# Trobades amb l'Estadística: pensar-la abans de començar

Santiago Pérez-Hoyos Unitat d'Estadística i Bioinformàtica(UEB) Vall d'Hebron Institut de Recerca (VHIR) santi.perezhoyos@vhir.org

# Objectiu

Donar algunes pinzellades dels principals usos, abusos, errors estadístics i les maneres de resoldre'ls a l'hora de publicar articles

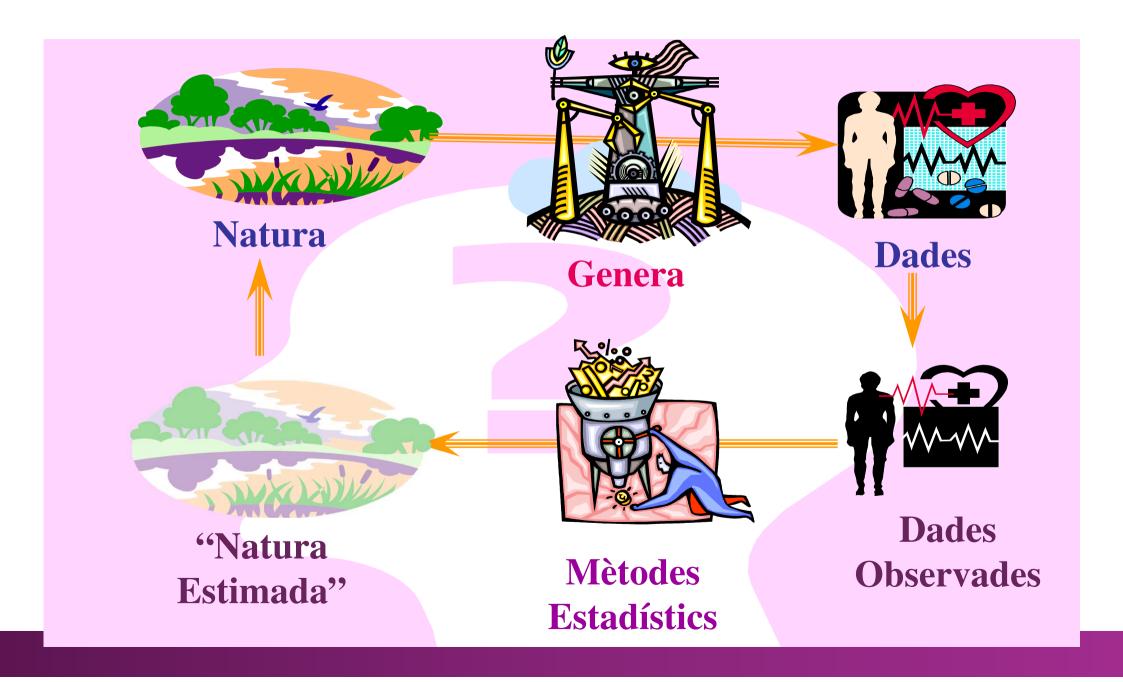
# Perquè trobar-se amb l'Estadística Per evitar el que deia Fisher ...



"Hiring a statistician after de data has been collected is like hiring a physician when a patient is in the morgue: He might be able to tell you what went wrong, but he is unlikely to be able to fix it"

R.A. Fisher

## Com treballa l'estadística



# Què fa l'estadística

Descriu

Anàlisi exploratori de dades. Estimació

Relaciona o Modelització estadística

## Objectius

Preguntes Pacients Objectius Medicina Objectius Ciència

Atemporal / Ara

Què em passa?

Diagnòstic

Classificació

Estudis de seguiment

Què em passarà?

Pronòstic

Predicció

Perquè em passa?

Etiologia

Causes/Epidemiologia

Què em pot a ajudar a millorar?

Tractament / Prevenció

Efectes / Relació causal



Guidelines for reporting statistics in journals published by the American Physiological Society: the sequel

	Table 1. Statistical Concerns in Declined Papers. <sup>a</sup>		
Commentary	Concern	No.	
Statistical Considerations in Otolaryngology Journals  Brian W. Blakley, MD, PhD, FRCSC <sup>1</sup> , and Bryan Janzen, MD <sup>2</sup>	Conclusions not supported by data in the paper Confusing correlation with causation Fishing trip—raw data with no interpretation, unexplained significant P value, often from an administrative database	13 13 12	
AMERICAN ACADEMY OF OTOLARYNGOLOGY— HEAD AND NECK SURGERY  FOUNDATION	Overgeneralization  Multiple t tests for related outcomes with no correction for multiple tests  Multiple, novel, or inappropriate tests. "If you torture	9 8 8	
Otolaryngology— Head and Neck Surgery 1–3 © American Academy of Otolaryngology—Head and Neck Surgery Foundation 2017 Reprints and permission:	data with enough statistics it will confess to anything." Parametric tests when underlying assumptions are clearly violated	7	
sagepub.com/journalsPermissions.nav DOI: 10.1177/0194599817704395 http://otojournal.org	Conclusion absent or inappropriate  Lack of power analysis for negative results  Inappropriate or overlapping assignment to groups that	7 4 3	
	are not mutually exclusive  Data only partially analyzed  Other logical flaws	3 12	

## Que revisen els revisors estadístics

- Grandària Mostral
- Són vàlides les variables resultat?
- Són els mètodes adequats per als objectius?
- Estan les dades bàsiques ben descrites?
- Les proves estadístiques estan ben descrites i a més del valor p, es presenten els intervals de confiança?
- S'han tingut en compte els confusors i els biaixos?

