

## Metodi statistici per lo studio dei fenomeni biologici

Alla fine di questa lezione dovreste essere in grado di:

- ❑ conoscere la differenza fra parametri e stime
- ❑ conoscere il concetto di variabilità campionaria
- ❑ spiegare la differenza fra errore sistematico e casuale



## Statistica e matematica

### MATEMATICA

- funzione nota

$$y = a + bx$$

- parametri (a, b) noti
- x o y ignota

Stime dei parametri

### STATISTICA

- modello ignoto

*lineare?*

$$Y = \alpha + \beta x$$

- parametri ( $\alpha$ ,  $\beta$ ) ignoti
- x o y entrambe note

$$y_i = \alpha + \beta x_i + \varepsilon_i$$
$$\hat{y}_i = \hat{\alpha} + \hat{\beta} x_i$$

## Parametri

Quantità che definiscono un modello teorico.

Non si riferiscono ai dati effettivamente osservati, ma possono essere usati per predire il comportamento atteso nel singolo soggetto

## Stime

Quantità, ottenute a partire dai dati effettivamente osservati, per stimare ('indovinare') i valori dei parametri

## Statistica

---

L'obiettivo è riconoscere le cause rilevanti di variabilità

- dimostrare che esiste un'associazione tra fattore studiato (esposizione) e fenomeno in esame (esito)
- dimostrare che l'eventuale associazione osservata non è in realtà un artefatto
- stimare l'entità dell'associazione fra esposizione ed esito

## Esempi (da Oxman et al. JAMA 1998)

### Terapia

Il sumatriptan (**esposizione**) riduce l'intensità del dolore (**esito**) nei pazienti con emicrania ricorrente?

### Diagnosi precoce

La ricerca del PSA (**esposizione**) negli anziani asintomatici riduce il rischio di morte per cancro della prostata (**esito**)?

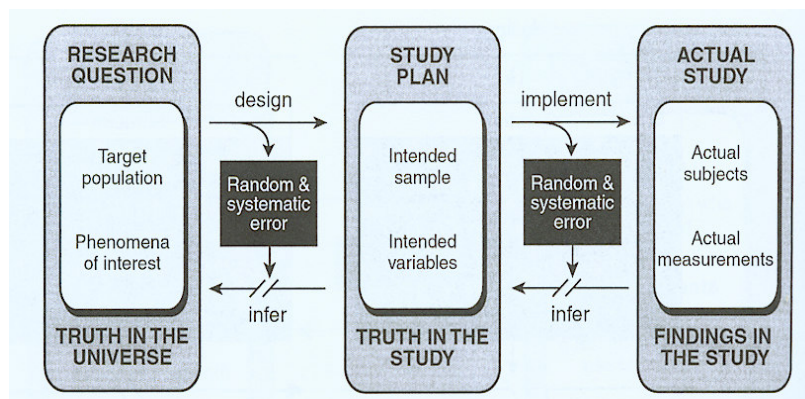
### Prognosi

Le convulsioni febbrili nel bambino (**esposizione**) modificano la probabilità di sviluppare una qualche forma di epilessia (**esito**)?

