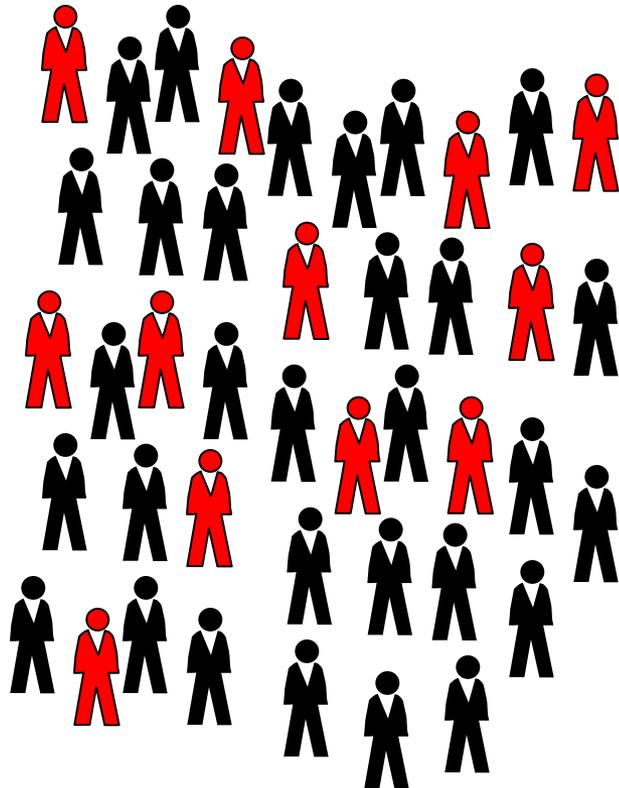


# La relazione causa-effetto

- valutare il rischio in epidemiologia
- associazione statistica e rapporto causa-effetto
- gli studi analitici: generalità

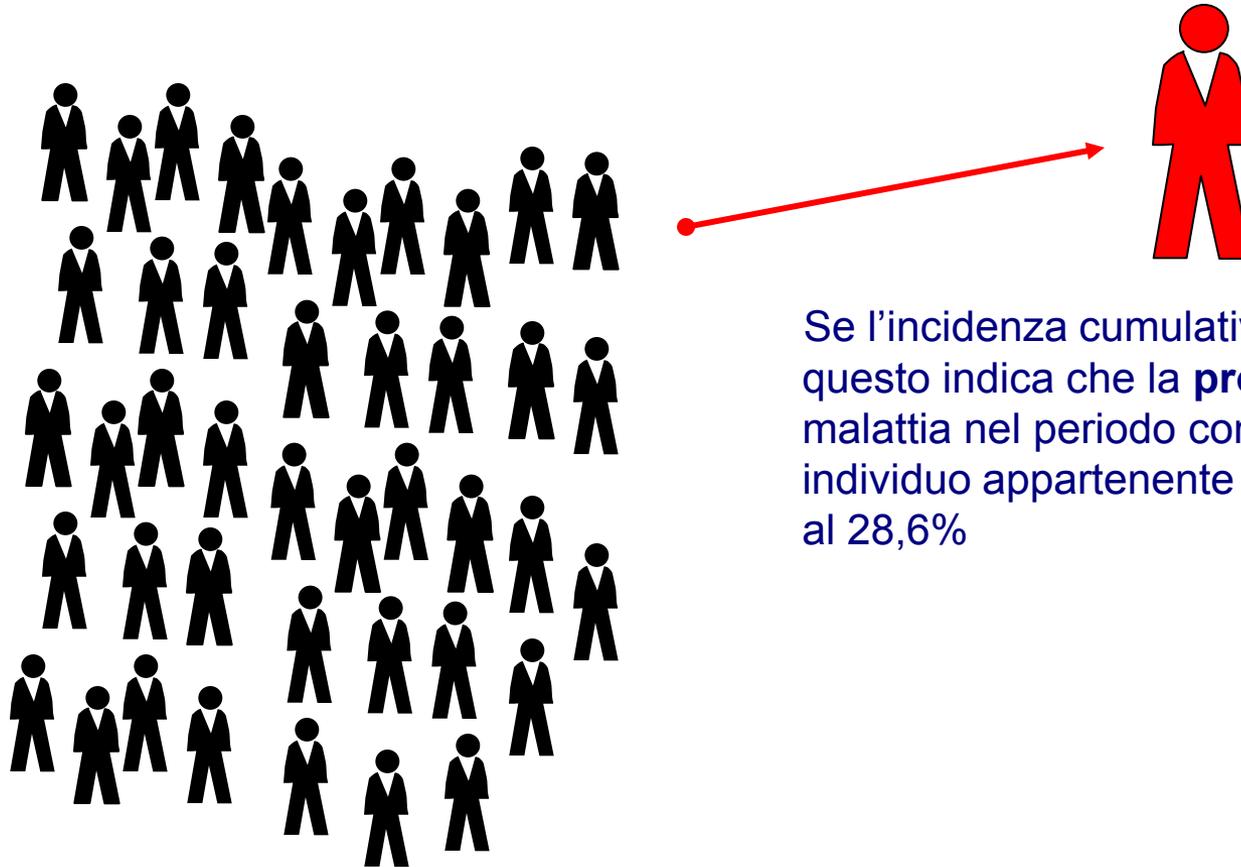
# Misurare l'incidenza di malattia in un gruppo di popolazione...