# Variables

## Tipus de variables

#### **QUALITATIVES**

Recullen qualitats o atributs

#### **QUANTITATIVES**

Prenen valors numèrics i aquests tenen sentit

#### **Nominals**

Les categories descriuen una qualitat

#### Sexe

home dona Malaltia, Síntoma,...

#### **Ordinals**

Les categories descriuen una qualitat i segueixen un ordre

#### Nivell de estudis

Sense estudis o primaris Secundaris Universitaris Gravetat, Escala Likert

#### Discretes

Pren valors enters

Nº fills

 $\mathsf{C}$ 

1

2

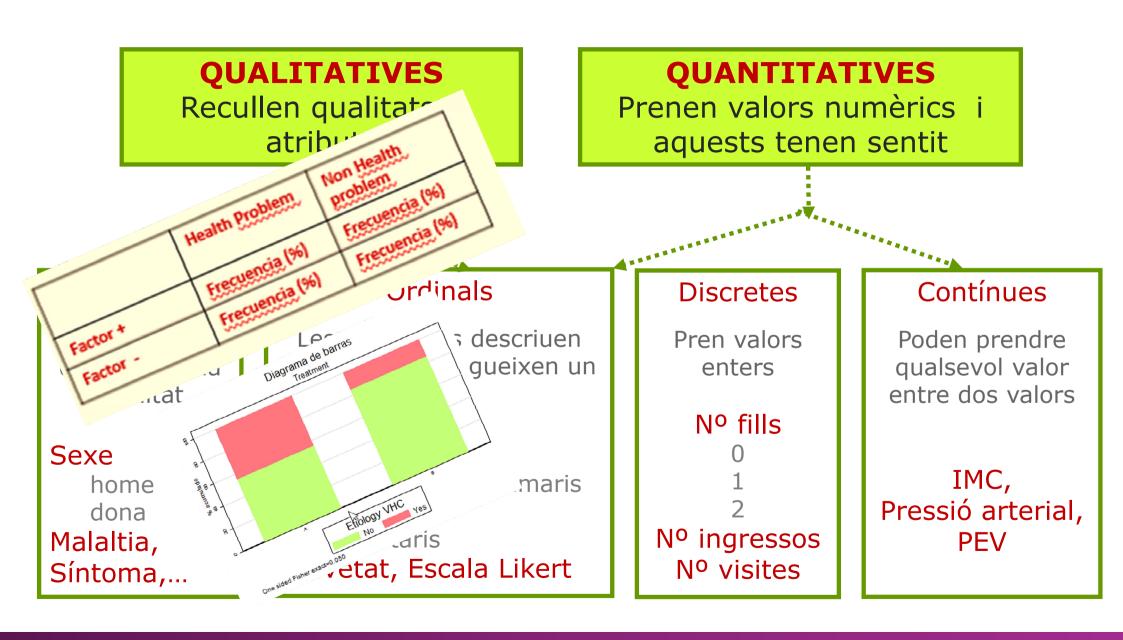
Nº ingressos Nº visites

#### Continues

Poden prendre qualsevol valor entre dos valors

IMC, Pressió arterial, PEV

## Tipus de variables



## Tipus de variables

#### **QUALITATIVES**

Recullen qualitats o atributs

#### QUANTI

Prenen val aque

TOAH

No

No

#### Nominals

Les categories descriuen una qualitat

#### Sexe

home dona Malaltia, Síntoma,...

#### **Ordinals**

Les categories des una qualitat i segueixe ordre

#### Nivell de estudis

Sense estudis o primaris Secundaris Universitaris

Gravetat, Escala Likert

de Shapiro Wilk per edat diagnostic That pro 1000 20,00 110,00 36,00 130,00 118,00 39,00 11 Control ontinues 21.42(14.87) [Min; Max] 1000:48,00] prendre 1. Miller test. Test de normalitat de sh

TOAH

√ arterial, PEV

col valor

valors

## La taula I. Descripció de les dades basals

- Les taules molt grans són difícils de seguir pels autors
- Les taules molt simples no tenen tota la informació.
- Cal utilitzar les mesures descriptives adients per a cada variable
- Les dades que no són essencials i que no aporten res no s'haurien de posar
- Si la variable no te una distribució simètrica, la mitjana no es la millor mesura descriptiva. Es millor utilitzar la mediana, els percentils o el min/max
- Utilitzar la forma "mitjana (SD)" en comptes de "mitjana (±SD)"
- NO s'utilitza l'Error Estándar per descriure la variable
- Posar el nombre de casos en el que es basa la mesura

#### Estadística: pensar-la abans de començar

JAMA Otolaryngology-Head & Neck Surgery | Original Investigation

#### Analgesic Effects of Intravenous Acetaminophen vs Placebo for Endoscopic Sinus Surgery and Postoperative Pain

A Randomized Clinical Trial

Matthew A. Tyler, MD; Kent Lam, MD; Faramarz Ashoori, MD, PhD; Chunyan Cai, PhD; Joshua J. Kain, MD; Samer Fakhri, MD; Martin J. Citardi, MD; Davide Cattano, MD, PhD; Amber Luong, MD, PhD

Table 1. Preoperative and Intraoperative Parameters for Patients Receiving Intravenous Acetaminophen (IVAPAP) and Placebo

