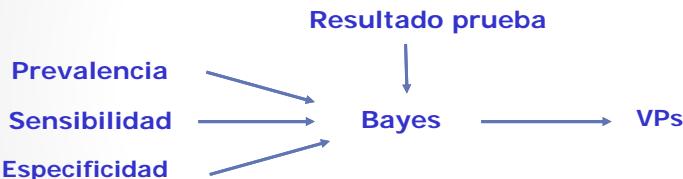
$$Sens = P(+|E) = \frac{90}{100} = 90\%$$

$$Esp = P(-|S) = \frac{70}{100} = 70\%$$

~~$$VP+ = P(E|+) = \frac{90}{120} = 75\%$$~~

~~$$VP- = P(S|-) = \frac{70}{80} \approx 87.5\%$$~~

Obtención de los VP

$$\text{Expectativa previa} + \text{Información prueba} = \text{Expectativa posterior}$$


Resultado prueba

Prevalencia
Sensibilidad
Especificidad

Bayes

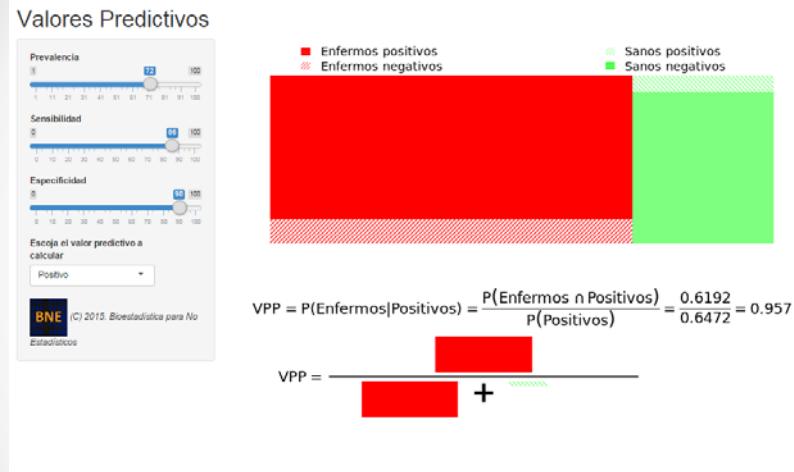
VPs

<http://shiny-eio.upc.edu/bne/VPs/>

Bioestadística para no estadísticos

13

Valores Predictivos



Prevalencia
Sensibilidad
Especificidad

Enfermos positivos
Enfermos negativos
Sanos positivos
Sanos negativos

Escoga el valor predictivo a calcular
Positivo

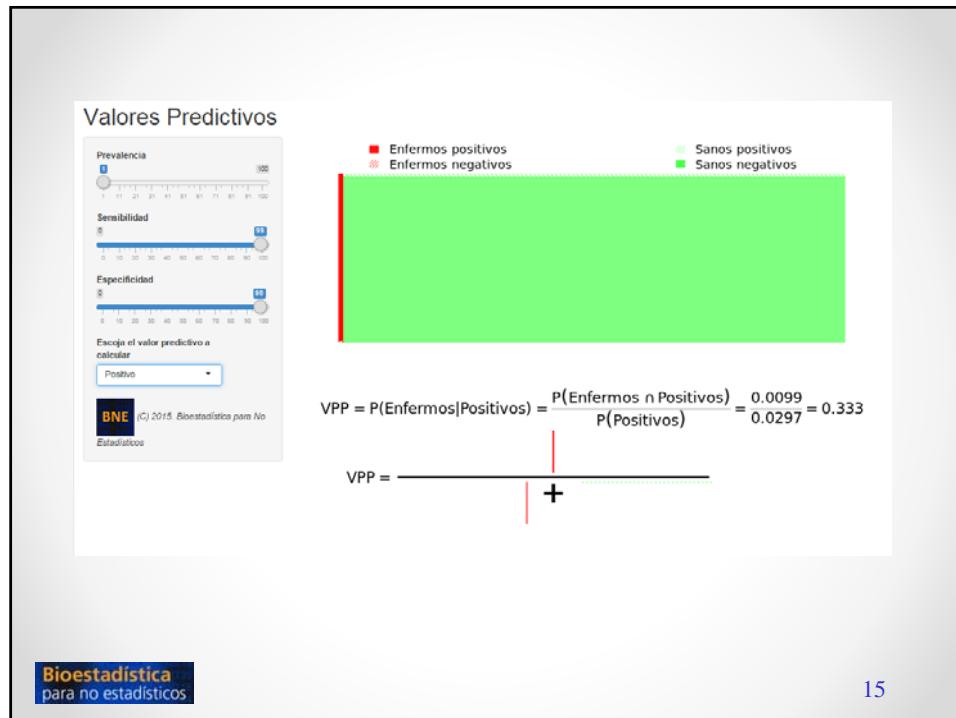
BNE © 2015. Bioestadística para No Estadísticos

$$\text{VPP} = \frac{P(\text{Enfermos} | \text{Positivos})}{P(\text{Positivos})} = \frac{P(\text{Enfermos} \cap \text{Positivos})}{P(\text{Positivos})} = \frac{0.6192}{0.6472} = 0.957$$

