



CAUSALITÀ *versus* ASSOCIAZIONE

Maria Cecilia Giron

Dipartimento di Scienze del Farmaco

Edificio di Farmacologia

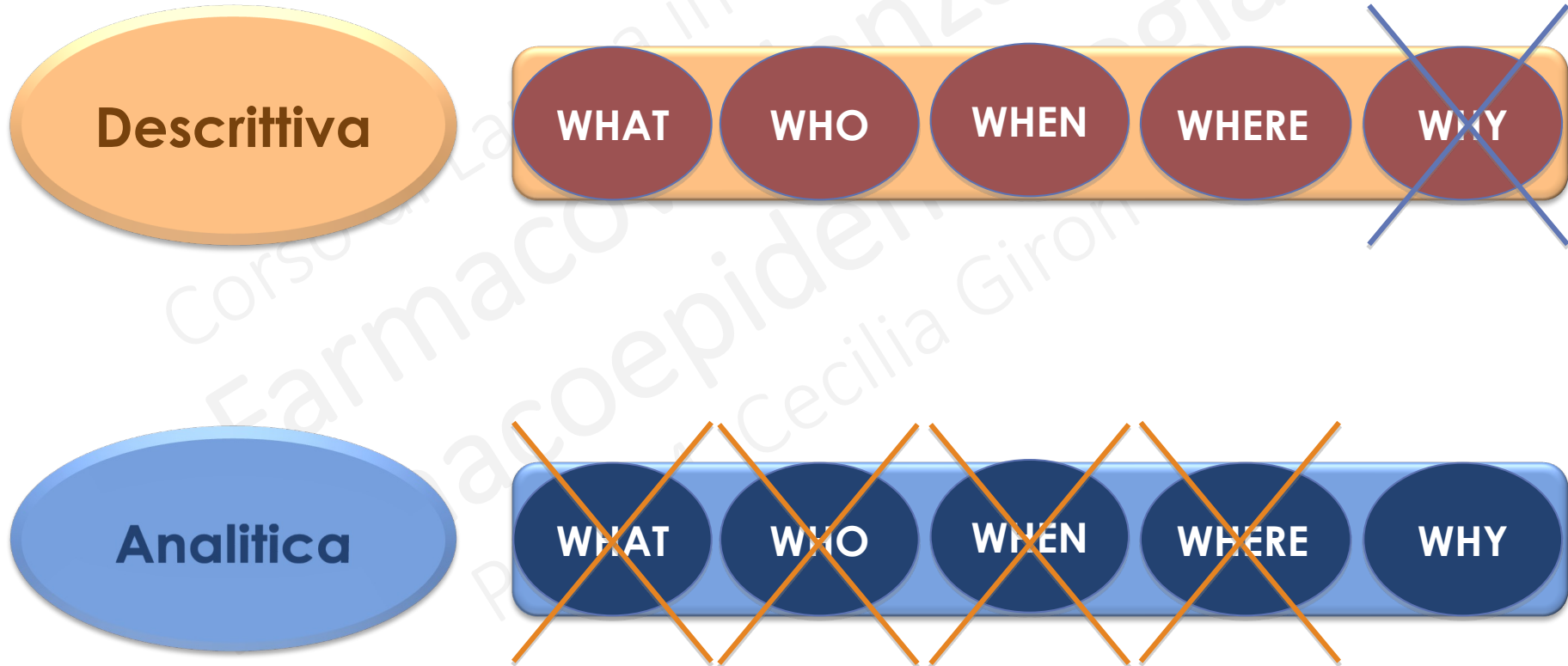
Email: cecilia.giron@unipd.it

Tel. 049-8275091



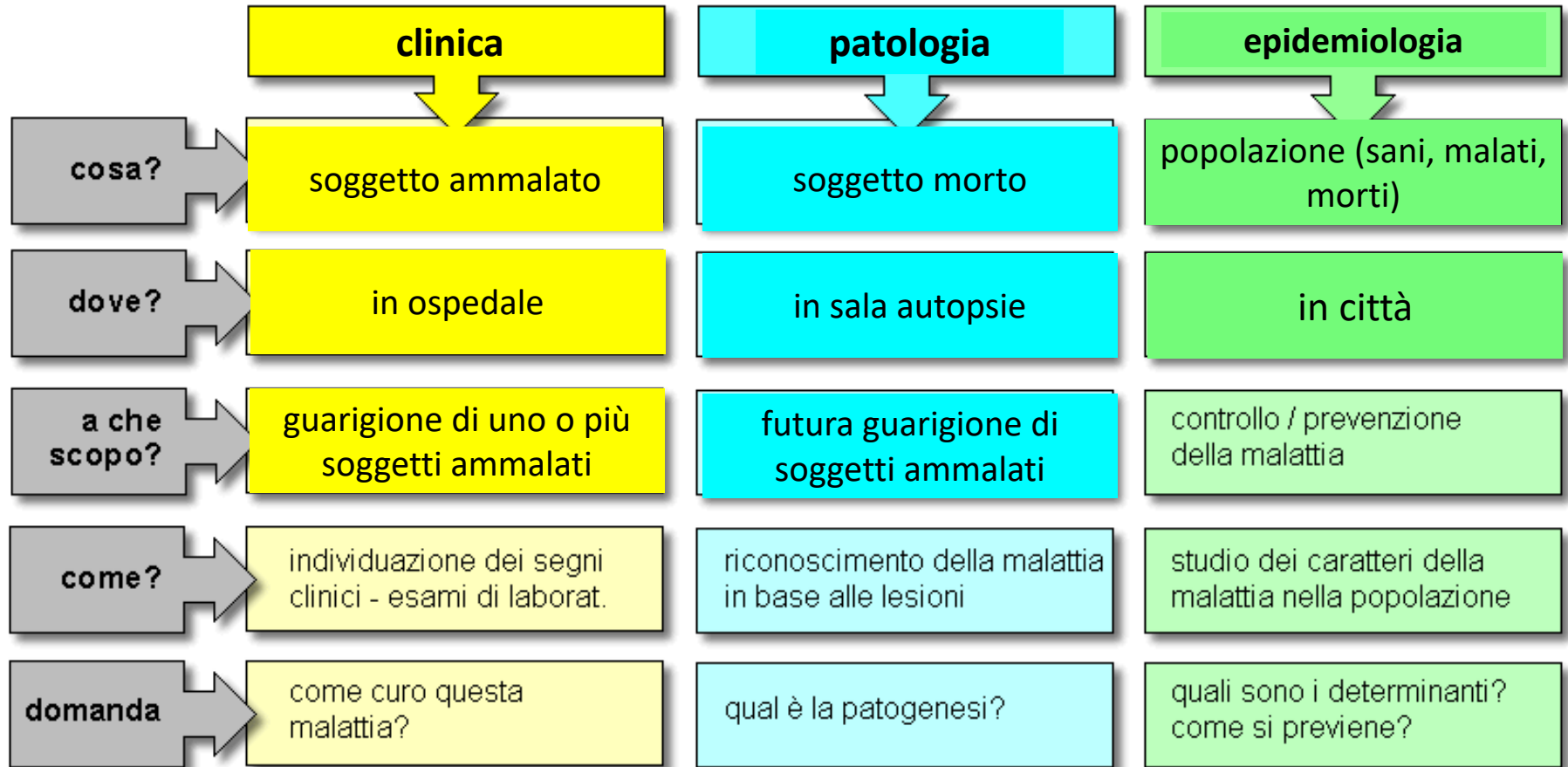
EPIDEMIOLOGIA

descrittiva vs analitica





EPIDEMIOLOGIA



Conoscere la malattia per prevenirla meglio



Ciclo di vita di un medicinale

Preclinica

Clinical
trials

Immissione in
commercio

Pratica
clinica

Sicurezza clinica

Sorveglianza

Passiva

Attiva

Farmacovigilanza

- Identificazione
- Segnali di allarme

Strumenti:

- Segnalazioni spontanee
- Letteratura
- Valutazione di serie di casi

Farmacoepidemiologia

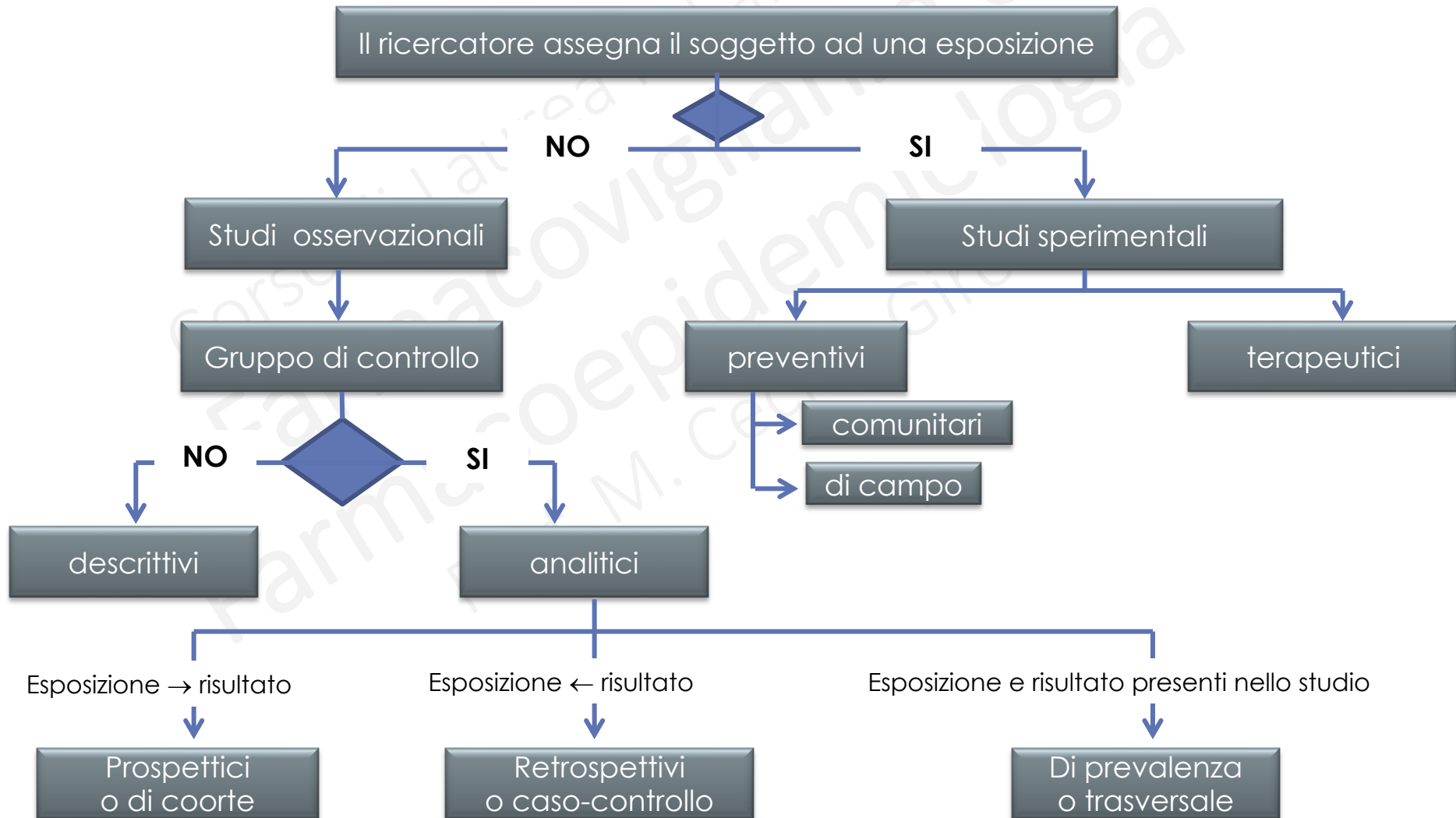
- Verifica / conferma
- Quantificazione dei rischi

Strumenti:

- Studi descrittivi
- Studi analitici

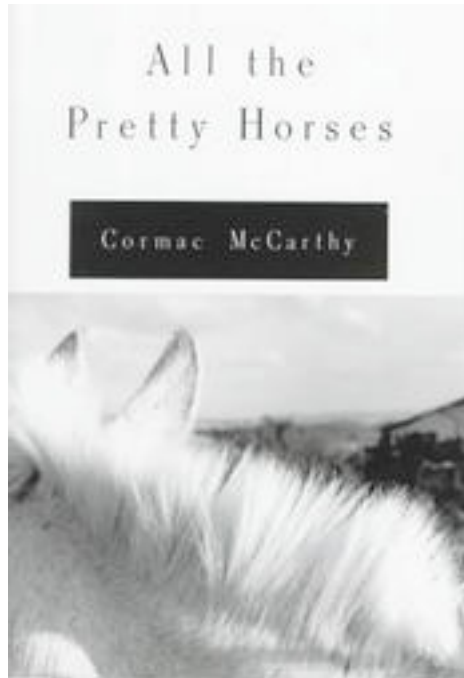


ALGORITMO per la classificazione degli studi di FARMACOEPIDEMIOLOGIA



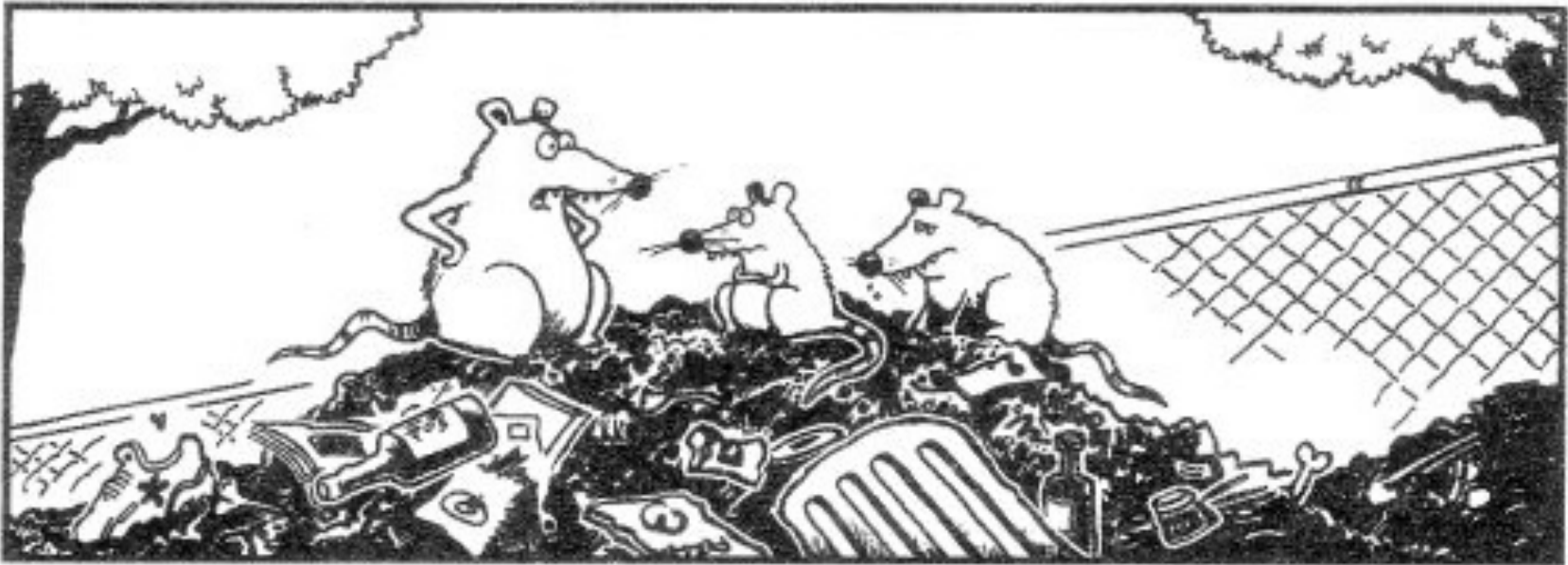


La STORIA e i gruppi di controllo



«A scuola, studiando biologia, ho imparato che negli esperimenti gli scienziati prendono un gruppo di batteri, o di topi, o di uomini, e lo sottopongono a condizioni particolari, poi confrontano i risultati ottenuti con un gruppo analogo che non ha subito condizionamenti. Quest'ultimo si chiama gruppo di controllo, ed è quello che consente agli scienziati di misurare gli effetti dell'esperimento, di valutare la portata dei cambiamenti. Ma nella storia non ci sono gruppi di controllo e nessuno può dire cosa sarebbe successo altrimenti. Noi piangiamo pensando a ciò che avrebbe potuto succedere al posto di ciò che è successo, ma questa eventualità non esiste, non è mai esistita. Dicono che chi non conosce la storia è condannato a ripeterla, ma io non credo che conoscerla serva a qualcosa. L'avidità, la follia e l'attrazione per il sangue sono una costante della storia, e questa è una cosa che persino Dio - che sa tutto quel che si può sapere - sembra incapace di modificare.»

la saggezza degli animali da laboratorio

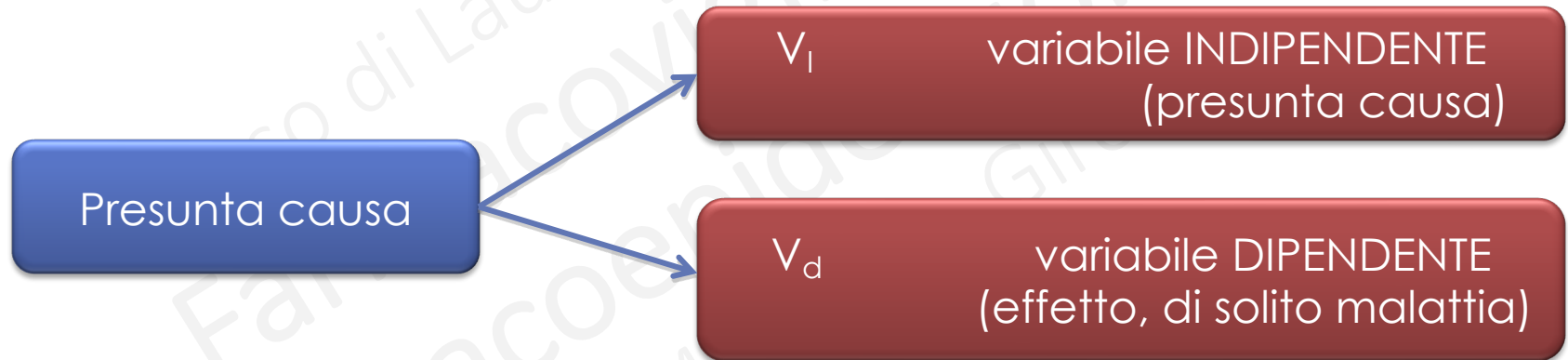


— ... e un'altra cosa, figlioli: state lontano dagli scienziati. Quelli vi fanno ammalare!