Parameter	IVAPAP Group (n=31)	Placebo Group (n=29)
Preoperative variables		
Age, mean (SD), y	53.7 (14.6)	53.7 (14.7)
Sex, No. (%) <sup>a</sup>		
Men	22 (71)	13 (45)
Women	9 (29)	16 (55)
Ethnicity, No. (%) <sup>a</sup> of patients		
Asian	2 (7)	1 (3)
Black	4 (13)	3 (10)
Hispanic	2 (7)	1 (3)
Other	0	1 (3)
White	23 (74)	23 (79)
Disease phenotypes, No. (%) <sup>a</sup> of patients		
CRSwNP	19 (61)	20 (69)
CRSsNP	12 (38)	9 (31)
Lund Mackay score,	13.8 (6.7)	15.3 (6.1)

Tabla 1 Variables significativas asociadas a la presencia de depresión en pacientes con DM1.

Variable	Media y desviación típica	Rangos medios	р
Sexo Hombre (n = 99) Mujer (n = 108)	97,7 + 44,51 122,27 51,01		< 0,001*
Fumador/a Sí (n = 64) No (n = 143)		118,98 97,29	0,016**
Población activa Sí (n = 115) No (n= 91)	99,35 + 41,13 124 / 1 <u>+</u> 55,52		< 0,001*

# Mitjana o Mediana?

# Mitjana

μ

- •Serveix per localitzar les dades i és la suma dels valor observats partit per la grandària de la mostra
- •És una mesura que es veu alterada per valors extrems