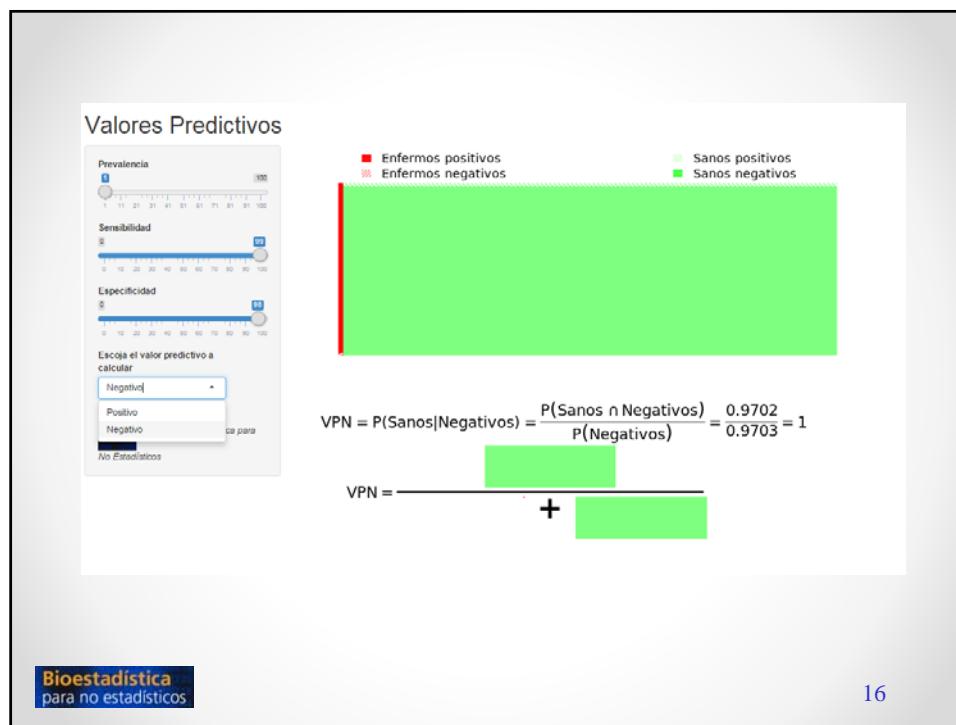
$$\text{VPP} = \frac{\text{Enfermos positivos}}{\text{Enfermos positivos} + \text{Sanos positivos}}$$