Misurare l'incidenza di malattia equivale a valutare la probabilità (ossia il rischio) di avere nuovi casi di malattia nella popolazione nel periodo prescelto

$$12/42 * 100 \Rightarrow 28,6\%$$

# Misurare l'incidenza di malattia in un gruppo di popolazione...



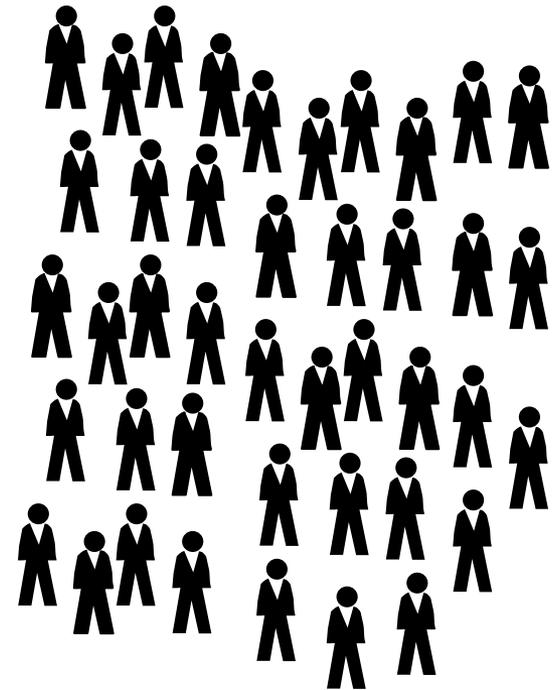
Se l'incidenza cumulativa risulta pari a 28,6%, questo indica che la **probabilità** di sviluppare la malattia nel periodo considerato, per ciascun individuo appartenente alla popolazione, è pari al 28,6%

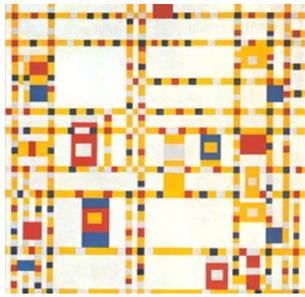
# Valutare il rischio individuale...

*Ma gli individui che compongono questa popolazione sono tutti uguali?*

*La probabilità che ha ciascun individuo di contrarre una malattia è effettivamente pari all'incidenza cumulativa calcolata su tutta la popolazione?*

**La probabilità di contrarre una malattia per un individuo può aumentare o diminuire in base alla presenza/assenza di particolari condizioni**

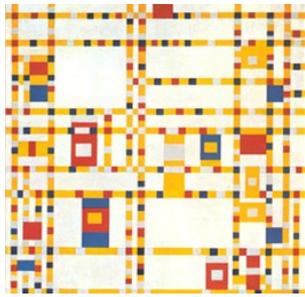




# Cos'è un fattore di rischio?

È un fattore che aumenta la probabilità dei soggetti a esso esposti di contrarre una determinata malattia.

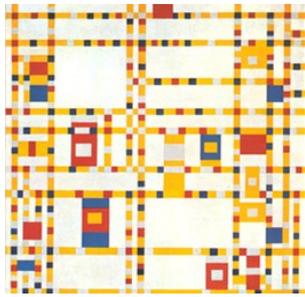
Può essere rappresentato da una condizione geneticamente determinata, da una abitudine personale, da un particolare stile di vita, da un evento accidentale, ecc.



# L'epidemiologia analitica...

indaga le cause o i fattori che determinano l'insorgere di una malattia o ne influenzano la diffusione

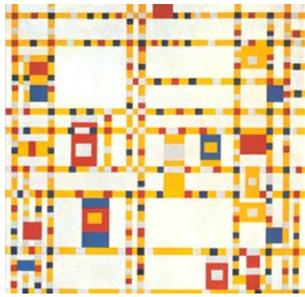
... ovvero indaga la **relazione causa-effetto** fra fattori di rischio e patologie



# Indagare una relazione causa-effetto

Fra un evento variabile (*causa o fattore di rischio*) e una malattia (*o una determinata condizione in studio*) può esistere una certa associazione statistica

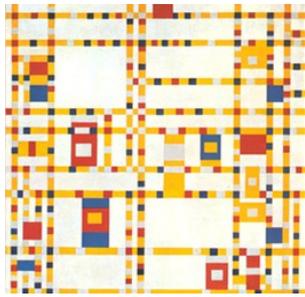
Per *associazione* intendiamo il grado di dipendenza statistica tra due o più eventi variabili



# Indagare una relazione causa-effetto

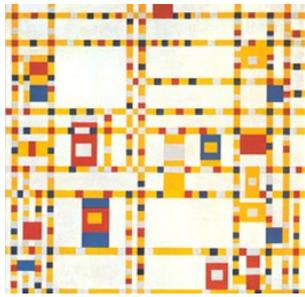
*post hoc = propter hoc*

?