#### **Exemple Dies d'Estada**

3, 4, 6, 9, 12

Mitjana=6,8

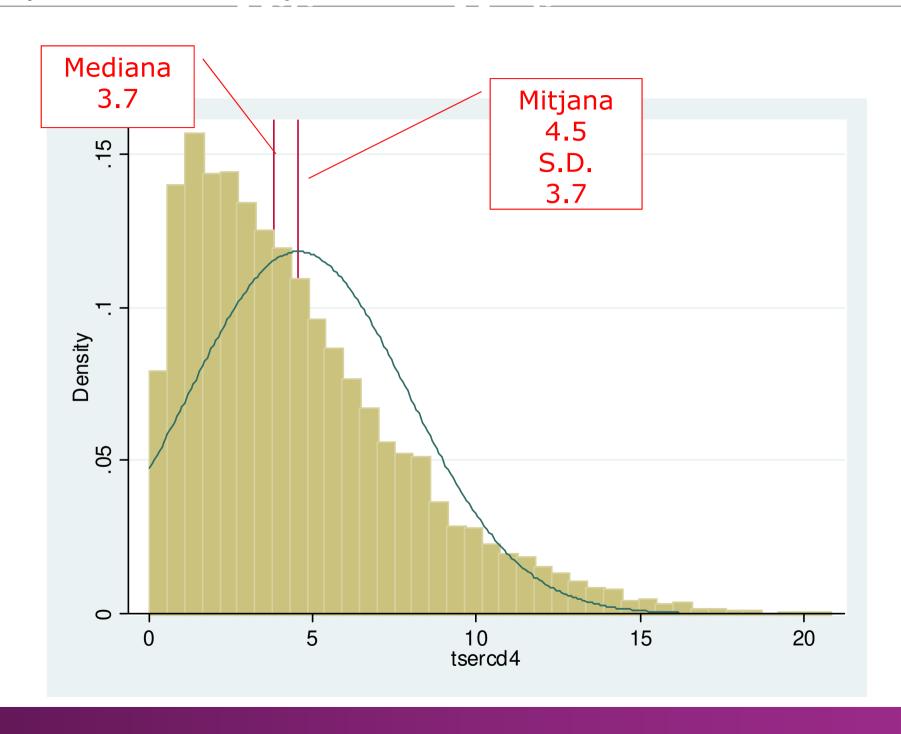
3, 4, 6, 9, 20

Mitjana=8,4

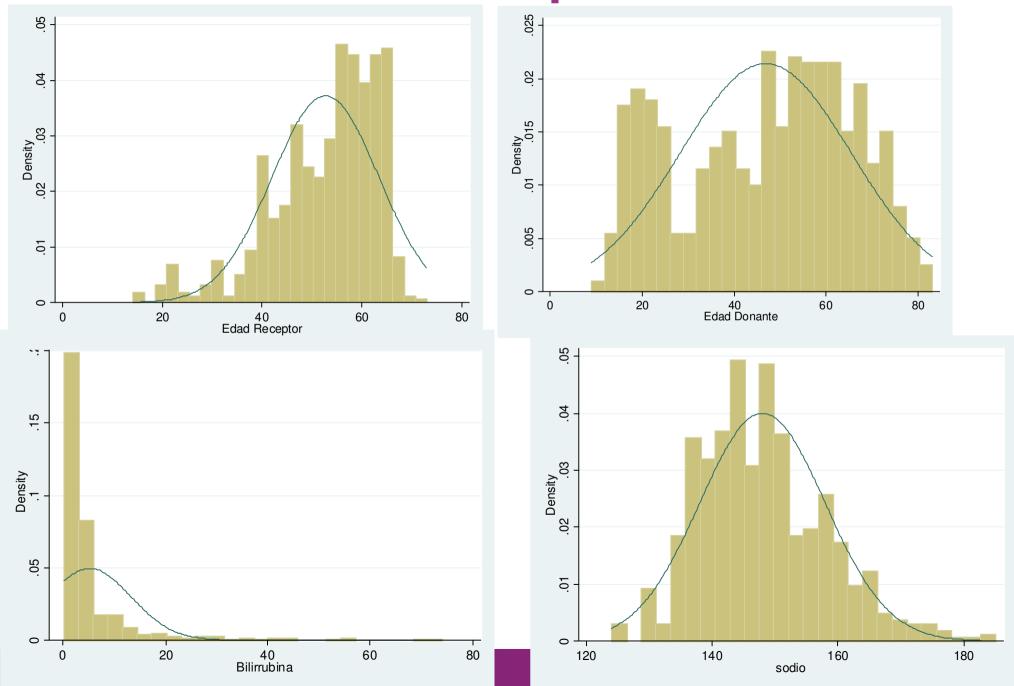


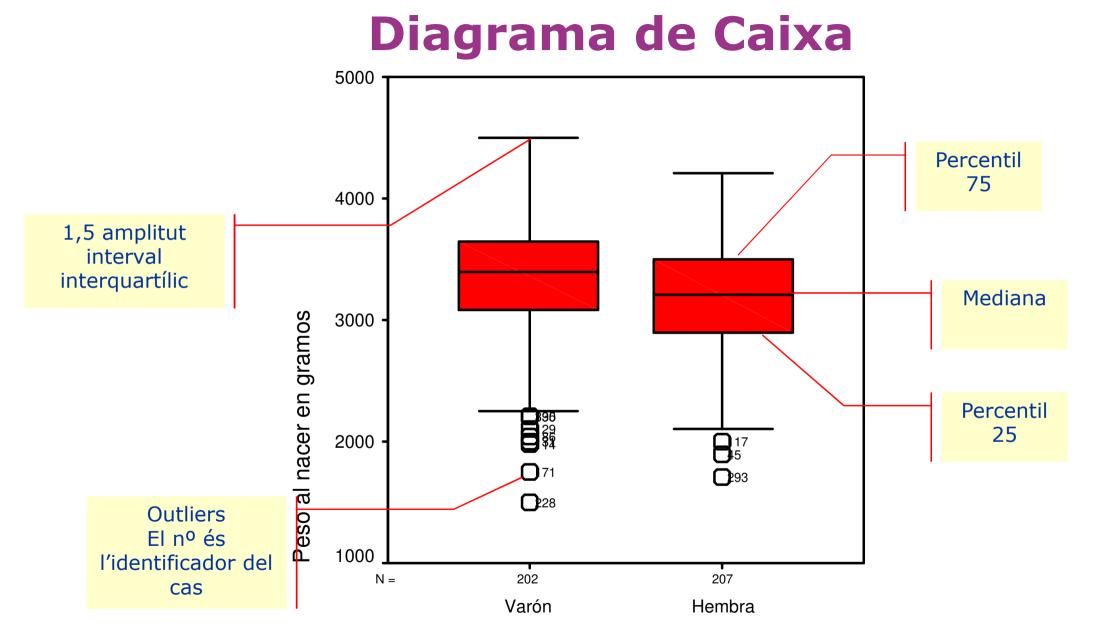
- •Es el punt que divideix les observacions en dues parts iguals
- •S'ordenen les observacions de la més petita a la més gran i es tria el punt central
- •No es veu alterada per observacions extremes

<b>Exemple Dies d'Estada</b>	
3, 4, 6, 9, 12	Mediana=6
3, 4, 6, 9, 20	Mediana=6



# Estudi de transplantaments





Sexo del recien nacido

# Factores que influyen en el síndrome de apneas-hipopneas durante el sueño

Damián Malia, Fernando Sánchez-Gascón, José Antonio Ros, María Gómez-Gallego, Pedro Castell, Cristina Nombela, Pedro Méndez, María del Carmen Abellán y José Hernández

Grupo de Fisiopatología del Sueño. Universidad de Murcia. Servicio de Neumología. Hospital General Universitario Virgen de la Arrixaca. El Palmar. Murcia. España.

TABLA 1

Datos de las 127 historias clínicas de los pacientes con síndrome de apneas-hipopneas durante el sueño

	Mínimo	Máximo	Mediana	Medla (DE)	IC del 95% de la media
Edad (años)	27	77	52	52,39 (10,92)	50,48-54,31
Peso (kg) Talla (m)	48 1.40	140 1.92	89 1,67	89,09 (10,4) 1,63 (0,08)	86,28-91,91 1,64-1,67
IMC (kg/mol)	20,8	49,3	32	32,48 (5,49)	31,51-34,4
Perímetro cervical (cm) Escala de Epworth	30 5	59 21	44 15	44,34 (4,50) 14,44 (4,72)	43,55-45,13 13,61-15,67
IAH	10	90	40	40,30 (21,58)	36,50-44,08
Consumo de alcohol (g)	0	150	30	31,50 (30,27)	26,18-36,01

IMC: índice de masa corporal, IAH: índice de apneas-hipopneas; DE: desviación estándar; IC: intervalo de confianza.