Bioestadística para no estadísticos

14

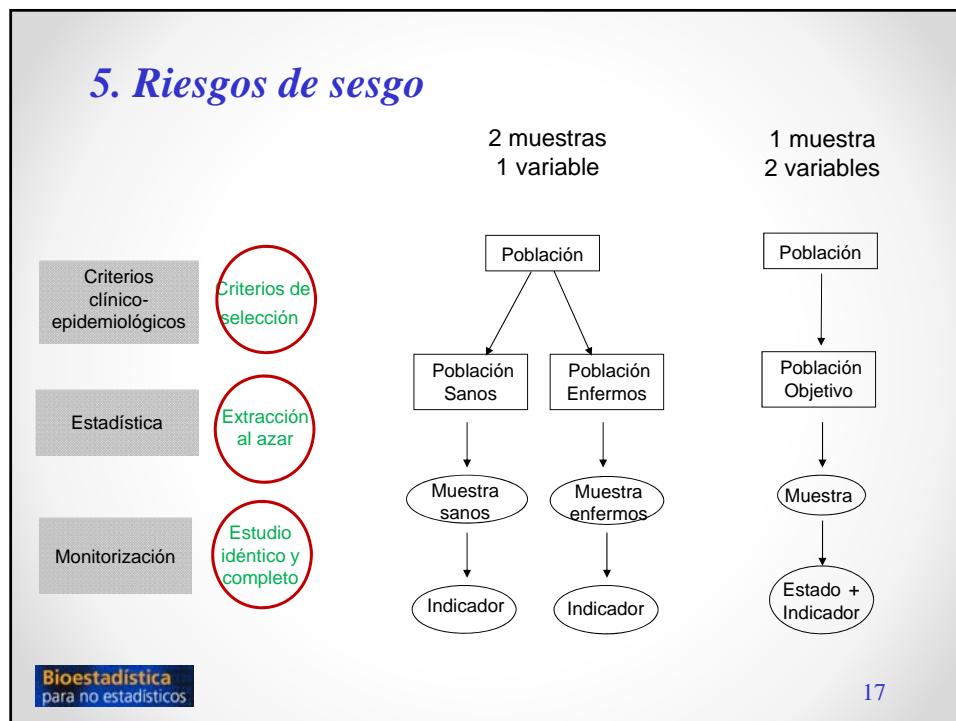


15

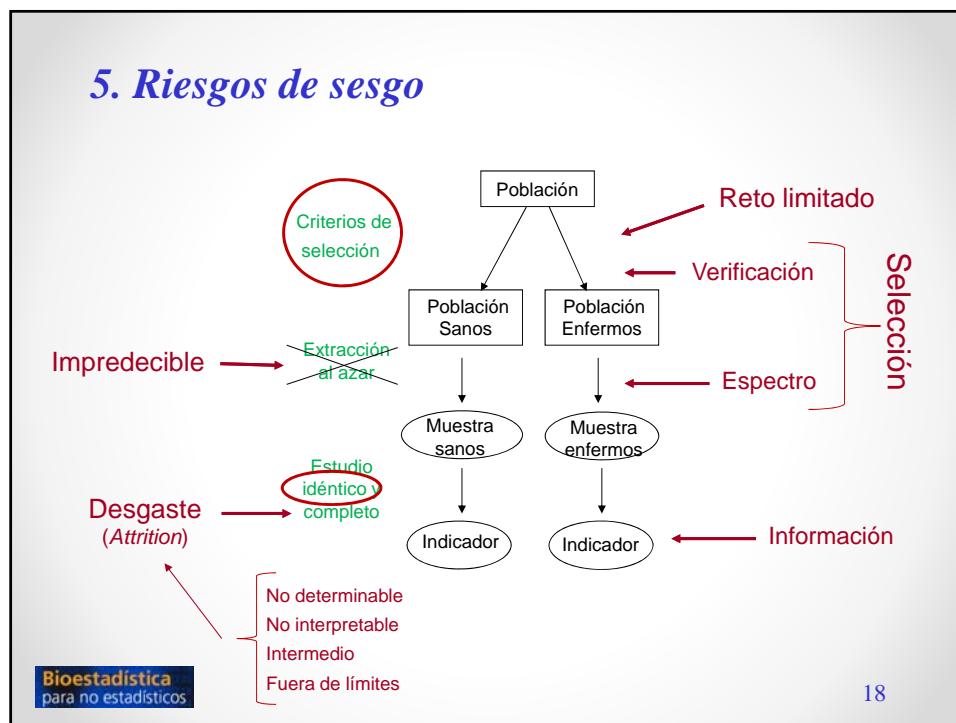


16

5. Riesgos de sesgo



5. Riesgos de sesgo





A black and white cartoon illustration. A man with a cast on his right leg is falling from a plane, looking up at another man who has already parachuted away. The man falling says, "NUNCA MÁS ME APUNTO A UN ENSAYO CLÍNICO SOBRE PARACÁIDAS!" (Never again will I sign up for a clinical trial about parachutes!).

*Diagnóstico y
Pronóstico
STARD*

Bioestadística
para no estadísticos

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Departament d'Estadística
i Investigació Operativa

Bioestadística
para no estadísticos

9. Diagnóstico y pronóstico TRIPOD

Construcción 1 escala



UNIVERSITAT POLITÈCNICA DE CATALUNYA

BARCELONATECH

Departament d'Estadística
i Investigació Operativa



1. Modelo Predictivo
2. Razón de tasas (HR: *Hazard Ratio*)
3. Capacidad Predictiva
4. Sobre-ajuste
5. Aprender, Valorar y Actualizar
6. Discriminar y Calibrar

1. Modelo Predictivo

1. Type of Model

Examples

We used the Cox proportional hazards model in the derivation dataset to estimate the coefficients associated with each potential risk factor [predictor] for the first ever recorded diagnosis of cardiovascular disease for men and women separately (278). [Prognosis]

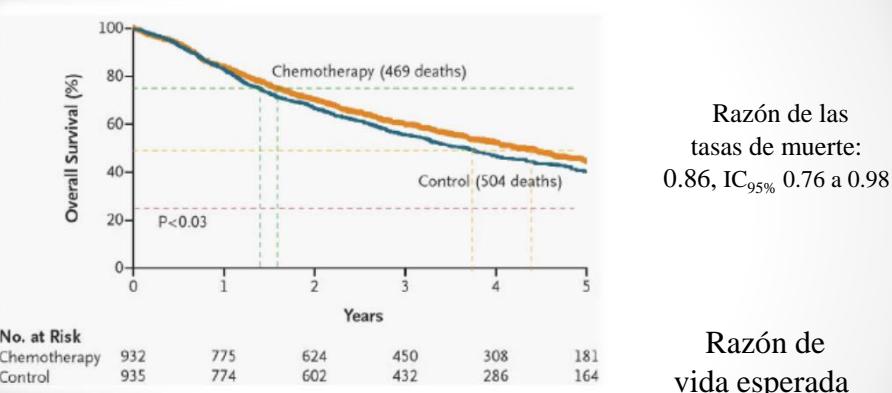
All clinical and laboratory predictors were included in a multivariable logistic regression model (outcome: bacterial pneumonia) (279). [Diagnosis]



Bioestadística
para no estadísticos

3

2. HR: Hazard ratio o razón de tasas



A hazard ratio was estimated by a ratio of median survival times, but with considerable uncertainty

Jordi Cortés^{a,b}, José A. González^a, Michael J. Campbell^b, Erik Cobo^a

4

BMJ

Research

Cite this article as: BMJ. doi:10.1136/bmj-2006-38895-5543177C [published 23 October 2006]