# Associazione fra due eventi

- causale (o eziologica)
- indiretta (o secondaria)
- spuria (o non causale)

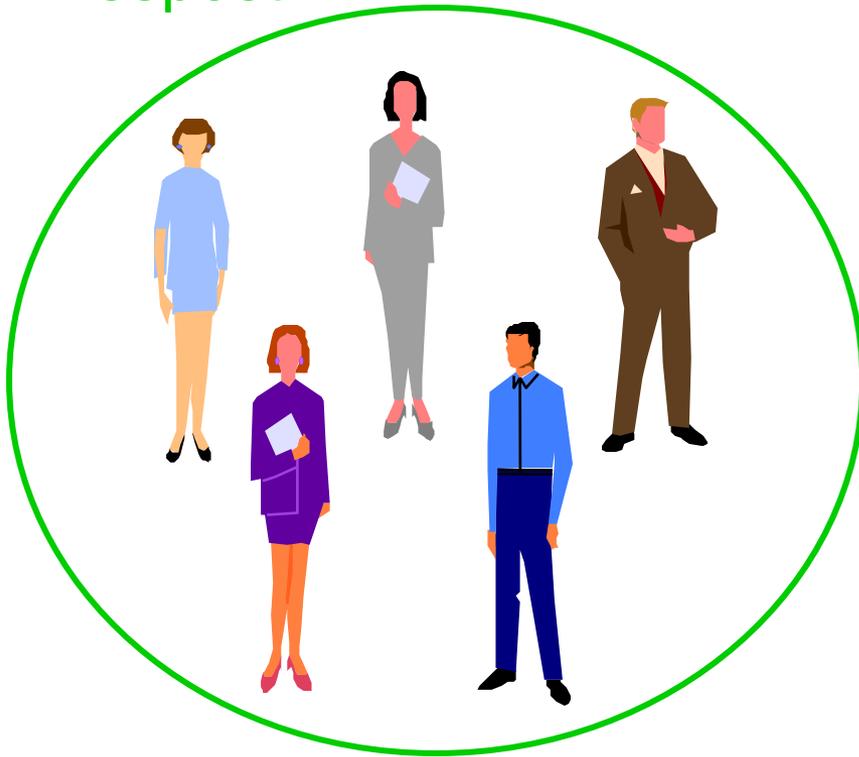


# Come valutare un'associazione causale

- forza dell'associazione
- consistenza dell'osservazione
- relazione temporale
- plausibilità biologica
- relazione di dipendenza dose-risposta

# Schema generale di uno studio analitico

esposti



non esposti

# Schema generale di uno studio analitico

esposti

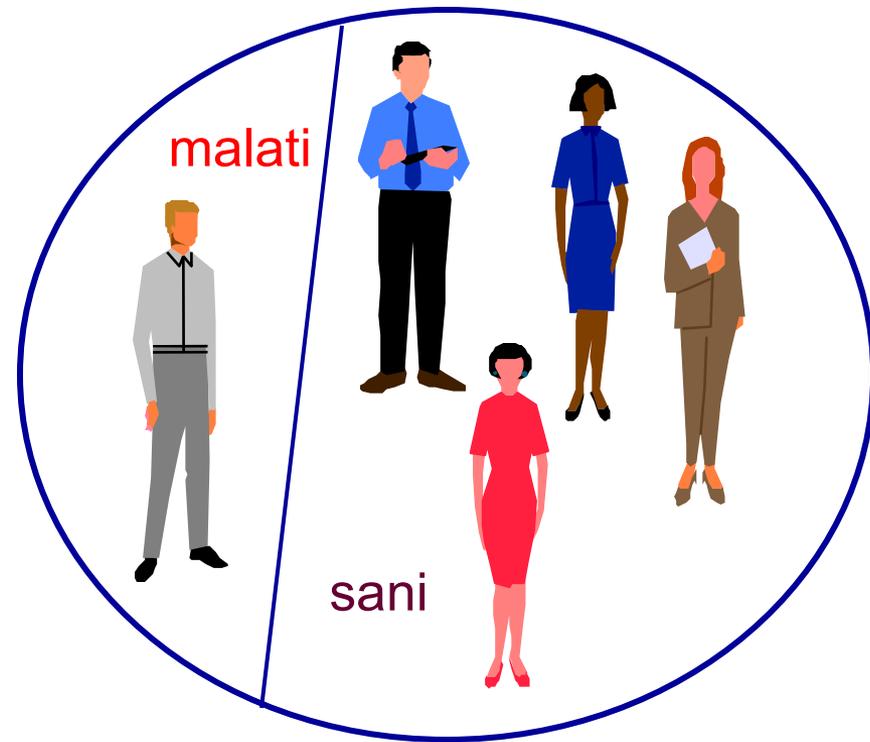
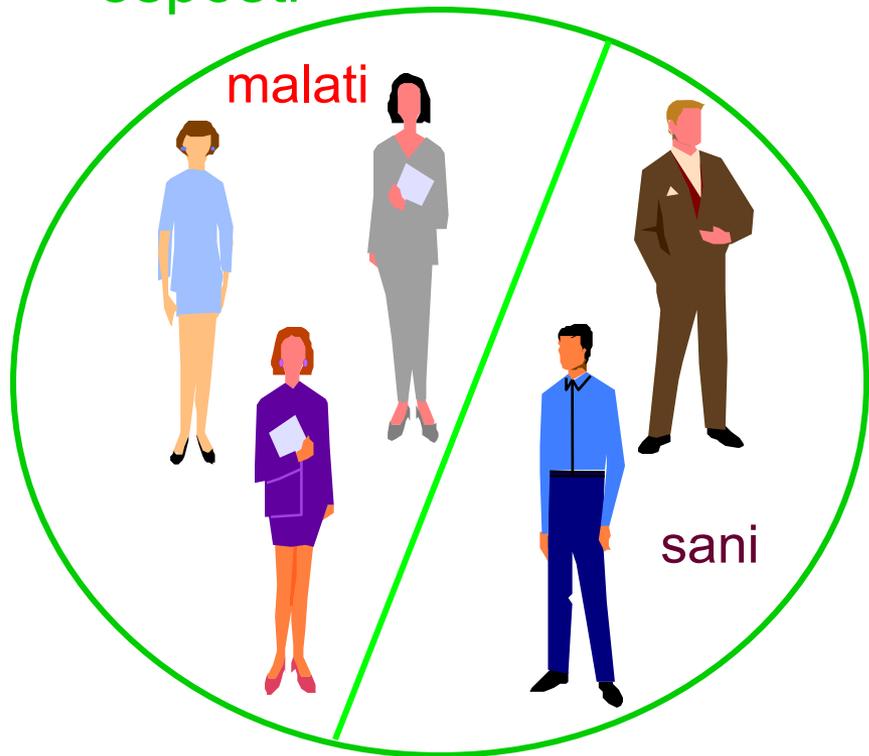
malati