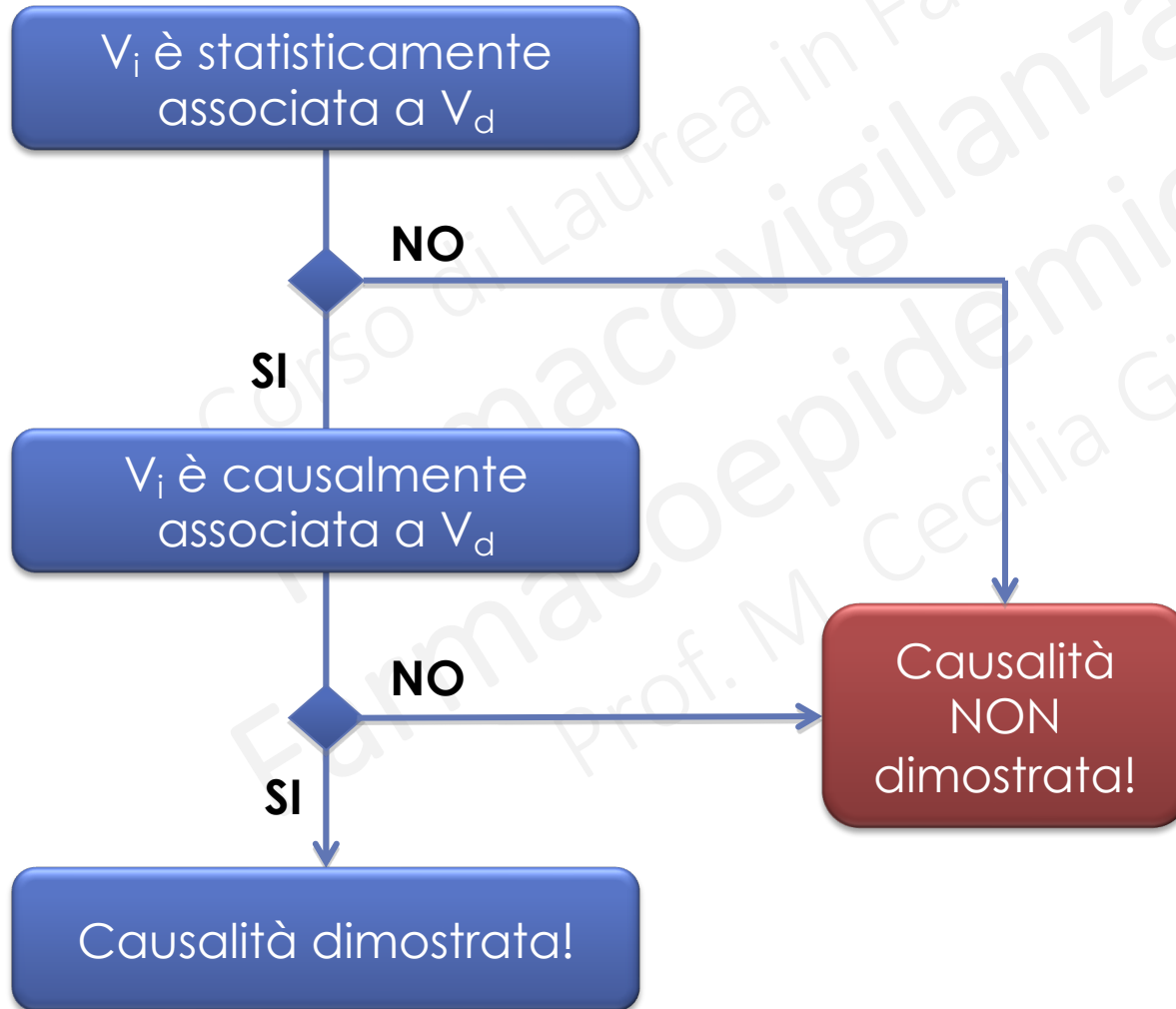
CAUSALITA vs CASUALE

Causalità: procedimento logico studi osservazionali



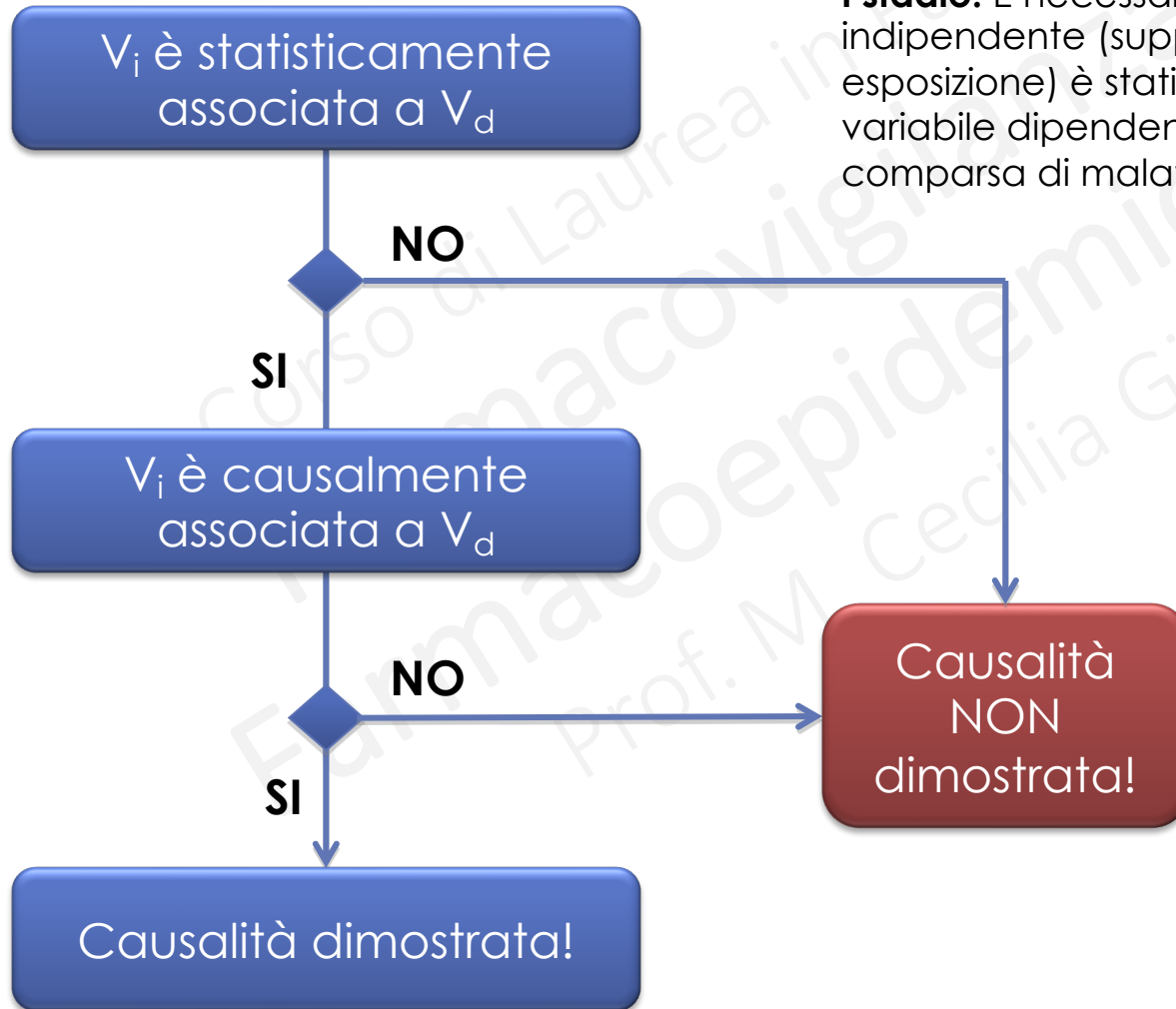


procedimento logico per gli studi osservazionali





procedimento logico per gli studi osservazionali

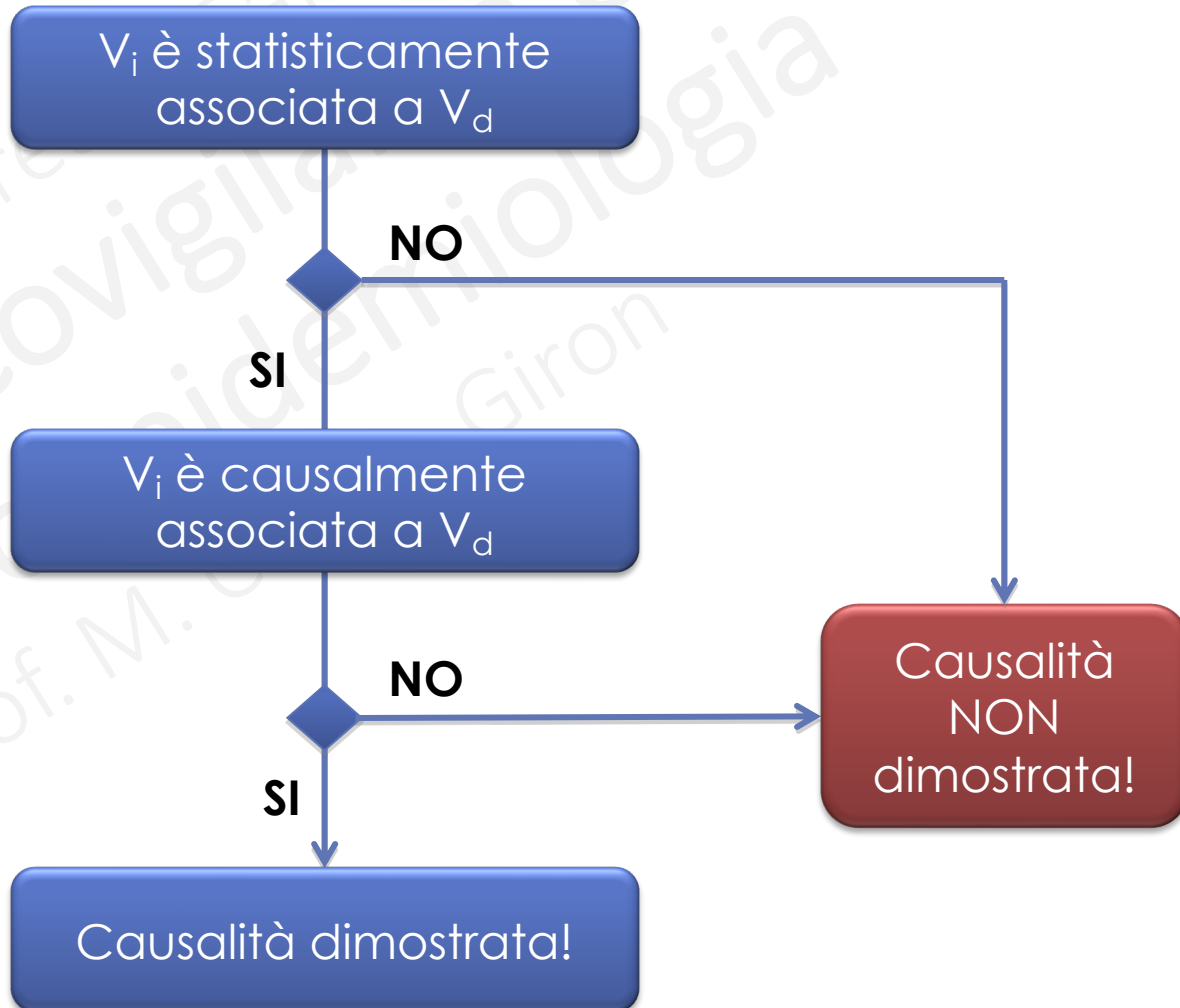


I stadio. È necessario accertare se la variabile indipendente (supposta causa o fattore di esposizione) è statisticamente associata con la variabile dipendente (l'«effetto», come ad es. la comparsa di malattia).



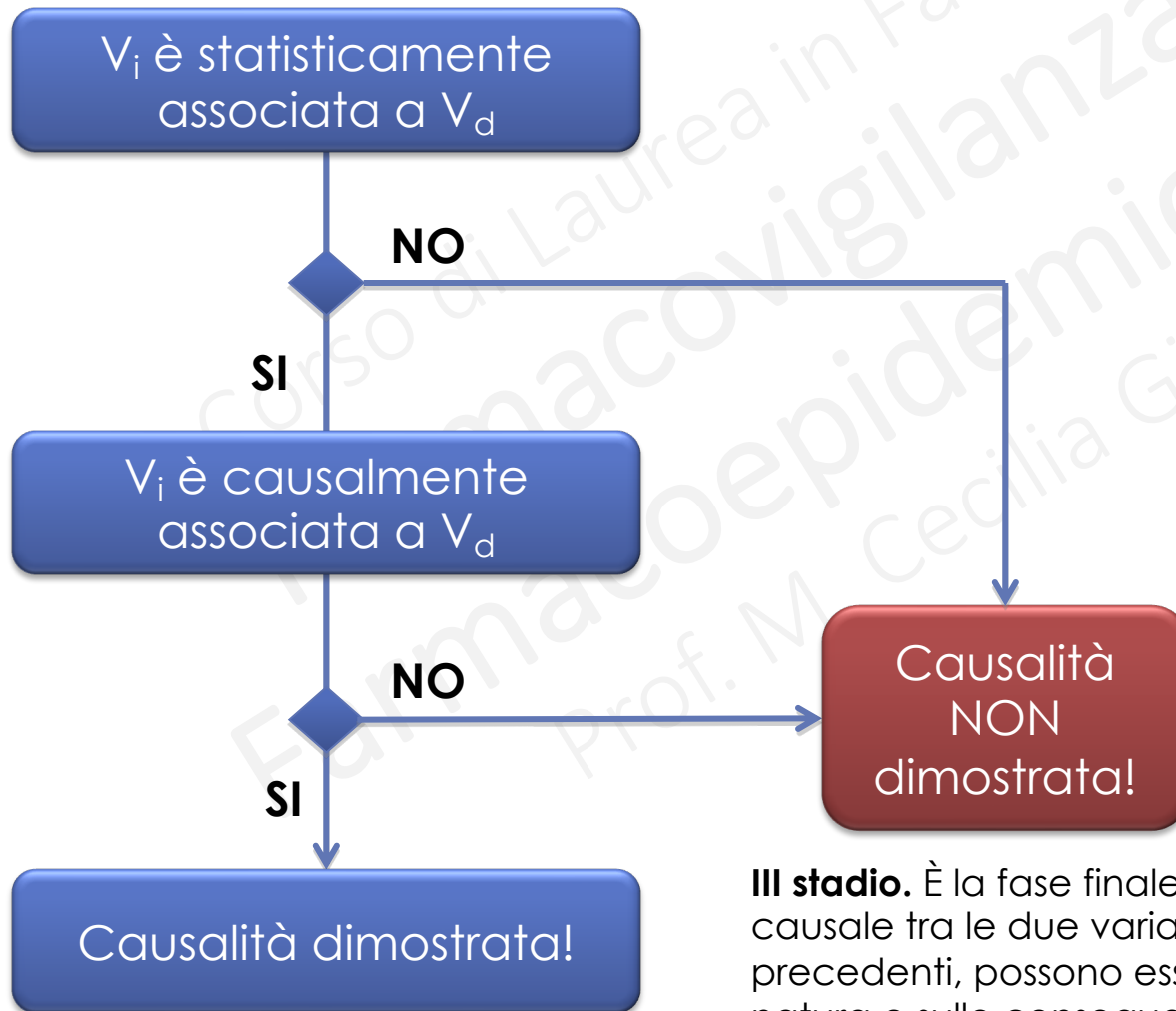
CAUSALITÀ: procedimento logico per gli studi osservazionali

Il stadio. Soltanto nel caso in cui le variabili siano statisticamente associate, è necessario accertare, mediante una serie di criteri accettati, che le due variabili siano causalmente associate





CAUSALITÀ: procedimento logico per gli studi osservazionali



III stadio. È la fase finale in cui, sulla base del rapporto causale tra le due variabili dimostrato nelle due fasi precedenti, possono essere effettuate elaborazioni sulla natura e sulle conseguenze dell'associazione, utilizzando modelli teorici, simulazioni, esperimenti in laboratorio o in campo ecc.