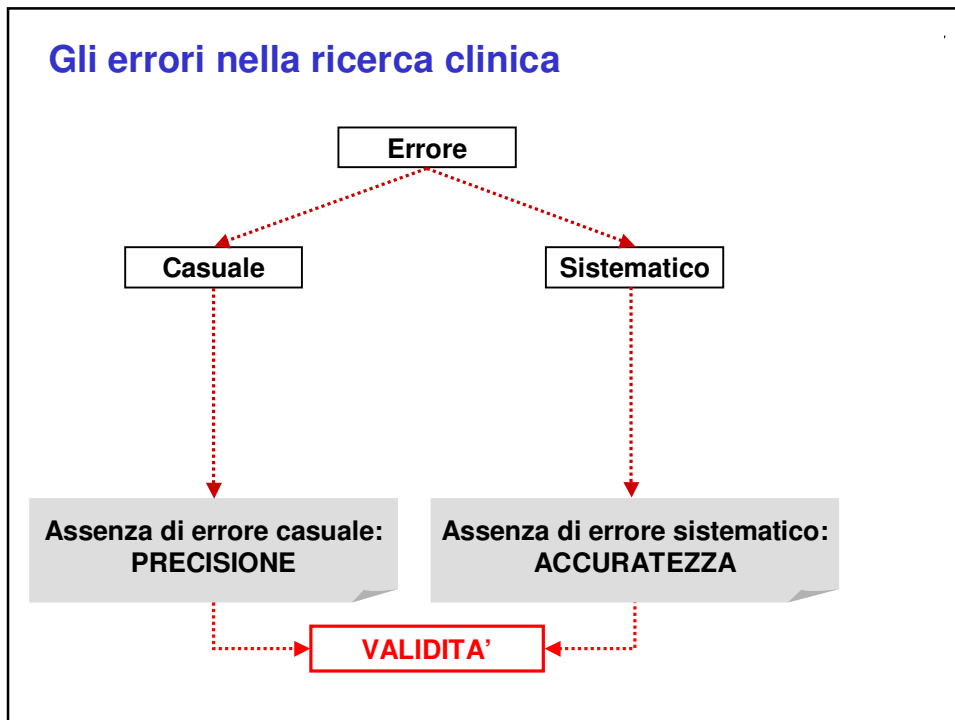
### Danno

I  $\beta$ -agonisti (**esposizione**) modificano la probabilità di morte nei pazienti asmatici (**esito**)?