sani

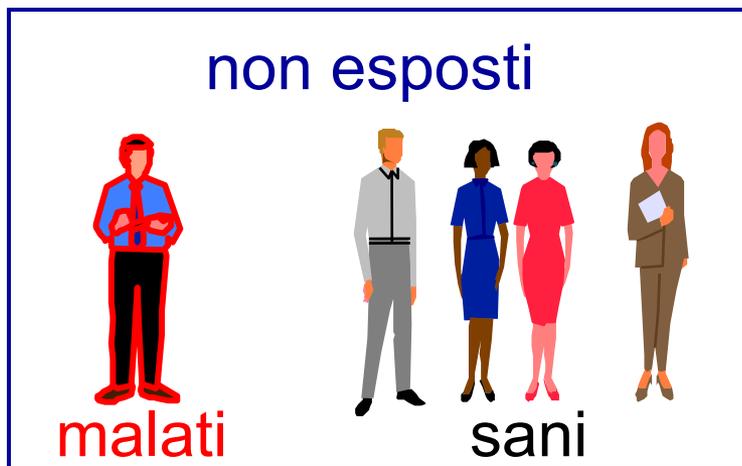
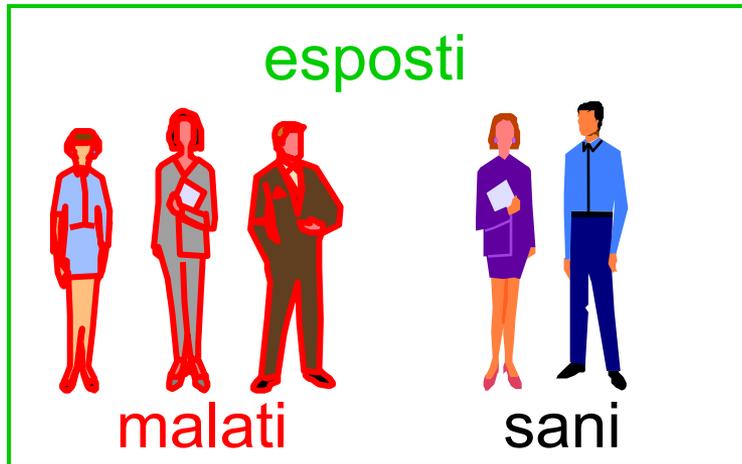
malati