Bizzarra concatenazione di causa ed effetti



Il primo incontro col mio maestro è andato così:
"Conosce qualche proverbio giapponese?" mi chiede.

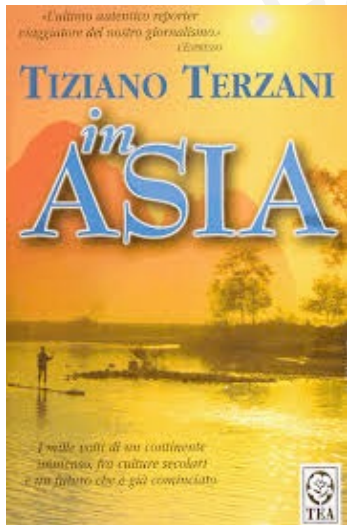
"No."

"Allora incominci a imparare questo: "Se tira vento i fabbricanti di tini di legno diventano ricchi"."

"E perché?"

"Semplice" dice lui. In Asia "Il vento solleva la polvere, la polvere acceca, i ciechi per guadagnarsi da vivere suonano strumenti a corda, le corde sono fatte di budella dei gatti, quanti più gatti vengono ammazzati tanti più topi scorrazzano per la città, i topi fanno buchi nei tini di legno, più tini di legno devono essere rifatti più i loro fabbricanti diventano ricchi. Logico, non le pare?"

da: Tiziano Terzani "In Asia", TEA, Tascabili degli Editori Associati, 1999, pag. 84.





TEST DEL CHI-QUADRATO

Maria Cecilia Giron

Dipartimento di Scienze del Farmaco

Edificio di Farmacologia

Email: cecilia.giron@unipd.it

Tel. 049-8275091



ASSOCIAZIONE: significato

Supponiamo che in un esperimento è stato estratto il 38% di *Haemophilus influenzae* in individui ammalati di polmonite

Domanda

L'isolamento del batterio *H. influenzae* e la polmonite sono due venti associati?