Ratio measures in leading medical journals: structured review of accessibility of underlying absolute risks
Lisa M Schwartz, Steven Woloshin, Evan L Dowlin, H Gilbert Welch

Synonyms used to describe various ratio measures

- Ratio of risks**
Relative risk (or RR)*; post-trial relative risk, risk ratio; "times more likely"; and unadjusted risk ratio
- Ratio of risks: adjusted**
Adjusted relative risk*; multivariable adjusted relative risk; multivariable analysis relative risk; multivariable relative risk; multivariate relative risk; after adjustment risk ratio; and adjusted risk ratio
- One minus the ratio of risks**
Relative risk reduction*; reduction in relative risk*; risk reduction; proportional reduction*; and per cent reduction
- Ratio of rates**
Rate ratio*; incidence rate ratio; and incidence density ratio
- Ratio of rates: adjusted**
Adjusted rate ratio* and standardised incidence ratio
- Ratio of odds**
Odds ratio*; unadjusted odds ratio; and exposure odds ratio
- Ratio of odds: adjusted**
Adjusted odds ratio*; multivariate odds ratio; and multivariate adjusted odds ratio

*Most commonly used language

RR ≈ HR ≈ OR Hazard rate ratio and prospective epidemiological studies
M.J. Symons*, D.T. Moore

RR < HR < OR

Proporciones y Probabilidades

Tasas

Momios

5

1. Modelo Predictivo

1. Type of Model

Examples

We used the Cox proportional hazards model in the derivation dataset to estimate the coefficients associated with each potential risk factor [predictor] for the first ever recorded diagnosis of cardiovascular disease for men and women separately (278). [Prognosis]

All clinical and laboratory predictors were included in a multivariable logistic regression model (outcome: bacterial pneumonia) (279). [Diagnosis]

Intercept and Predictors	$\beta\pm$	Odds Ratio	95% CI
Intercept	-3.66		
Traditional baker	0.67	2.2	1.2-3.9
Nasocconjunctival symptoms in the past 12 mo	0.72	2.3	1.2-4.5
Asthma symptoms in the past 12 mo	0.63	2.0	0.9-4.4
Shortness of breath and wheeze	0.61	2.3	1.3-3.8
Work-related upper respiratory symptoms	0.47	1.7	0.9-3.1
Work-related lower respiratory symptoms	0.61	2.2	1.1-4.4
ROC area (95% CI)	0.75 (0.71-0.81)		

ROC = receiver-operating characteristic.
From reference 319.

*The predicted probability of wheat sensitization can be calculated using the formula: $P(\text{sensitization}) = 1/(1 + e^{-(-3.66 + \text{traditional baker} \times 0.67 + \text{nasocconjunctival symptoms in the past 12 mo} \times 0.72 + \text{asthma symptoms in the past 12 mo} \times 0.63 + \text{shortness of breath and wheeze} \times 0.61 + \text{work-related upper respiratory symptoms} \times 0.47 + \text{work-related lower respiratory symptoms} \times 0.61)})$. Predictor value is one when present and zero when absent.

† Regression coefficient multiplied with a shrinkage factor (obtained from the bootstrapping procedure) of 0.89.

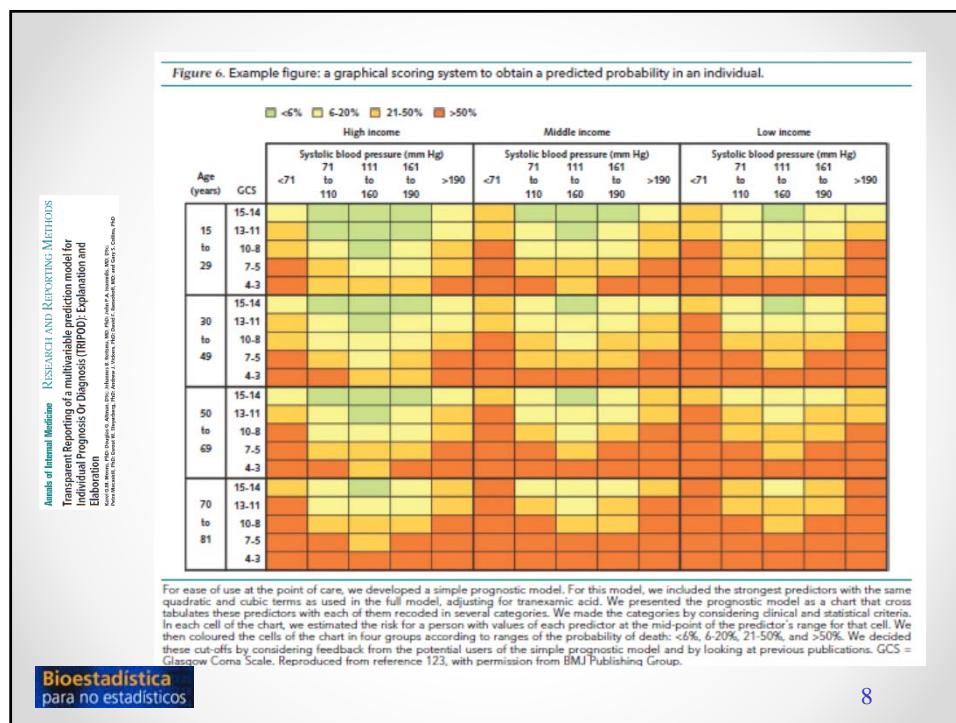
Annals of Internal Medicine RESEARCH AND REPORTING METHODS