sani

non esposti

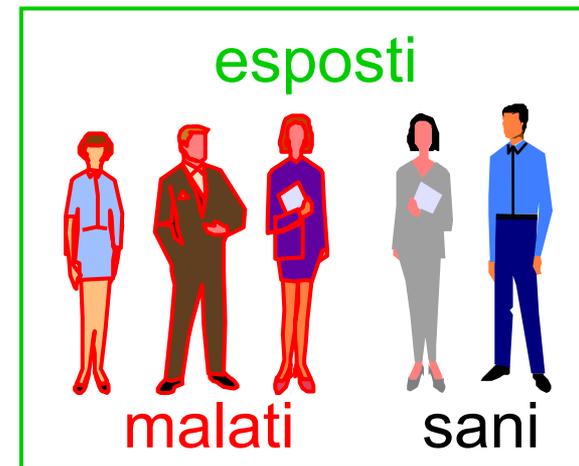


# La tabella 2x2

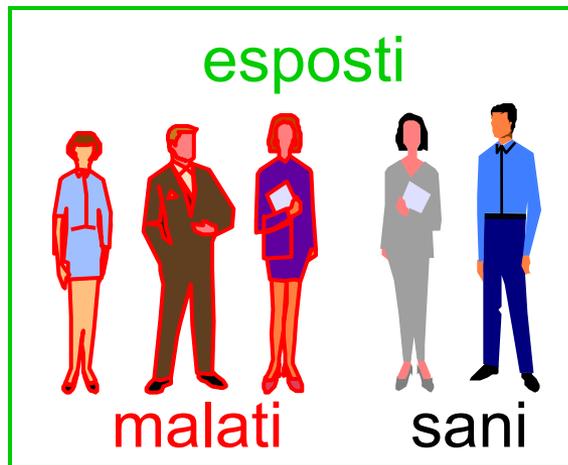


	M+	M-	
E+	3 <i>a</i>	2 <i>b</i>	5
E-	1 <i>c</i>	4 <i>d</i>	5
	4	6	10

# Un modello di studio: lo studio di coorte

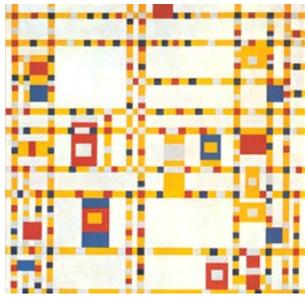


# Un modello di studio: lo studio di coorte



	M+	M-	
E+	3 <i>a</i>	2 <i>b</i>	5
E-	1 <i>c</i>	4 <i>d</i>	5
	4	6	10

$$\text{Rischio Relativo} = \frac{I_{E+}}{I_{E-}} = \frac{a/(a+b)}{c/(c+d)} = \frac{3/5}{1/5} = 3$$



# Interpretare l'RR

- **RR >1**
- **RR =1**
- **RR <1**

# Significatività dell'RR

	M+	M-	
E+	200 <i>a</i>	150 <i>b</i>	350
E-	170 <i>c</i>	250 <i>d</i>	420
	370	400	770