SI, secondo il senso comune

...perché nel senso comune due eventi sono associati se sono fra loro correlati (funzionalmente e temporalmente etc.)

NO, o meglio non si sa in epidemiologia

...perché i dati raccolti per l'esperimento sono insufficienti per rispondere alla domanda

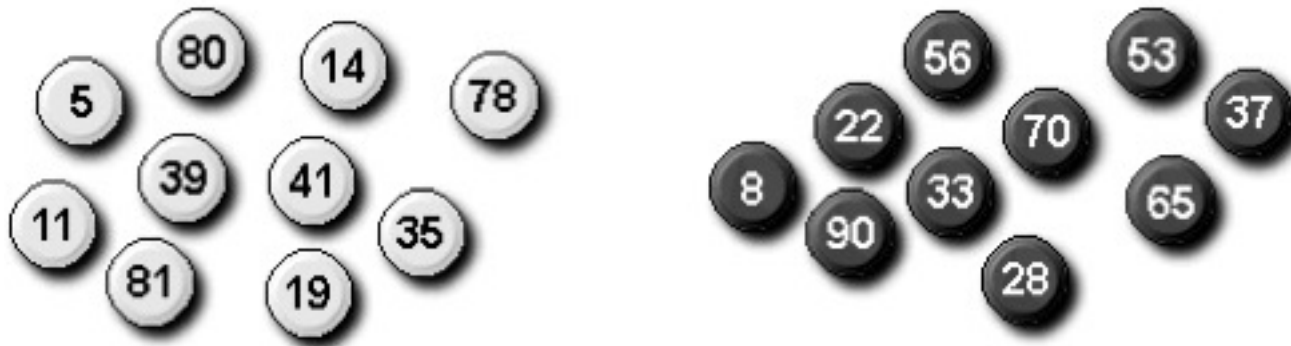


ASSOCIAZIONE: significato





ASSOCIAZIONE: significato



Supponiamo di avere due sacchetti di numeri dischetti numerati; in un sacchetto ci sono dischetti bianchi, nell'altro neri.

Immaginiamo ora di estrarre 10 numeri da ciascun sacchetto, e di ottenere il risultato della figura a lato. *

Fra i numeri bianchi, 7 sono dispari e 3 sono pari; fra i numeri neri, 5 sono dispari e 5 sono pari. Puoi forse concludere che i numeri bianchi sono «associati» alla disparità?

Mi spiego meglio: i risultati della estrazione (i dati!) sono sufficienti a concludere che l'ignota «popolazione» di numeri bianchi è composta prevalentemente da numeri dispari?

Ovviamente la risposta è «no», poiché i risultati dell'estrazione sono influenzati dal caso.

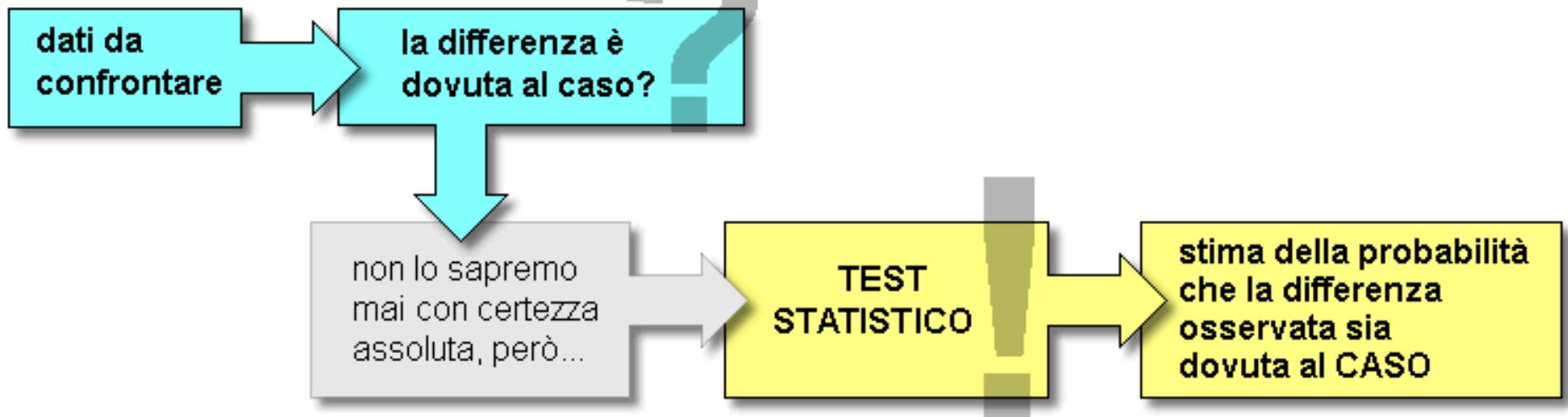


ASSOCIAZIONE: significato

Se è vero che il caso può influenzare i risultati,
come posso dimostrare
l'esistenza di associazioni ed
essere ragionevolmente sicuro che
le differenze osservate
NON SONO DOVUTE AL CASO?



ASSOCIAZIONE: significato





ASSOCIAZIONE: significato

Quando la statistica dimostra che «l'associazione fra due eventi NON è casuale» (esiste una «differenza significativa»), resta ancora da dimostrare che i due eventi siano legati da un rapporto causa-effetto.

**IN ALTRE PAROLE, «ASSOCIAZIONE»
NON È SINONIMO DI «CAUSALITÀ»!!!**



CAUSALITÀ vs CASUALE

CAUSA vs CASO

Somiglianza fonetica ed ortografica fra i termini «casuale» e «causale», ed anche «casualità» e «causalità».

I due termini hanno un significato assolutamente diverso!

Casuale = dovuto al caso;

Causale = legato da un rapporto di causa-effetto.



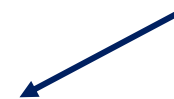
CAUSALITÀ vs CASUALE



I pazienti trattati con A sono guariti più spesso di quelli trattati con B

A è migliore di B non solo sul mio campione, ma anche sulla mia «popolazione di interesse» ?

inferenza





CAUSALITÀ vs CASUALE



Colpa dello sperimentatore !

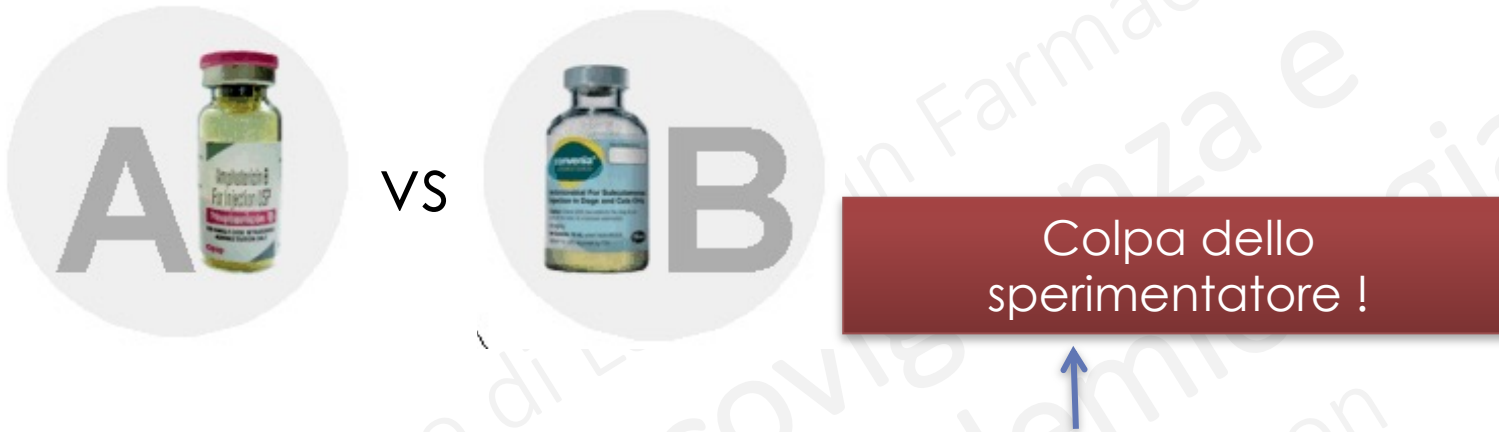
3 ipotesi perché A potrebbe apparire migliore di B

Il confronto fra i due gruppi è VIZIATO (**BIAS**)

La differenza è dovuta al CASO (variabilità biologica)

A è realmente più EFFICACE di B (questa ipotesi è accettabile solo dopo aver escluso le altre due)

CAUSALITÀ vs CASUALE



Il confronto fra i due gruppi è VIZIATO (**BIAS**)

Deriva da una "colpa" dello sperimentatore, che ha selezionato un **campione NON rappresentativo** della popolazione oggetto dello studio.

Un campione non rappresentativo si dice distorto (o affetto da «**bias**», dall'inglese bias=distorsione). Ovviamente un campione distorto fornisce risultati non affidabili.

CAMPIONE RAPPRESENTATIVO: è quello in cui la distribuzione degli individui rispetto al criterio scelto è molto simile alla distribuzione dell'intera popolazione rispetto allo stesso criterio.

BIAS: errore sistematico che conduce a risultati errati; è costituito dalla differenza tra ciò che si è misurato e ciò che si intendeva misurare. (*bias*=distorsione).