# Desviació Típica i Error Estàndar

# Variança



 Mesura de la dispersió mitjana de les dades respecte a la mitjana

	Dies (d)	d-mitjana	(d-mitjana)^2	d	d-mitjana	(d-mitjana)^2
	3	-3,8	14,44	3	-5,4	29,16
	4	-2,8	7,84	4	-4,4	19,36
	6	-0,8	0,64	6	-2,4	5,76
	9	2,2	4,84	9	0,6	0,36
	12	5,2	27,04	20	11,6	134,56
Suma		0	54,8		0	189,2
Suma/5	6,8	0	10,96	8,4	0	37,84

# **Desviació Típica**

- És l'arrel quadrada de la variança
- Es mesura en les mateixes unitats que la variable

#### **Exemple Dies de Estancia**

3,	4,	6,	9,	<b>12</b>	<b>Variança = 10.96</b>
_	_	_	-		

### **Error Estándar**

- Es la desviació típica si s'obtinguessin totes les mostres possibles d'una població
- Normalment és el quocient entre la desviació típica de la variable i l'arrel quadrada de la grandària mostral

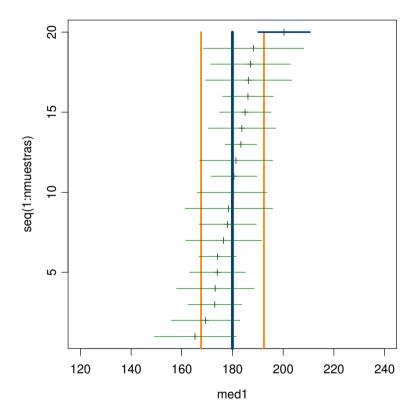
error estàndar = 
$$\frac{\sigma}{\sqrt{n}}$$

# Intervals de confiança

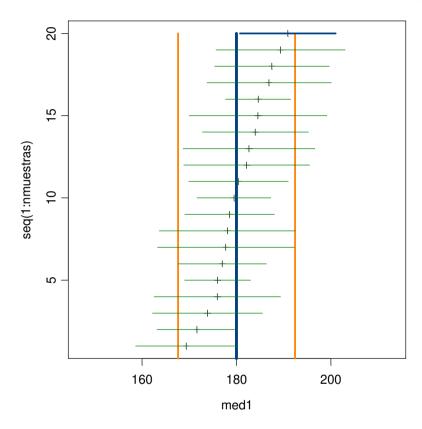
# Interval de confiança

El 95% dels intervals construits com m $\pm 1,96 * \sigma/\sqrt{n}$  contenen la mitjana poblacional

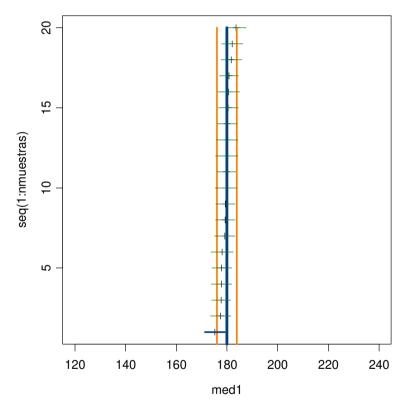
En general l'expressió per a qualsevol estimador és Estimador  $\pm$  Coef.<sub>1-a/2</sub> x Error Estándar 20 muestras de tamaño 10 media 180 desv.tip. 20

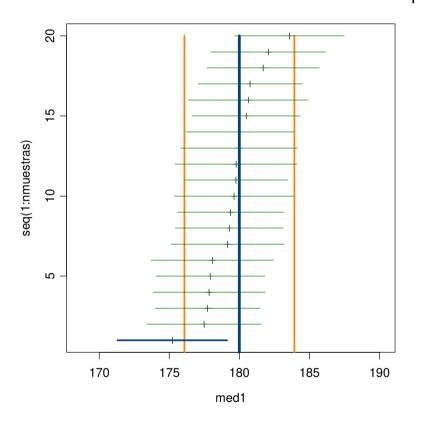


20 muestras de tamaño 10 media 180 desv.tip. 20

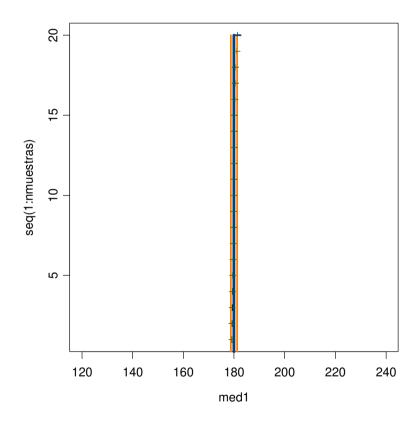


20 muestras de tamaño 100 media 180 desv.tip. 20 muestras de tamaño 100 media 180 desv.tip. 20