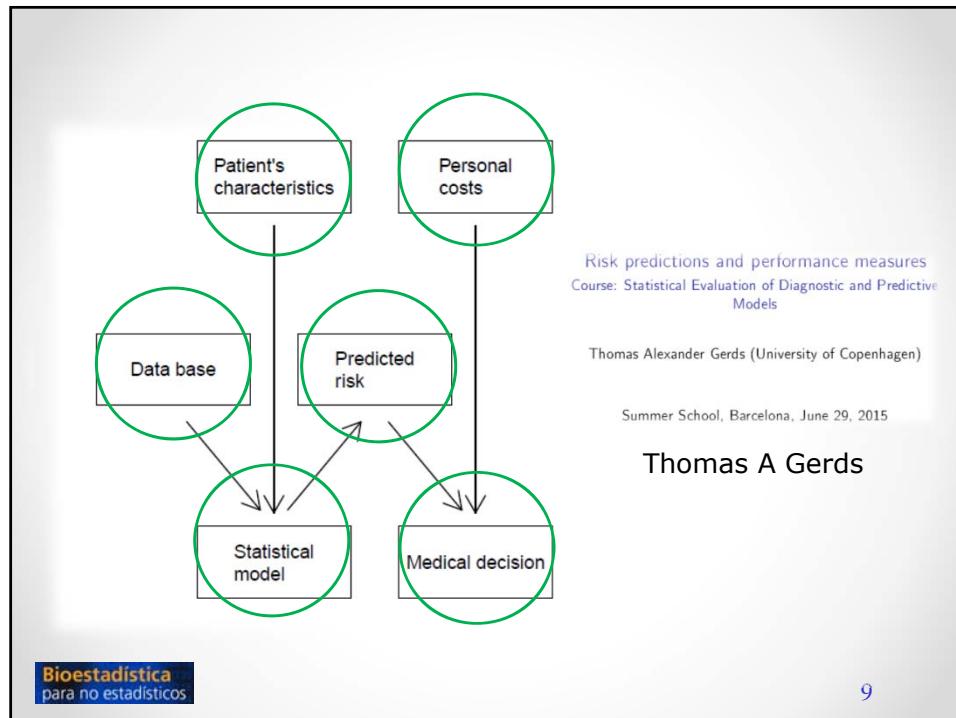
Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration

Karel G.M. Moons, PhD; Douglas G. Altman, DSc; Johannes B. Reitsma, MD, PhD; John P.A. Ioannidis, MD, DSc; Petra Macaskill, PhD; Ewout W. Steyerberg, PhD; Andrew J. Vickers, PhD; David F. Ransohoff, MD; and Gary S. Collins, PhD