CAUSALITÀ vs CASUALE



La differenza è dovuta al CASO (variabilità biologica)

La seconda ipotesi chiama in causa il caso e la variabilità biologica.

Se eseguiamo di nuovo, però su altri soggetti, l'esperimento sull'efficacia dei due farmaci già fatto: siamo sicuro che otterremmo gli stessi risultati?

No, appunto perché nell'andamento dell'esperimento si inseriscono **fattori diversi, noto e ignoti**, dovuti appunto al caso.

Per isolare l'effetto del caso, e quindi per escludere l'ipotesi 2, si deve ricorrere alla statistica, **applicando un test che ci permetta di verificare la «significatività» dei nostri risultati.**

Confrontare due proporzioni o due percentuali: TEST CHI-QUADRATO



L'applicazione di un test di significatività statistica è un passo *indispensabile* nel confronto fra due gruppi o popolazioni riguardo a un parametro.

Confrontiamo due percentuali ottenute in un esperimento, allo scopo di verificare se la differenza fra tali percentuali è dovuta al caso oppure no. Se riusciamo a dimostrare che la **differenza non è dovuta al caso**, allora possiamo affermare che essa è «**statisticamente significativa**».

Il processo logico è molto semplice, si parte con una **sorta di pregiudizio**: **qualsiasi sia la differenza esistente tra le due percentuali da confrontare, inizialmente devi ritenere valida l'«ipotesi zero»**.

L'ipotesi zero (detta anche ipotesi nulla) afferma semplicemente che **la differenza osservata - di qualsiasi entità essa sia - è dovuta al caso**.

Questa ipotesi (che può essere vera o falsa) verrà accettata oppure rifiutata sulla base del risultato di un appropriato test statistico. Nel confronto di due percentuali o di due proporzioni il test appropriato è, appunto, il test del chi-quadrato.

In sintesi:

Le due percentuali sono DAVVERO DIVERSE
oppure la differenza è dovuta al CASO?

Confrontare due proporzioni o due percentuali: TEST CHI-QUADRATO



Le due percentuali sono DAVVERO DIVERSE
oppure la differenza è dovuta al CASO?

Ipotesi zero: la differenza è dovuta al caso

Accettare o rifiutare l'ipotesi zero?

eseguo il TEST del χ^2

Confrontare due proporzioni o due percentuali: TEST del CHI-QUADRATO



TEST del χ^2

% guariti: $52/62 = 0,84 = 84\%$

Trattamento ↓	Esito →	guariti	non-guariti	totale
xmicina		52 a	10 b	62
streptomicina		40 c	21 d	61
totale		92	31	123

% guariti: $40/61 = 0,66 = 66\%$

% guariti TOTALI: $92/123 = 0,748 = 74,8\%$

Confrontare due proporzioni o due percentuali: TEST del CHI-QUADRATO



TEST del χ^2

% guariti: $52/62 = 0,84 = 84\%$

Trattamento	Esito	guariti	non-guariti	totale
xmicina		52 <i>a</i>	10 <i>b</i>	62
streptomicina		40 <i>c</i>	21 <i>d</i>	61
totale		92	31	123

% guariti: $40/61 = 0,66 = 66\%$

	xmicina	streptom.	totale
guariti	52 <i>a</i>	40 <i>b</i>	92
non-guariti	10 <i>c</i>	21 <i>d</i>	31
totale	62	61	123

Variabile indipendente (trattamento xmicina/streptomicina) sulle righe

Variabile dipendente (nel nostro caso: guarito/non guarito) sulle colonne

Confrontare due proporzioni o due percentuali: TEST del CHI-QUADRATO



% guariti: $52/62 = 0,84 = 84\%$

Trattamento	Esito	guariti	non-guariti	totale
xmicina		52 <i>a</i>	10 <i>b</i>	62
streptomicina		40 <i>c</i>	21 <i>d</i>	61
totale		92	31	123

% guariti: $40/61 = 0,66 = 66\%$

Supponiamo che, in realtà, **NON** esistano differenze nell'efficacia dei due trattamenti.

Che probabilità c'è di osservare - in uno studio di dimensioni simili a questo - differenze nell'efficacia dei due antibiotici uguali o superiori a quelle che abbiamo osservato?

Confrontare due proporzioni o due percentuali: TEST del CHI-QUADRATO



Quanto i dati ottenuti si discostano dai dati che sarebbe lecito attendersi:

- ✓ se i trattamenti avessero la stessa efficacia
- ✓ se i dati fossero influenzati soltanto dalla **variazione casuale**?

se $x_{micina} = streptomicina$ (ipotesi zero!), è possibile ottenere risultati simili a quelli che abbiamo osservato?

La risposta è **sì**, perché i tuoi dati potrebbero essere stati influenzati dal caso.

Qual è la probabilità di osservare una differenza uguale o superiore a quella che hai ottenuto nel tuo esperimento?

Confrontare due proporzioni o due percentuali: TEST del CHI-QUADRATO



% guariti: $52/62 = 0,84 = 84\%$

Trattamento ↓	Esito →	guariti	non-guariti	totale
xmicina		52 <i>a</i>	10 <i>b</i>	62
streptomicina		40 <i>c</i>	21 <i>d</i>	61
totale		92	31	123

% guariti: $40/61 = 0,66 = 66\%$

Guariti totali = $(52 + 40) / 123 * 100 = 74,8\%$

**complessivamente (indipendentemente dall'antibiotico usato)
il trattamento è risultato efficace nel 74,8% dei casi**

Trattamento ↓	Esito →	guariti	non-guariti	totale
xmicina		46.37 <i>a</i>	15.63 <i>b</i>	62.00
streptomicina		45.63 <i>c</i>	15.37 <i>d</i>	61.00
totale		92.00	31.00	123.00

$$b = 62 - 46,37 = 15,63$$

$$d = 61 - 45,63 = 15,37$$

% di guarigione con xmicina = $62 * 74,8 / 100 = 46,37$
 % di guarigione con streptomicina = $61 * 74,8 / 100 = 45,63$

Confrontare due proporzioni o due percentuali: TEST del CHI-QUADRATO



Trattamento	Esito	guariti	non-guariti	totale
xmicina		52 a	10 b	62
streptomicina		40 c	21 d	61
totale		92	31	123

% guariti: $52/62 = 0,84 = 84\%$

Guariti totali = $(52 + 40) / 123 * 100 = 74,8\%$

% guariti: $40/61 = 0,66 = 66\%$

Trattamento	Esito	guariti	non-guariti	totale
xmicina		46.37 a	15.63 b	62.00
streptomicina		45.63 c	15.37 d	61.00
totale		92.00	31.00	123.00

Guariti con xmicina = $62 * 74,8 / 100 = 46,37$
 Guariti con streptomicina = $61 * 74,8 / 100 = 45,63$

$b = 62 - 46,37 = 15,63$
 $d = 61 - 45,63 = 15,37$

Il valore del chi-quadrato quantifica la differenza fra i dati osservati e quelli attesi

$$\frac{(\text{dato osservato} - \text{dato atteso})^2}{\text{dato atteso}}$$

$$\chi^2 = \frac{(52 - 46.37)^2}{46.37} + \frac{(10 - 15.63)^2}{15.63} + \frac{(40 - 45.63)^2}{45.63} + \frac{(21 - 15.37)^2}{15.37} = 5.46$$

Confrontare due proporzioni o due percentuali: TEST del CHI-QUADRATO



Il valore del chi-quadrato
quantifica la differenza fra i
dati osservati e quelli attesi

$$\frac{(\text{dato osservato} - \text{dato atteso})^2}{\text{dato atteso}}$$

$$\chi^2 = \frac{(52 - 46.37)^2}{46.37} + \frac{(10 - 15.63)^2}{15.63} + \frac{(40 - 45.63)^2}{45.63} + \frac{(21 - 15.37)^2}{15.37} = 5.46$$

Il valore del chi-quadrato **aumenta** con
l'aumentare della differenza dei dati posti a
raffronto!!!

Confrontare due proporzioni o due percentuali: TEST del CHI-QUADRATO



$$\chi^2 = \frac{(52 - 46.37)^2}{46.37} + \frac{(10 - 15.63)^2}{15.63} + \frac{(40 - 45.63)^2}{45.63} + \frac{(21 - 15.37)^2}{15.37} = 5.46$$

Il valore del chi-quadrato aumenta con l'aumentare della differenza dei dati posti a raffronto!!!

Tabella dei valori di χ^2

Gradi di libertà	Probabilità	
	5%	1%
1	3.841	6.635
2	5.991	9.210
3	7.815	11.345
4	9.488	13.277
5	11.070	15.086
6	12.592	16.812
7	14.067	18.475
ecc.

il grado di libertà è uguale a
(n. di righe-1)*(n. di colonne-1)

Quindi:

$(2-1) * (2-1) = 1$ grado di libertà

Se esso supera certi valori prefissati la differenza viene ritenuta **significativa**; in caso contrario, non si può affermare l'esistenza di una significativa differenza tra i due eventi considerati.

Confrontare due proporzioni o due percentuali: TEST del CHI-QUADRATO



Ora, confrontando il nostro valore di chi-quadrato (**5.46**) con quelli tabulati, vediamo che è **>3.841 e <6.635**. Quindi possiamo ritenere che la differenza fra i due gruppi **sia significativa al livello di probabilità 5%** ma non al livello di probabilità 1%.