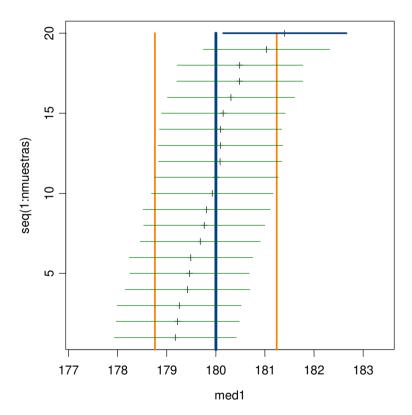




20 muestras de tamaño 1000 media 180 desv.tip. 20



20 muestras de tamaño 1000 media 180 desv.tip. 20



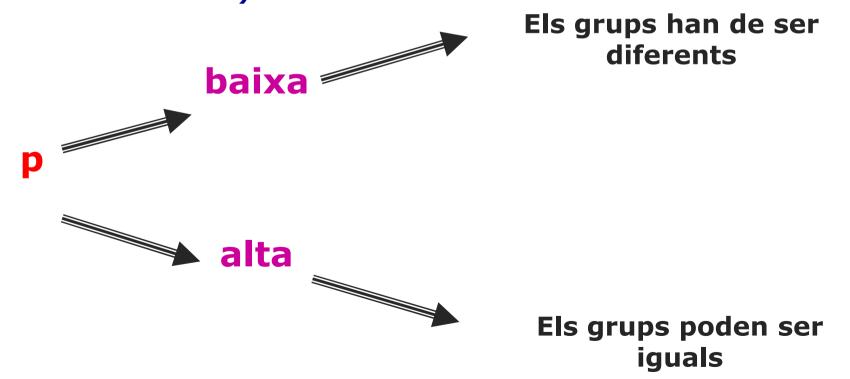
# Contrastos d'Hipòtesis

# Contrastos d'Hipòtesis

- La interpretació del test depèn de la grandària mostral
- Indicar el test seleccionat
- Indicar la mínima diferencia clínicament important
- En un estudi de equivalència o no inferioritat, indicar la màxima diferència per a dir que són equivalents
- Calcular mesures d'impacte (diferència de mitjanes, RR, OR i el seus IC95%)
- Indicar si s'han efectuat ajustos per comparacions múltiples
- Especificar el p-valor a no ser que sigui menor de 0.001.
- NO utilitzar NS o \*p<0.05</li>

# Què és el valor p?

La probabilitat (p) d'observar les diferències que mostren les dades entre els dos grups suposant que no existeixen diferències reals (No associació)



# Errors en un contrast d'hipòtesis

	Hipòtesis nul·la certa	Hipòtesis nul·la falsa
Test accepta hipòtesis nul·la		Error tipus II β
Test rebutja hipòtesis nul·la	Error tipus I α	<b>Poder (1- β)</b>

### El valor p depèn de la grandària mostral

1			
•	col		
row	1	2	Total
1	2	98	100
	2.00	98.00	100.00
2	+   4	96	100
	4.00	96.00	100.00
Total	+   6	194	200
	3.00	97.00	100.00

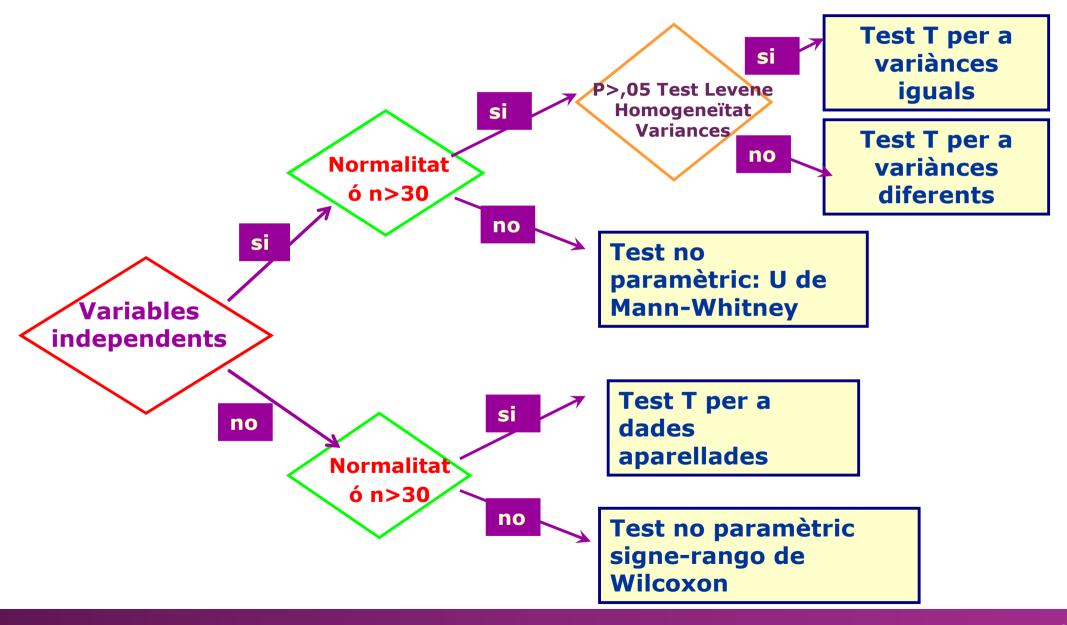
Valor p Pearson 0.407

row	col	2	Total
1	20	980	1,000
	2.00	98.00	100.00
2	40	960	1,000
	4.00	96.00	100.00
Total	60 3.00	1,940 97.00	2,000