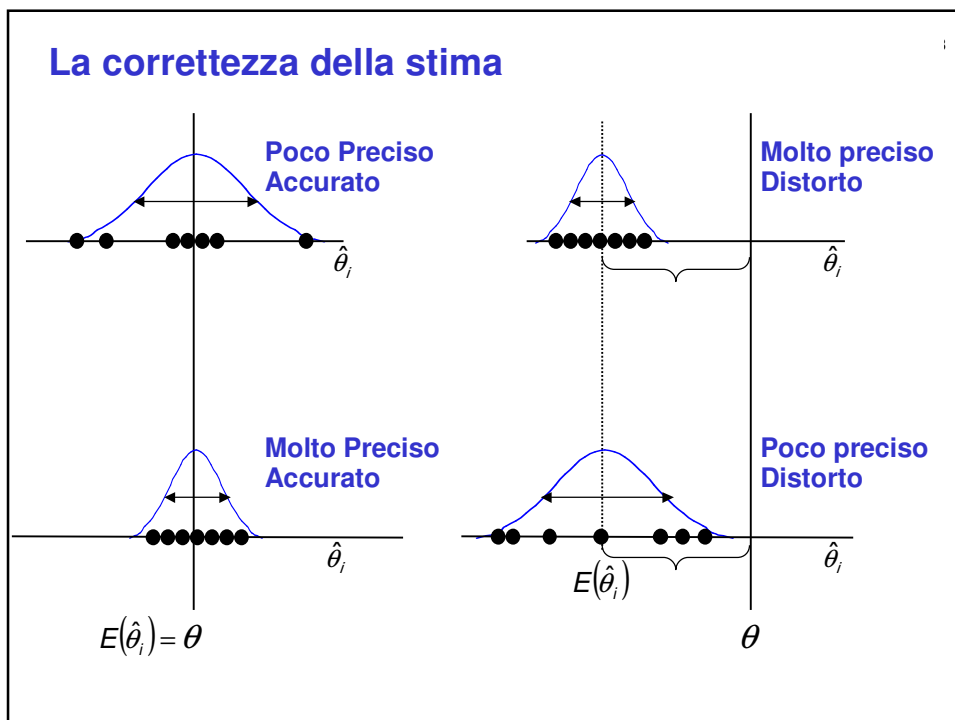
## La struttura della ricerca clinica

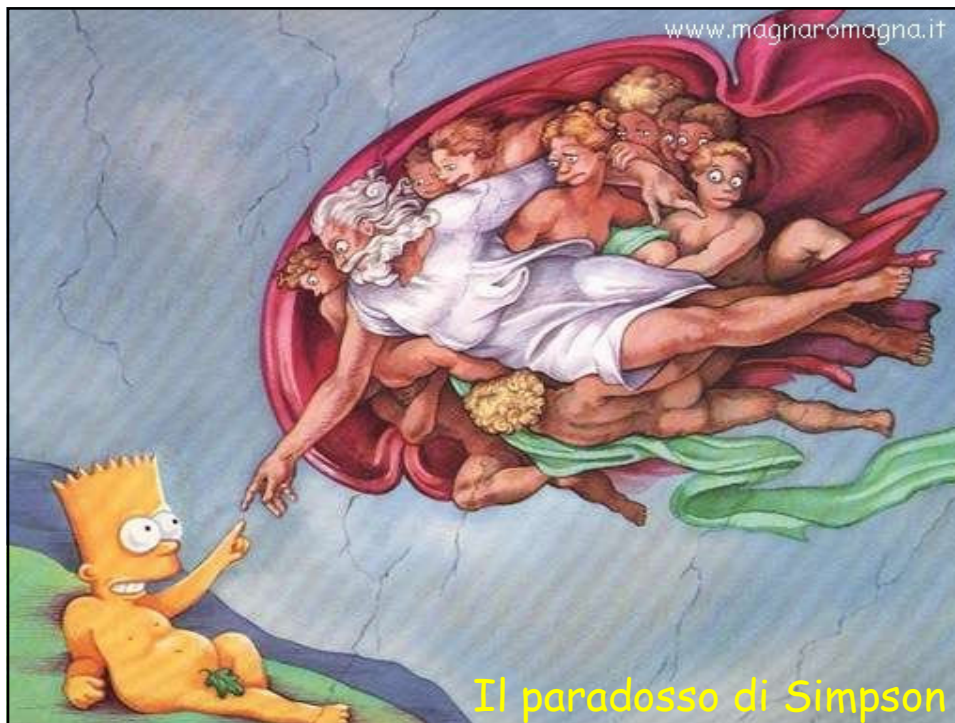


## Gli errori nella ricerca clinica



## La correttezza della stima





**In quale di due ospedali A e B si ha una maggiore mortalità?**

---

	<b>Dimessi vivi</b>	<b>Morti</b>	<b>Totale</b>	
<b>Ospedale A</b>	<b>800</b>	<b>200</b>	<b>1,000</b>	<b>20%</b>
<b>Ospedale B</b>	<b>900</b>	<b>100</b>	<b>1,000</b>	<b>10%</b>

## In quale di due ospedali A e B si ha una maggiore mortalità?

### Pazienti in buone condizioni

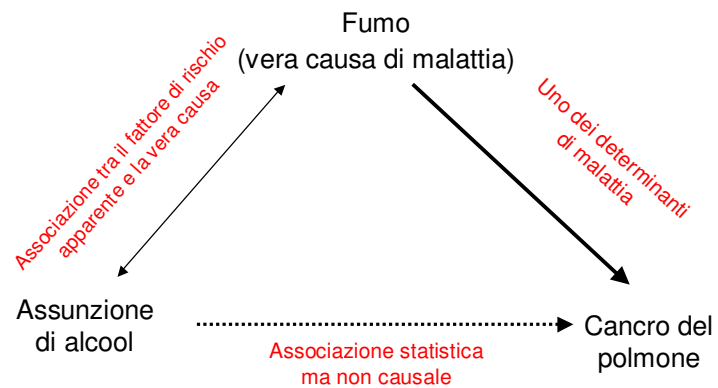
	Dimessi vivi	Morti	Totale	
Ospedale A	590	10	600	1.7%
Ospedale B	870	30	900	3.3%

### Pazienti in cattive condizioni

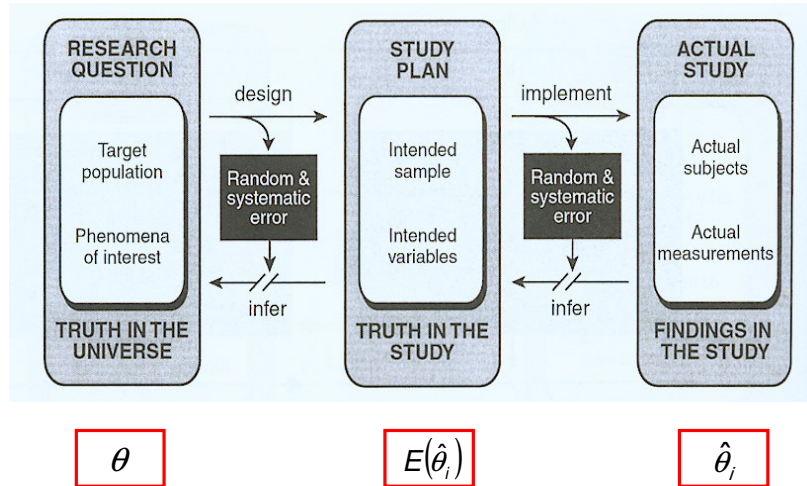
	Dimessi vivi	Morti	Totale	
Ospedale A	210	190	400	47.5%
Ospedale B	30	70	100	70%