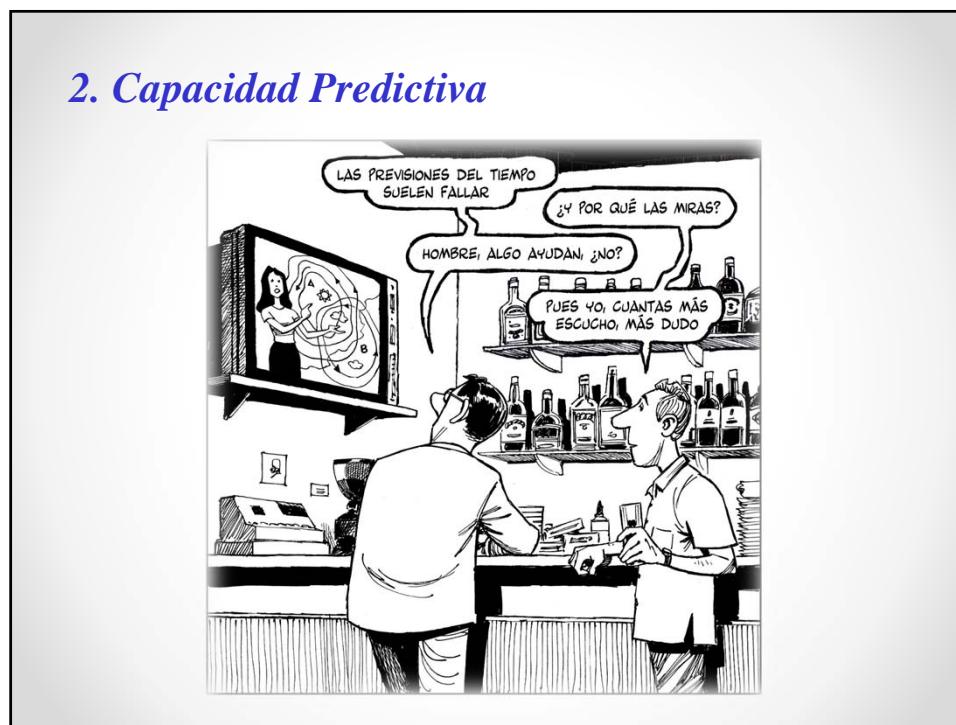
Bioestadística para no estadísticos

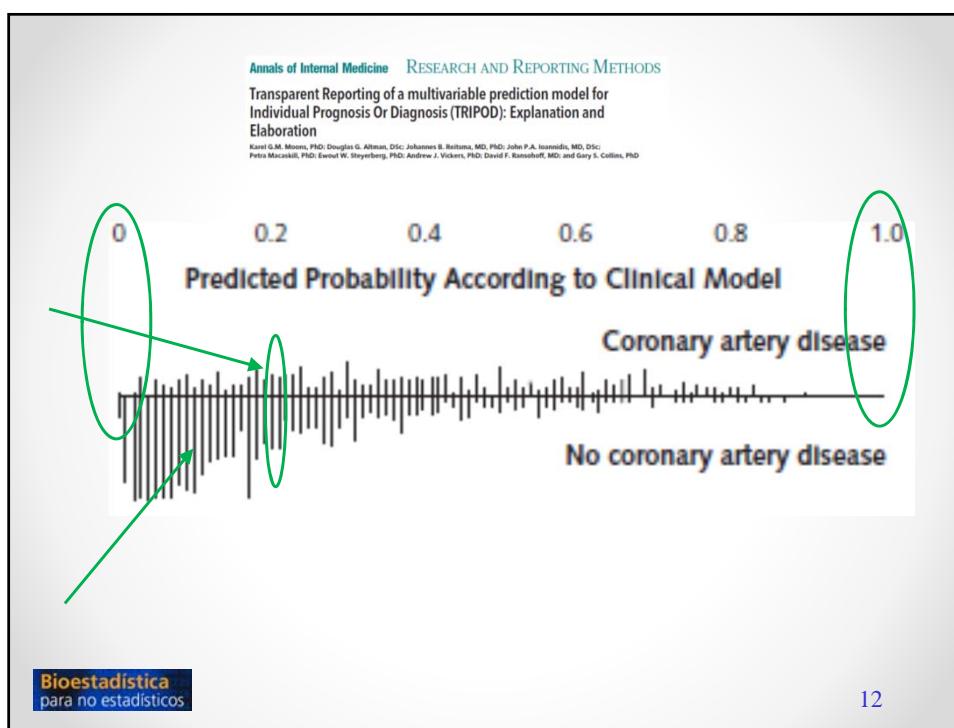
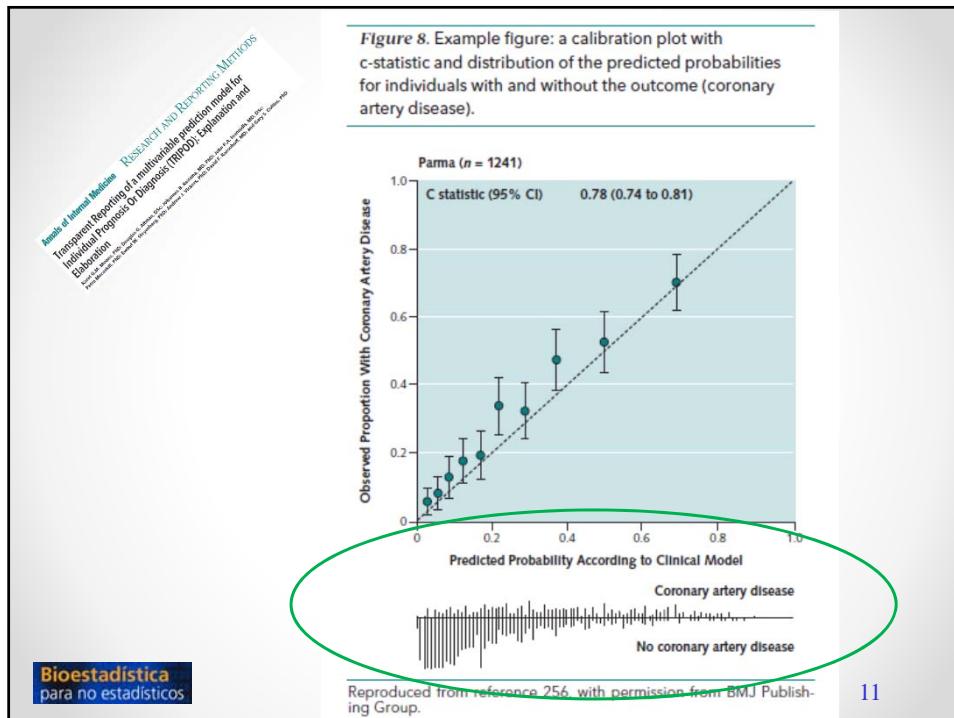
6





9





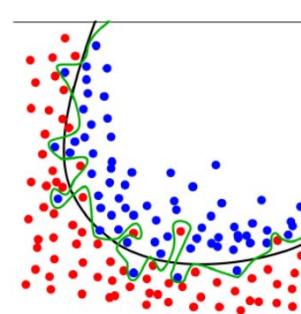
Risk predictions and performance measures
Course: Statistical Evaluation of Diagnostic and Predictive Models

3. Sobre-ajuste

Thomas Alexander Gerdts (University of Copenhagen)

Overfitting

Summer School, Barcelona, June 29, 2015



Sources of overfitting

- ▶ unshranked parameter estimates
- ▶ variable selection
- ▶ data-driven cut-offs to group continuous variables
- ▶ relatively many parameters with respect to amount information in data

Bioestadística para no estadísticos

13

4. Aprender, Valorar y Actualizar

Figure 1. Types of prediction model studies covered by the TRIPOD statement.