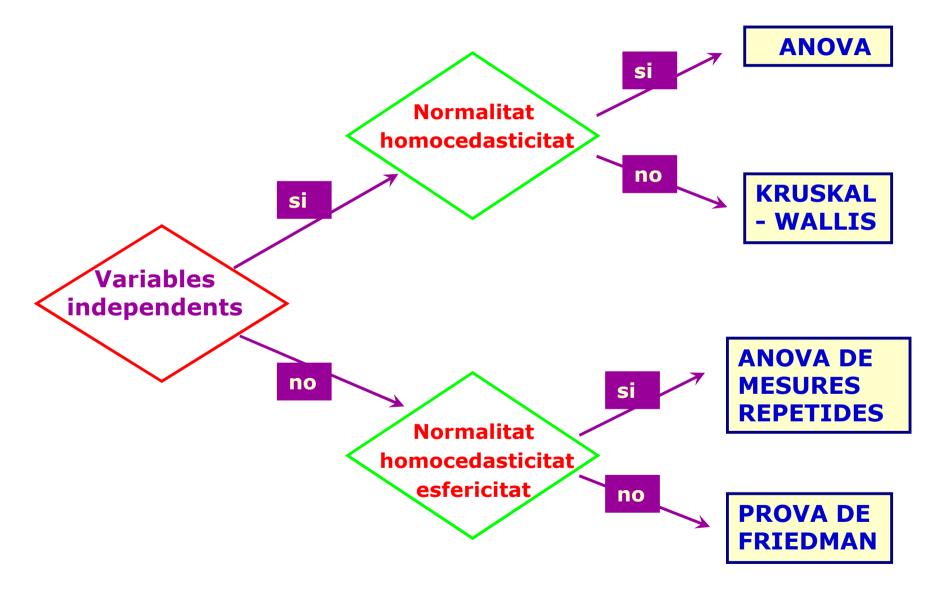
row	col   1	2	Total
1	4 2.00	196   98.00	200
2	8 4.00	192   96.00	200
Total	12 3.00	388   97.00	400 100.00

Valor p Pearson 0.241

#### Es disposa de 2 grups



#### Es disposa de 3 grups



A prospective cohort study assessing the clinical utility of the Cottle maneuver in

nasal septal surgery

James P. Bonaparte<sup>1\*</sup> ond Ross Campbell<sup>2</sup> □

Variable		All Patients	Positive Cottle	Negative Cottle	<i>p</i> -value
Patient Count	N	141	95	46	n/a
Age	Mean	41.3	40.4	43.4	0.17 <sup>a</sup>
	SD	13.4	12	16.3	
Gender	% Female	28.5	23.1	41.9	0.007 <sup>b</sup>
Baseline NOSE	Mean	13.3	13.1	13.7	0.43 <sup>a</sup>
	SD	4.06	4.1	3.9	
12 Month NOSE	Mean	4.1	4.2	4.2	0.57 <sup>a</sup>
	SD	4.6	4.6	4.5	
Change in NOSE	Mean	9.1	8.9	9.5	0.78 <sup>a</sup>
	SD	5.7	5.8	5.6	
% Change in	Mean	67.5	66.8	68.8	0.81 <sup>a</sup>
NOSE	SD	35.8	36.9	33.5	
Surgical Failure	n	14	9	5	0.99 <sup>b</sup>
	% of Total	9.9	9.5	10.5	

N number of patients, SD standard deviation, NOSE nasal obstruction symptom index, n/a not applicable

<sup>&</sup>lt;sup>a</sup>two-sample t-test, <sup>b</sup> Chi-square test

JOURNAL OF APPLIED CLINICAL MEDICAL PHYSICS, VOLUME 13, NUMBER 6, 2012

# Dose variations in tumor volumes and organs at risk during IMRT for head-and-neck cancer

Mercè Beltran, <sup>1a</sup> Mónica Ramos, <sup>2</sup> Juan José Rovira, <sup>1</sup> Santiago Perez-Hoyos, <sup>3</sup> Marc Sancho, <sup>1</sup> Enrique Puertas, <sup>2</sup> Sergi Benavente, <sup>2</sup> Mercè Ginjaume, <sup>4</sup> Jordi Giralt <sup>2</sup>

TABLE 2. Volumes with significant changes on  $CT_{15}$  and  $CT_{25}$  in relation to  $CT_{plan}$ .

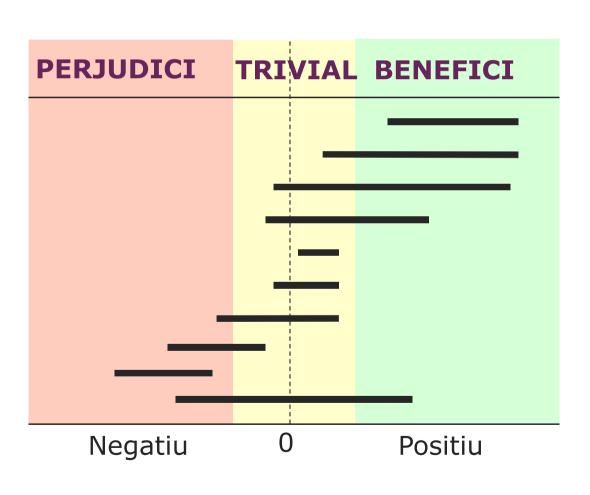
	PTV2		Limited Ex Contor		Contrala Parotid (		Ipsilate Parotid (	
	Mean <sup>a</sup> (SD)	p	Mean (SD)	p	Mean (SD)	p	Mean (SD)	p
$\mathrm{CT}_{\mathrm{plan}}$	130.8 (73.1)		1692 (412)		18.71 (10.3)		19.2 (10.6)	
CT <sub>15</sub>	116.7 (54.8)	0.07	1639 (402)	0.017	14.2 (6.6)	< 0.001	15.3 (8.3)	0.003
CT <sub>25</sub>	113.6 (53.5)	0.03	1638 (415)	0.014	13.1 (6.4)	< 0.001	13.5 (7.2)	<0.001

<sup>&</sup>lt;sup>a</sup> The mean value is the average volume between 16 patients, in cm<sup>3</sup>.

SD is the standard deviation; p < 0.05 is statistically significant.

Estadísticament significatiu, és realment rellevant?