## Il confondimento

Si ha quando i gruppi confrontati differiscono per caratteristiche diverse dal fattore studiato



## La struttura della ricerca clinica



## L'osservazione

Metodi statistici per lo studio dei fenomeni biologici/14

$$\begin{array}{rcl}
 \text{Effetto studiato} & + & \\
 \text{Errore sistematico} & + & \\
 \text{Errore casuale} & = & \\
 \hline
 \text{Effetto osservato} & & 
 \end{array}$$

## Il modello statistico

**Effetto studiato** +

**Errore casuale** =

---

**Effetto osservato**

**N.B. In assenza di errore sistematico!**

## Statistica

---

L'obiettivo è riconoscere le cause rilevanti di variabilità

- dimostrare che esiste un'associazione tra fattore studiato (esposizione) e fenomeno in esame (esito)
- dimostrare che l'eventuale associazione osservata non è in realtà un artefatto
- stimare l'entità dell'associazione fra esposizione ed esito

La statistica fornisce gli strumenti per valutare le ipotesi a partire dalla variabilità casuale dei fenomeni

Table 1. Statistical Content of Original Articles in the *New England Journal of Medicine* over Time.\*

Statistical Procedure	Original Articles Containing Methods			Accumulation by Article†
	1978–1979	1989	2004–2005	2004–2005
	number (percent)			
No statistical methods or descriptive statistics only	91 (27)	14 (12)	39 (13)	39 (13)
t-Tests	147 (44)	43 (39)	80 (26)	42 (14)
Contingency tables	91 (27)	41 (36)	166 (53)	47 (15)
Nonparametric tests	38 (11)	24 (21)	85 (27)	53 (17)
Epidemiologic statistics	33 (10)	25 (22)	110 (35)	55 (18)
Pearson's correlation	40 (12)	22 (19)	10 (3)	56 (18)
Simple linear regression	28 (8)	10 (9)	19 (6)	56 (18)
Analysis of variance	25 (8)	23 (20)	50 (16)	61 (20)
Transformation	23 (7)	8 (7)	31 (10)	62 (20)
Nonparametric correlation	13 (4)	1 (1)	14 (5)	65 (21)
Survival methods	36 (11)	37 (32)	190 (61)	74 (24)
Multiple regression	15 (5)	16 (14)	160 (51)	122 (39)

NEJM 2005; 353: 1977-9 1/2

Table 1. Statistical Content of Original Articles in the *New England Journal of Medicine* over Time.\*

Multiple comparisons	11 (3)	10 (9)	70 (23)	127 (41)
Adjustment and standardization	9 (3)	10 (9)	3 (1)	128 (41)
Multiway tables	12 (4)	11 (10)	39 (13)	136 (44)
Power analyses	10 (3)	4 (3)	121 (39)	211 (68)
Cost-benefit analysis	3 (1)	0	1 (<1)	212 (68)
Sensitivity analysis‡	0	0	18 (6)	223 (72)
Repeated-measures analysis	—	—	37 (12)	249 (80)
Missing-data methods	—	—	26 (8)	272 (87)
Noninferiority trials	—	—	11 (4)	283 (91)
Receiver-operating characteristic	—	—	7 (2)	288 (93)
Resampling	—	—	5 (2)	293 (94)
Principal component analysis and cluster analysis	—	—	5 (2)	298 (96)
Other methods§	12 (4)	10 (9)	13 (4)	311 (100)
Total				
Articles	332	115	311	311
Article-method uses	637	311	1510	
Average uses of methods per article	1.9	2.7	4.2	

NEJM 2005; 353: 1977-9 1/2