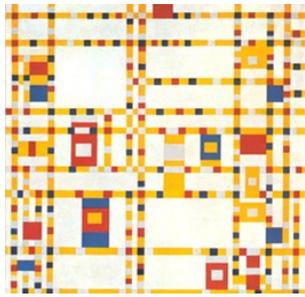
$$RR = 1,41$$

$$IC\ 95\% = (1,22-1,64)$$

	M+	M-	
E+	20 <i>a</i>	15 <i>b</i>	35
E-	17 <i>c</i>	25 <i>d</i>	42
	37	40	77

$$RR = 1,41$$

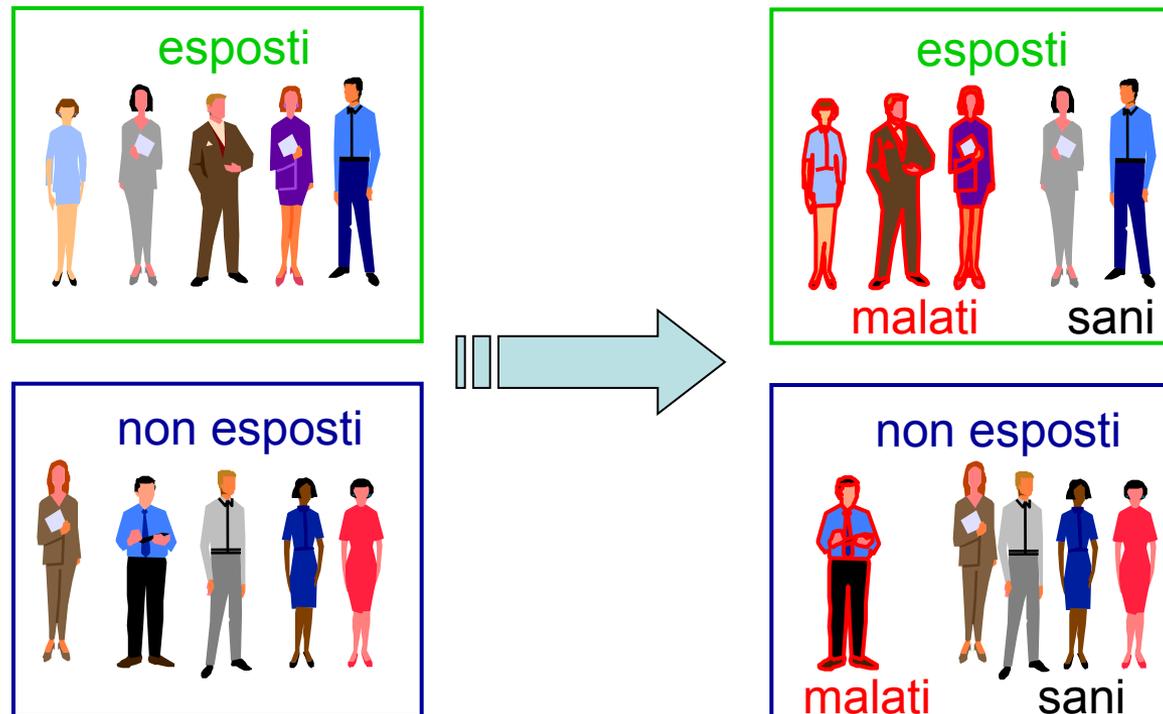
$$IC\ 95\% = (0,89-2,25)$$



## La relazione causa-effetto 2

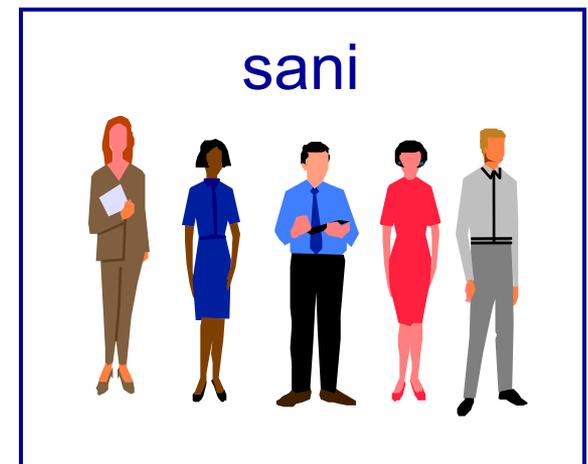
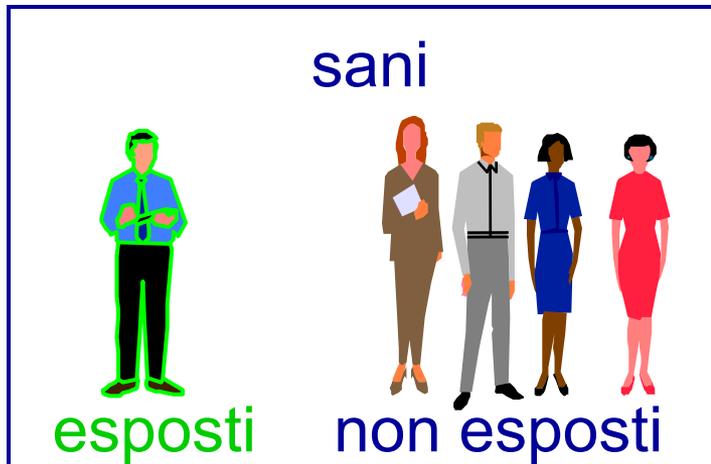
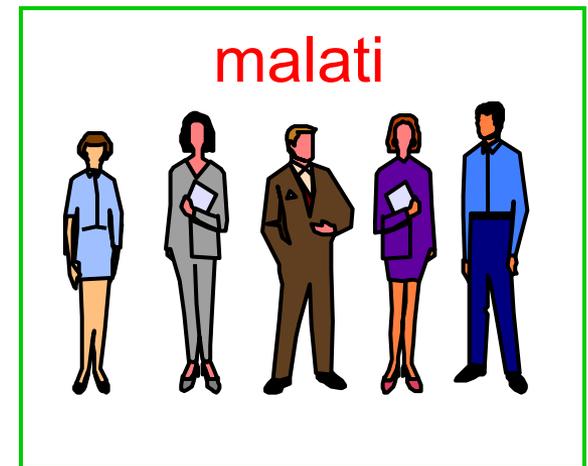
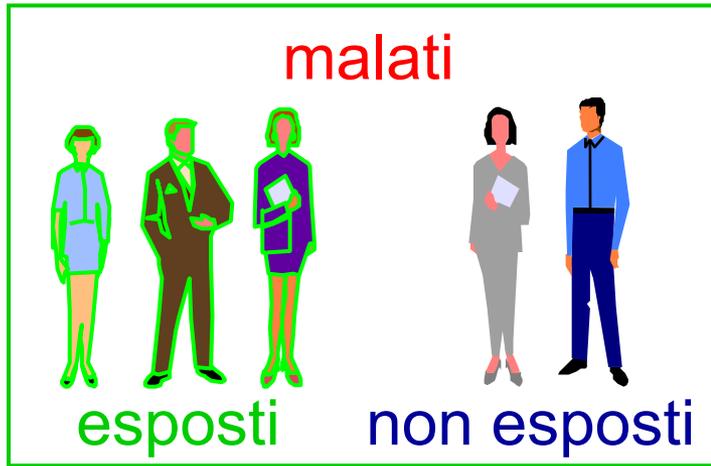
- gli studi *caso-controllo*  
e altri modelli di studio
- vantaggi e limiti degli studi analitici
- misure di impatto sulla popolazione