# Clínicament vs estadísticament significatiu



Clínicament Estadísticament Significatiu Significatiu SI SI SI SI SI NO Depèn NO Depèn SI Depèn NO Depèn NO Depèn SI SI SI(no usar)

NO

NO

#### **Models estadístics**

Resposta Y	Exposició X	Model estadístic
Continua	Continua	Regressió Linial
Continua	Categòrica	Anova, Regressió Linial
Dicotòmica (si/no)	Continua/Categòrica	Regressió Logística
Categòrica (multinomial)	Continua/Categòrica	Regressió multinomial o politòmica
Recompte de casos	Continua/Categòrica	Regressió de Poisson
Temps fins l'esdeveniment	Continua/Categòrica	Regressió de Cox
Continua	Mesures repetides (temps/subjectes)	Anova Mesures repetides. Models mixtes o efectes aleatoris
Conjunt de variables	Continua/Categòrica	Anàlisis multivariant (PCA, Cluster, Machine Learning)

**Tabla 4** OR ajustadas

	OR	IC 95%		р
		Límite inferior	Límite superior	
Antecedentes tabaco	0,665	0,301	1,471	0,314
Edad	0,990	0,965	1,017	0,476
Sexo (mujer)	1,442	0,735	2,748	0,296
DM	2,323	1,281	4,214	0,006
HTA	0,746	0,392	1,419	0,371
IM previo	1,846	1,012	3,369	0,046
TCI ó 3 vasos	2,523	1,409	4,520	0,002

DM: diabetes mellitus; HTA: hipertensión arterial; IC 95%; intervalo de confianza del 95%; IM: infarto de miocardio; OR: odds ratio; TCI: tronco coronario izquierdo.

# Risc Relatiu Odds Ratio

## Risc

	Malal Present	tia Absent
Factor Present	<u>a</u>	<u>b</u> -
Factor Absent	<u>C</u>	<u>d</u>

a, b, c, d són el nombre de subjectes en cadascuna de les categories Risc de malaltia Factor <u>present</u> =

$$a/(a+b)$$

Risc de malaltia Factor <u>absent</u> =

$$c/(c+d)$$

## Risc Relatiu (RR)

	Malaltia Present Absent				
Factor Present	<u>a</u>	<u>b</u>			
Factor Absent	<u>C</u>	<u>d</u>			

a, b, c, d són el nombre de subjectes en cadascuna de les categories Risc de malaltia amb el factor present RELATIU al risc de malaltia amb el factor absent=

$$\frac{a/(a+b)}{c/(c+d)}$$

# Interpretació RR

- RR>1 Més risc entre aquells que tenen el factor
- RR=1 Risc igual en els dos grups
- RR<1 Menys risc entre aquells que NO tenen el factor

#### **ODDS**

	Malaltia Present Absent			
Factor Present	<u>a</u>	<u>b</u> .		
Factor Absent	<u>C</u>	<u>d</u>		

a, b, c, d són el nombre de subjectes en cadascuna de les categories Odds de malaltia Factor <u>present</u> = a/(b)

Odds de malaltia Factor <u>absent</u> =  $\frac{c}{d}$ 

## Odds Ratio (OR)

	Malaltia Present Absent				
Factor Present	<u>a</u>	<u>b</u>			
Factor Absent	<u>C</u>	<u>d</u>			

a, b, c, d són el nombre de subjectes en cadascuna de les categories Odds de malaltia amb el factor present RELATIU a la odds de malaltia amb el factor absent=

$$\frac{\frac{a}{b}}{\frac{c}{d}} = \frac{ad}{bc}$$

# Interpretació OR

- OR>1 Més odds entre aquells que tenen el factor
- OR=1 Odds igual en els dos grus
- OR<1 Menys odds entre aquells que NO tenen el factor

Tabla 4 OR ajustadas

	OR	IC 95%		р
		Límite inferior	Límite superior	
Antecedentes tabaco	0,665	0,301	1,471	0,314
Edad	0,990	0,965	1,017	0,476
Sexo (mujer)	1,442	0,735	2,748	0,296
DM	2,323	1,281	4,214	0,006
HTA	0,746	0,392	1,419	0,371
IM previo	1,846	1,012	3,369	0,046
TCI ó 3 vasos	2,523	1,409	4,520	0,002

DM: diabetes mellitus; HTA: hipertensión arterial; IC 95%: intervalo de confianza del 95%; IM: infarto de miocardio; OR: odds ratio; TCI: tronco coronario izquierdo.

# Guies sugerides per reportar anàlisis estadístiques (Sterne and Davey Smith, *BMJ* 2001; 322: 226-231)

- No és acceptable la descripció de diferències com a "statistically significant"
- S'han d'incloure els intervals de confiança dels principals resultats. Aquests no s'han d'interpretar com un sucedani del nivell de significació
- Quan no es rebutja la hipòtesi nul·la, cal donar el valor p
- A l'hora d'analitzar subgrups cal ser escèptic, sobretot si no s'ha previst a l'hora de dissenyar el treball i calcular la grandària mostral
- En estudis observacionals cal recordar que els biaixos i les variables de confusió són tan importants com els resultats "significatius"

# Guies sugerides per reportar anàlisis estadístiques (Sterne and Davey Smith, BMJ 2001; 322: 226-231)

"There was a statistically significant odds ratio of 2.8 (95% CI 1.3 to 6.2)...."

"The odds ratio was 2.8 (95% CI 1.3 to 6.2, P=0.01)

"P values less than 0.05 were regarded as statistically significant."

### ueb.vhir.org