Si può quindi affermare che la differenza tra soggetti trattati con xmicina e quelli trattati con streptomina è statisticamente significativa al livello di probabilità 5%.

In altre parole, se ammettiamo che i due antibiotici abbiano pari efficacia e ripetendo l'esperimento infinite volte, potremo osservare piuttosto raramente (ossia 5 volte su 100 o meno!) dati simili a quelli ottenuti oppure ancor più favorevoli a xmicina.

In sostanza: in base ai risultati del test del chi-quadrato, l'affermazione «xmicina è più efficace di streptomina» ha il 95% di probabilità di essere vera (e quindi ha il 5% di probabilità di essere falsa).

Tabella dei valori di χ^2

Gradi di libertà	Probabilità	
	5%	1%
1	3.841	6.635
2	5.991	9.210
3	7.815	11.345
4	9.488	13.277
5	11.070	15.086
6	12.592	16.812
7	14.067	18.475
ecc.

In base ai risultati del test del chi-quadrato, possiamo concludere che xmicina è più attiva di streptomina (**$P < 0,05$**) dove il valore di **P** indica la probabilità di respingere l'ipotesi zero



TEST DEL CHI-QUADRATO: esempi

Maria Cecilia Giron

Dipartimento di Scienze del Farmaco

Edificio di Farmacologia

Email: cecilia.giron@unipd.it

Tel. 049-8275091



Tutti i test statistici assumono inizialmente la cosiddetta
ipotesi nulla H_0
cioè

Quando si effettua il confronto fra 2 o più gruppi, l'ipotesi
nulla prevede che **non esista alcuna differenza** tra i
gruppi riguardo al parametro considerato

Gruppo A

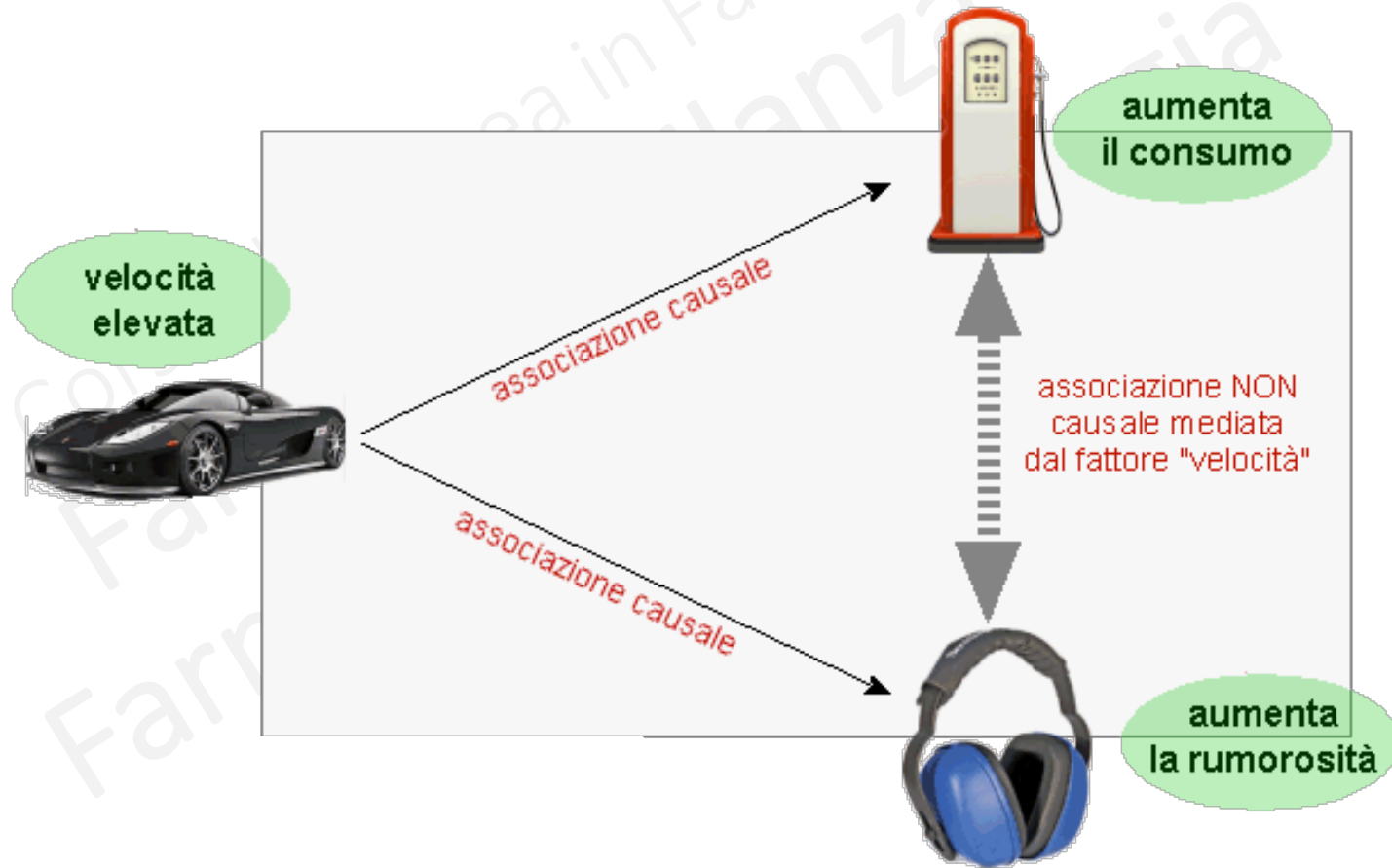
=

Gruppo B

Accetto o rifiuto l'ipotesi H_0 ?



Associazione non causale





C'è sempre il rischio che la decisione di rifiutare l'ipotesi H_0 **sia errata**

La misura di questo rischio è definita come

LIVELLO DI SIGNIFICATIVITÀ

- ✓ Può essere scelto arbitrariamente
- ✓ In genere si sceglie pari a 0,05 (5%)
- ✓ Il valore più basso al quale H_0 può essere respinta viene definito valore P

Confrontare due proporzioni o due percentuali: TEST del CHI-QUADRATO



Abbiamo trattato un buon numero di casi di tonsillite con il nuovo antibiotico (nome di fantasia, "mycomicina") ed un numero simile di casi analoghi con un antibiotico classico, la tetraciclina. In totale, **140 soggetti** dei quali **70 (scelti a caso) sono stati trattati con "mycomicina" ed i rimanenti 70 (sempre scelti a caso) con tetraciclina.**

Il fatto di aver scelto a caso gli individui ha ridotto la possibilità di assegnare ad un gruppo piuttosto che all'altro soggetti "particolari" (es. più giovani, con infezione più lieve ecc.). Infatti, i due gruppi risultano fra loro piuttosto simili.

Fra gli individui trattati con "mycomicina" ne sono guariti 62; fra quelli trattati con tetraciclina ne sono guariti 57.

Confrontare due proporzioni o due percentuali: TEST del CHI-QUADRATO



In totale, **140 soggetti** dei quali **70 (scelti a caso)** sono stati trattati con "mycomicina" ed i rimanenti **70 (sempre scelti a caso)** con tetraciclina. Fra gli individui trattati con "mycomicina" ne sono guariti **62**; fra quelli trattati con tetraciclina ne sono guariti **57**.

Trattamento	guariti	non-guariti	Totale
mycomicina	62	8	70
tetraciclina	57	13	70
Totale	119	21	140

Confrontare due proporzioni o due percentuali: TEST del CHI-QUADRATO



In totale, **140 soggetti** dei quali **70 (scelti a caso)** sono stati trattati con "mycomicina" ed i rimanenti **70 (sempre scelti a caso)** con tetraciclina. Fra gli individui trattati con "mycomicina" ne sono guariti **62**; fra quelli trattati con tetraciclina ne sono guariti **57**.

Trattamento	guariti	non-guariti	Totale
mycomicina	62 (88,6%)	8	70
tetraciclina	57 (81,4%)	13	70
Totale	119	21	140

Confrontare due proporzioni o due percentuali: TEST del CHI-QUADRATO



Tattamento	guariti	non-guariti	Totale
mycomicina	62 (88,6%)	8	70
tetraciclina	57 (81,4%)	13	70
Totale	119	21	140

$$\chi^2 = 0,896$$

Non esistono differenze nell'efficacia di mycomicina rispetto a quella di tetraciclina !

Questa conclusione è valida al livello di probabilità >5%.

Valori di χ^2

Gradi di libertà	Probabilità	
	5%	1%
1	3.841	6.635
2	5.991	9.210
3	7.815	11.345
4	9.488	13.277
5	11.070	15.086
6	12.592	16.812
7	14.067	18.475
...

Confrontare due proporzioni o due percentuali: TEST del CHI-QUADRATO



Quale vacanza preferite?

Valori di χ^2

Soggetti	mare	montagna	Totale
uomini	209	280	489
donne	225	248	473
Totale	434	528	962

Gradi di libertà	Probabilità	
	5%	1%
1	3.841	6.635
2	5.991	9.210
3	7.815	11.345
4	9.488	13.277
5	11.070	15.086
6	12.592	16.812
7	14.067	18.475
...

$$\chi^2 = 2.2638$$

$$P=0,132431$$

non significativo!