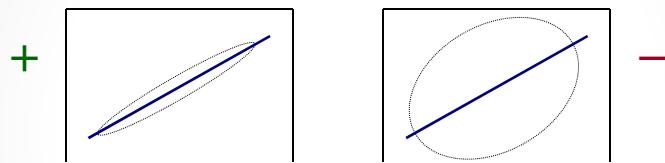


Bioestadística para no estadísticos

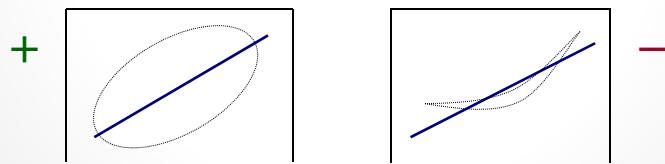
14

5. Discriminar y Calibrar

1) Discriminar: Clasificar bien, reducir incertidumbre



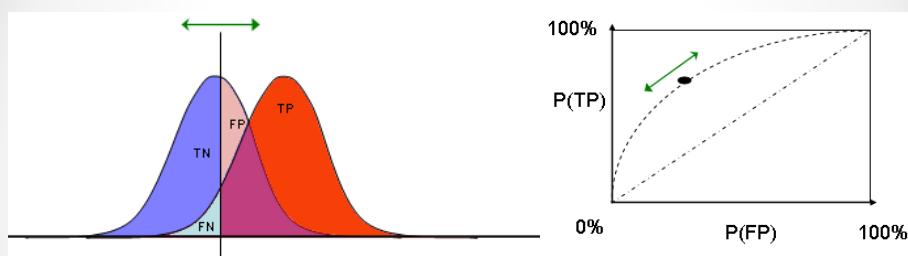
2) Calibrar: correspondencia entre modelo y datos



15

Discriminar

Estadístico c o área bajo la curva ROC



BUSQUE y VISITE: "youtube + UPC + BNE + ROC"

Proporción de parejas 'enfermo-sano' en las que el 'enfermo' tiene un valor mayor que el 'sano'.

Probabilidad que un enfermo tenga mayor valor que un sano.

Discriminar

NIH Public Access
Author Manuscript
Published in final edited form as:
Epidemiology. 2010 January ; 21(1): 128–138. doi:10.1097/EDE.0b013e3181c30fb2.

Assessing the performance of prediction models: a framework for some traditional and novel measures

Ewout W. Steyerberg¹, Andrew J. Vickers², Nancy R. Cook³, Thomas Gerdts⁴, Mithat Gonen², Nancy Obuchowski⁵, Michael J. Pencina³, and Michael W. Kattan⁵

Bioestadística para no estadísticos

17

Calibrar

Annals of Internal Medicine | ARTICLE

Development and Validation of a Multifactorial Risk Index for Predicting Postoperative Pneumonia after Major Noncardiac Surgery

Abbas M. Arasteh, MD, MPH; Shahn F. Khan, MD; William G. Henderson, PhD; and Jennifer Daley, MD, for the Participants in the National Veterans Affairs Surgical Quality Improvement Program