# Limiti di uno studio di coorte

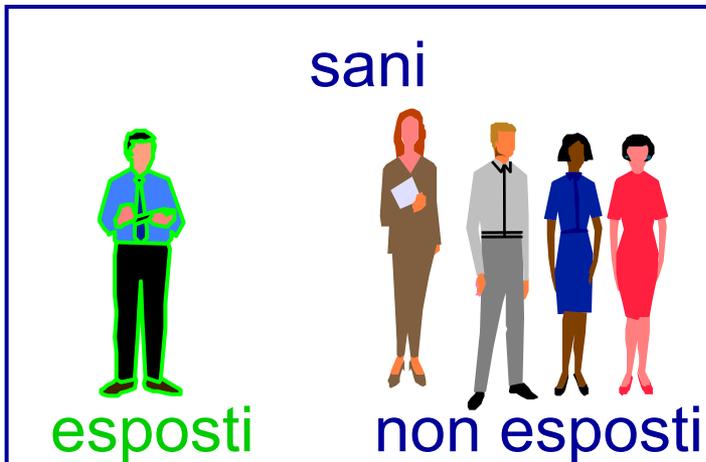
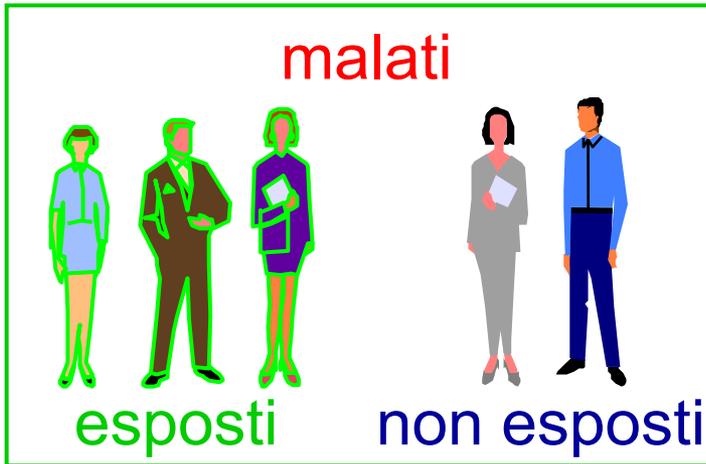


*nello studio di coorte si seleziona una popolazione sana e si valuta l'incidenza di malattia in funzione dell'esposizione a un fattore di rischio*

# Lo studio caso-controllo



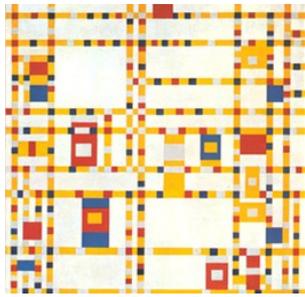
# Lo studio caso-controllo



	M+	M-	
E+	3 <i>a</i>	1 <i>b</i>	4
E-	2 <i>c</i>	4 <i>d</i>	6
	5	5	10

$$\text{Odds Ratio} = \frac{\text{odds}_{M+}}{\text{odds}_{M-}} = \frac{3/2}{1/4} = 6$$

$$\text{Odds Ratio} = \frac{a/c}{b/d} = \frac{a \times d}{b \times c}$$



# Vantaggi e limiti degli studi caso-controllo

costo

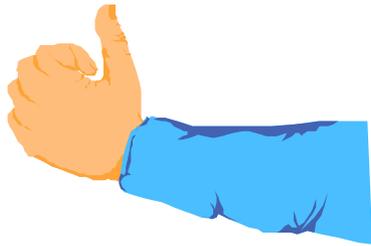
possibilità di valutare fattori di rischio multipli per un'unica patologia

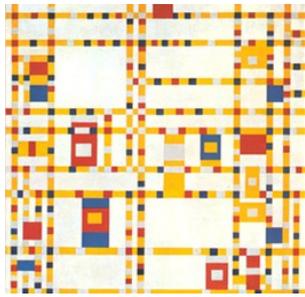
possibilità di studiare patologie rare

maggiore intervento di bias

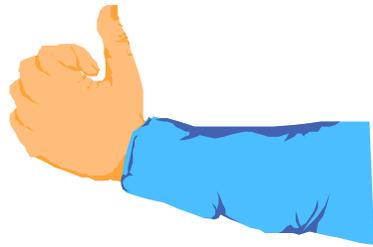
inadeguatezza allo studio di fattori di rischio rari

forniscono solo la stima del rischio nella popolazione





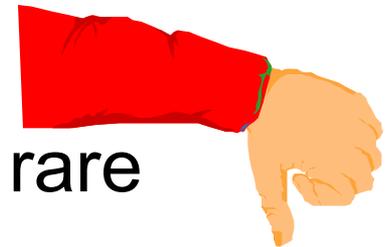
# Vantaggi e limiti degli studi di coorte

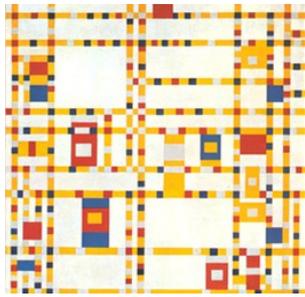


misura diretta dell'incidenza  
possibilità di valutare effetti multipli di un  
unico fattore di rischio  
possibilità di studiare fattori di rischio rari