Confrontare due proporzioni o due percentuali: TEST del CHI-QUADRATO



Quale animale preferite?

Soggetti	gatto	cane	Totale
uomini	207	282	489
donne	231	242	473
Totale	438	524	962

Valori di χ^2

Gradi di libertà	Probabilità	
	50%	1%
1	3.841	6.635
2	5.991	9.210
3	7.815	11.345
4	9.488	13.277
5	11.070	15.086
6	12.592	16.812
7	14.067	18.475
...

$$\chi^2 = 4,1035$$

$$P=0,042794$$

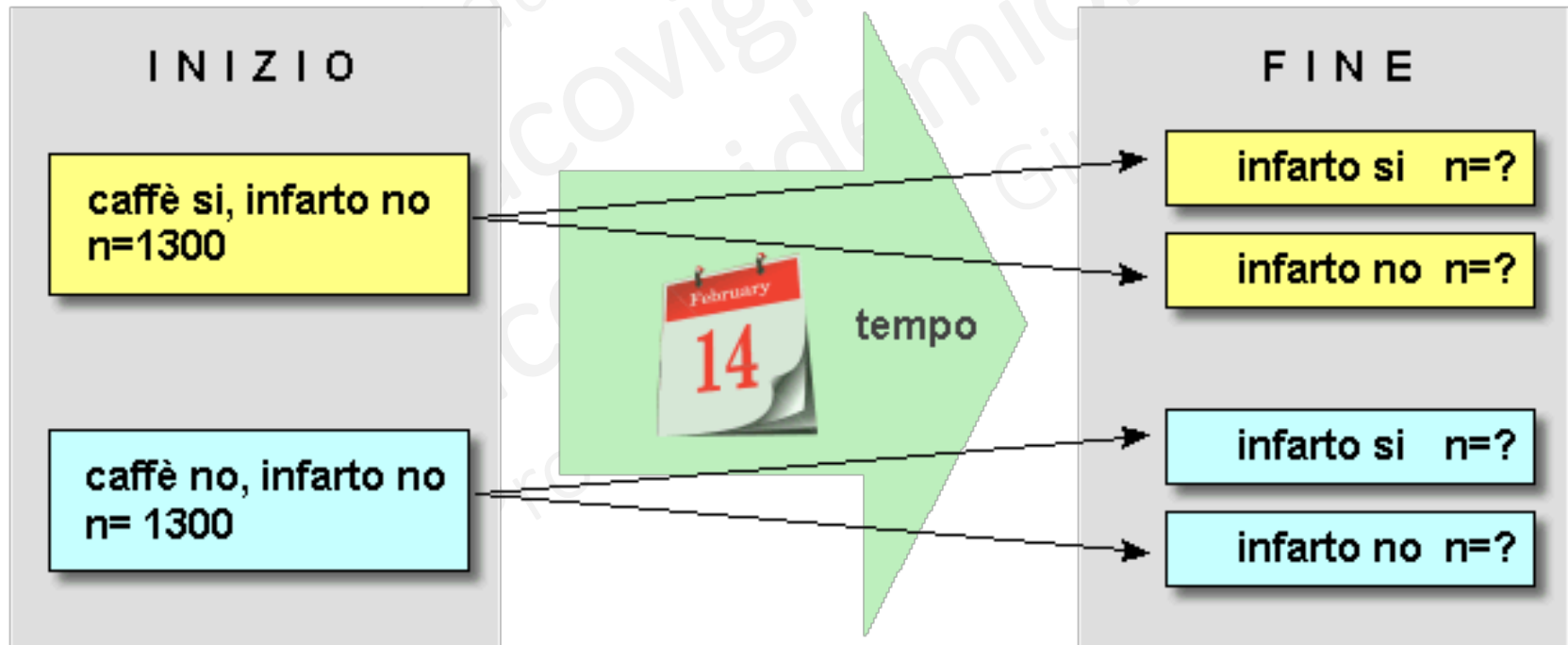
è significativo!



Ipotesi H_0 e significatività

Il consumo di caffè è un fattore di rischio di infarto del miocardio
Campione di maschi di età compresa fra 39 e 59 anni pari a 2600 persone
seguiti per 8 anni

Disegno dello studio





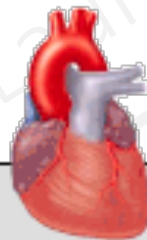
Ipotesi H_0 e significatività

Il consumo di caffè è un fattore di rischio di infarto del miocardio
Campione di maschi di età compresa fra 39 e 59 anni pari a 2600 persone
seguiti per 8 anni

Tabella 1

[dati fittizi]

		infarto		totale
		si	no	
caffè	si	215	1085	1300
	no	80	1220	1300
totale		295	2305	2600



fra i 1300 bevitori di caffè, se ne sono ammalati 215 (16.5%)

fra i 1300 non bevitori di caffè, se ne sono ammalati 80 (6.1%)



Ipotesi H_0 e significatività

Il consumo di caffè è un fattore di rischio di infarto del miocardio
Campione di maschi di età compresa fra 39 e 59 anni pari a 2600 persone
seguiti per 8 anni

Valori di χ^2

Verifica dell'ipotesi in uno studio retrospettivo
(con l'utilizzo del test del chi-quadrato)

	ammalati	sani	tot.
esposti	215	1085	1300
non-esposti	80	1220	1300
tot.	295	2305	2600

Gradi di libertà	Probabilità	
	5%	1%
1	3.841	6.635
2	5.991	9.210
3	7.815	11.345
4	9.488	13.277
5	11.070	15.086
6	12.592	16.812
7	14.067	18.475
...

	proporz.	%
proporzione di esposti negli ammalati:	$215/(215+80)$	0.73 72.88
proporzione di esposti nei sani:	$1085/(1085+1220)$	0.47 47.07

chi - q u a d r a t o:	69.69	→ differenza SIGNIFICATIVA (prob. 1%)
p:	0.0000	
chi - q u a d r a t o (corretto Yates):	68.66	→ differenza SIGNIFICATIVA (prob. 1%)
p:	0.0000	
o d d s r a t i o:	3.02	



TABLE 1 Baseline characteristics of the matched ulcerative colitis and Crohn's disease patient cohort at index

	Ulcerative colitis			Crohn's disease		
	Cases	Controls	Statistic	Cases	Controls	Statistic
Total, n	925	4106		551	2077	
Female, n (%)	458 (49.5)	1995 (48.6)	$\chi^2, P = 0.611$	301 (58.9)	1252 (60.3)	$\chi^2, P = 0.570$
Age at outcome, mean (SD)	68.9 (13.9)	66.3 (14.2)	$t, P < 0.001$	65.1 (15.3)	62.4 (15.7)	$t, P < 0.001$
Current smoker, n (%)	49 (5.3)	249 (6.1)	$\chi^2, P = 0.044$	75 (14.7)	375 (18.1)	$\chi^2, P = 0.001$
Risk-cardiovascular disease ^a , n (%)	257 (27.8)	756 (18.4)	$\chi^2, P < 0.001$	126 (24.7)	325 (15.5)	$\chi^2, P < 0.001$
Diabetes mellitus (type 1 or 2), n (%)	181 (19.6)	509 (12.4)	$\chi^2, P < 0.001$	90 (17.6)	202 (9.7)	$\chi^2, P < 0.001$
Disease duration ^b , mean (SD)	13.0 (11.9)	10.8 (11.6)	$t, P < 0.001$	13.8 (11.7)	11.0 (11.1)	$t, P < 0.001$
Flare at outcome, n (%)	32 (3.5)	88 (2.1)	$\chi^2, P = 0.018$	26 (5.1)	108 (5.2)	$\chi^2, P = 0.010$
5-ASA user status ^c						
≤30 d, n (%)	562 (60.8)	2380 (58.0)		298 (58.3)	1141 (54.9)	
31-≤180 d, n (%)	300 (32.4)	1457 (35.5)		172 (33.7)	765 (36.8)	
181-≤365 d, n (%)	63 (6.8)	269 (6.6)		41 (8.0)	171 (8.2)	
Use of concomitant nephrotoxic drugs ^d , n (%)	806 (87.1)	2940 (71.6)	$\chi^2, P < 0.001$	454 (88.8)	1560 (75.1)	$\chi^2, P < 0.001$
Use of thiopurine ^d , n (%)	99 (10.7)	393 (9.6)	$\chi^2, P = 0.295$	81 (15.8)	329 (15.8)	$\chi^2, P = 0.995$
Use of glucocorticosteroids ^d , n (%)	193 (20.9)	647 (15.8)	$\chi^2, P < 0.001$	89 (17.4)	318 (15.3)	$\chi^2, P = 0.241$
Use of methotrexate ^d , n (%)	16 (1.7)	36 (0.9)	$\chi^2, P = 0.020$	13 (2.5)	43 (2.1)	$\chi^2, P = 0.510$
Use of ciclosporin ^d , n (%)	^e	6 (0.1)		0	0	

^aHypertension, coronary heart disease, heart failure, stroke, transient ischemic attack, peripheral vascular disease, aortic aneurysms.

^bFrom first IBD diagnosis to index (nephrotoxic event/ matched outcome date).

^cTime of the last 5-ASA prescription prior to index.

^d ≤ 90 d prior to index.

^eSmall patient numbers (n = 1-5) masked.