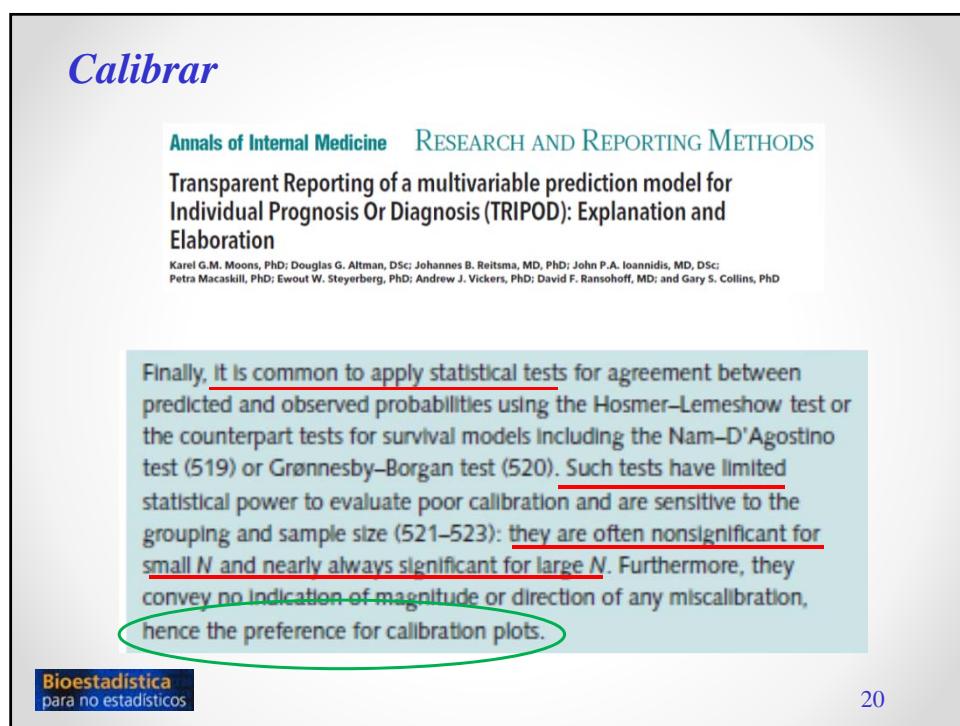
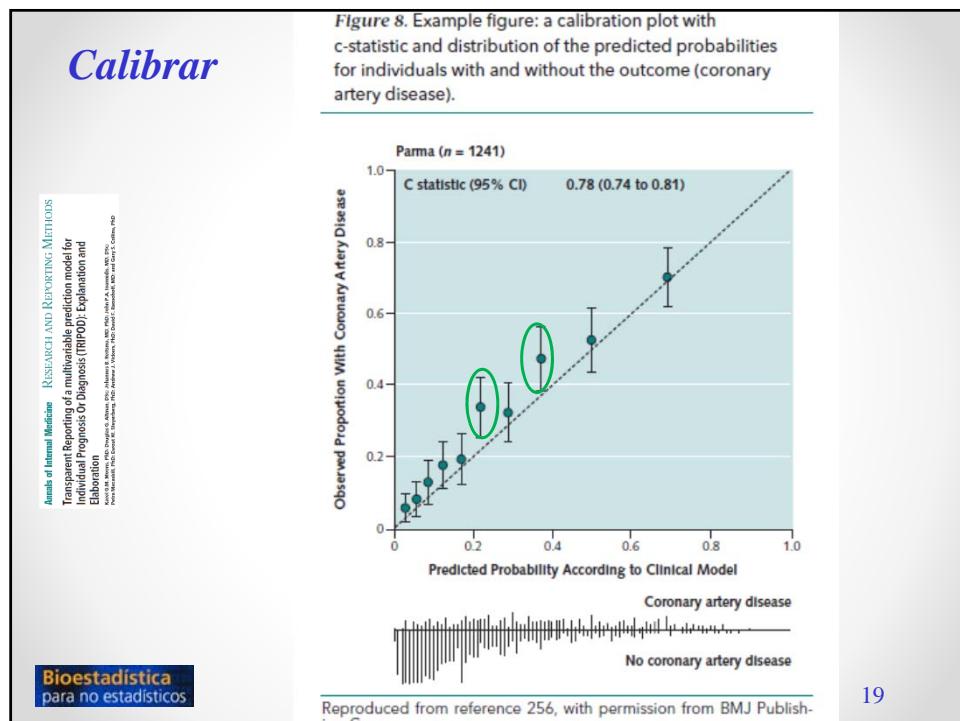
Table 5. Distribution of Postoperative Pneumonia Risk Index Scores in Patients in the Development and Validation Cohorts*

Variable	1 (0–15 Points)	2 (16–25 Points)	3 (26–40 Points)	4 (41–55 Points)	5 (> 55 Points)
Development cohort patients, n (%)	69 333 (43)	55 757 (35)	32 103 (20)	3517 (2)	95 (0.1)
Average predicted probability of postoperative pneumonia in development cohort patients (95% CI) % †	0.24 (0.24–0.25)	1.20 (1.19–1.20)	4.0 (3.98–4.01)	9.4 (9.34–9.42)	15.3 (15.1–15.5)
Rate of postoperative pneumonia in development cohort patients, % ‡	0.24	1.19	4.6	9.4	15.8
Rate of postoperative pneumonia in validation cohort patients, % §	0.24	1.18	4.6	10.8	15.9

* Patients in the development cohort were enrolled from 1 September 1997 to 31 August 1999; patients in the validation cohort were enrolled from 1 September 1995 to 31 August 1997.
† Based on logistic regression model; 95% CI calculated as average predicted probability \pm (1.96 \times SE).

Bioestadística para no estadísticos

18



nature International weekly journal of science

612 | NATURE | VOL 505 | 30 JANUARY 2014

NIH plans to enhance reproducibility

Francis S. Collins and Lawrence A. Tabak discuss initiatives that the US National Institutes of Health is exploring to restore the self-correcting nature of preclinical research.

PROPOSED NIH ACTIONS

As a funding agency, the NIH is deeply concerned about this problem. Because poor training is probably responsible for at least some of the challenges, the NIH is developing a training module on enhancing reproducibility and transparency of research findings, with an emphasis on good experimental design. This will be incorporated into the mandatory training

Several of the NIH's institutes and centres are also testing the use of a checklist to ensure a more systematic evaluation of grant applications. Reviewers are reminded to check, for example, that appropriate experimental design features have been

Diagnóstico y Pronóstico TRIPOD

Bioestadística para no estadísticos

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Departament d'Estadística i Investigació Operativa