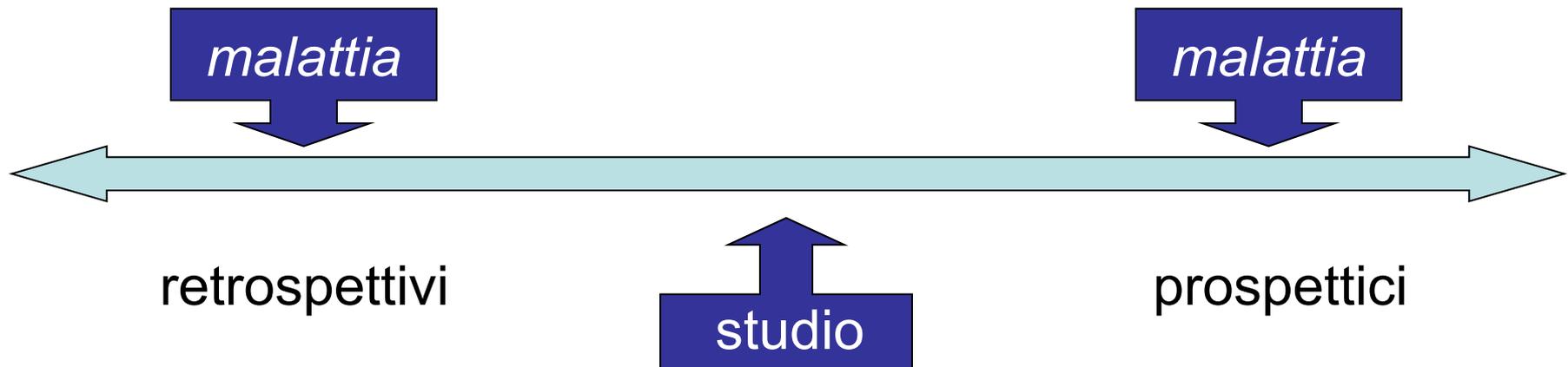
costo

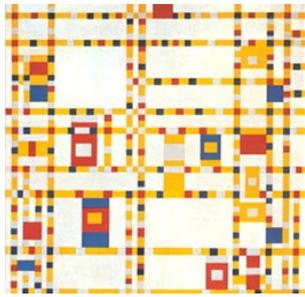
inadeguatezza allo studio di patologie rare





# Una riflessione sugli studi longitudinali





# Altri modelli di studio

- studi trasversali o di prevalenza
  - facili da realizzare
  - poco costosi
  - “l’esposizione precede l’effetto?”
- studi ecologici o di correlazione
  - facili da realizzare
  - l’unità di analisi è la popolazione
  - *fallacia ecologica*

# Indagine di un'epidemia di tossinfezione alimentare

## Approccio 1: studio di coorte

- intervista a tutti i partecipanti
- calcolo dei tassi d'attacco specifici per alimento
- valutazione del rischio relativo



## Approccio 2: studio caso-controllo

- intervista a un campione di partecipanti
- valutazione degli alimenti consumati da casi e controlli
- valutazione dell'odds ratio

# Indagine di un'epidemia di tossinfezione alimentare

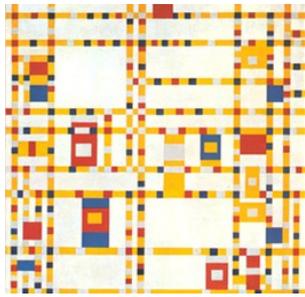
Nome	sessu	età	Data sintomi	Ora sintomi	Antipasto	Primo	Secondo	Contorno	Dolce
Intervistato 1	M	20	12/12/97	12:00	X	X	X	X	X
Intervistato 2	M	24	13/12/97	09:00	X	X			
Intervistato 3	F	35	13/12/97	11:30	X	X	X	X	X
Intervistato 4	M	12	12/12/97	10:30	X		X	X	X
Intervistato 5	F	45	13/12/97	08:15		X	X	X	X
Intervistato 6	F	48	13/12/97	15:00		X			X
Intervistato 7	F	10			X		X	X	X
Intervistato 8	M	25				X			X
Intervistato 9	M	41			X		X	X	
Intervistato 10	F	10			X	X	X	X	X
Intervistato 11	F	5				X			X
Intervistato 12	M	49			X		X	X	X

coorte	$RR_{\text{primo}} = \frac{?/?}{?/?} = ?$	$RR_{\text{secondo}} = \frac{?/?}{?/?} = ?$
caso controllo	$OR_{\text{primo}} = \frac{?/?}{?/?} = ?$	$OR_{\text{secondo}} = \frac{?/?}{?/?} = ?$

# Indagine di un'epidemia di tossinfezione alimentare

Nome	sexso	età	Data sintomi	Ora sintomi	Antipasto	Primo	Secondo	Contorno	Dolce
Intervistato 1	M	20	12/12/97	12:00	X	X	X	X	X
Intervistato 2	M	24	13/12/97	09:00	X	X			
Intervistato 3	F	35	13/12/97	11:30	X	X	X	X	X
Intervistato 4	M	12	12/12/97	10:30	X		X	X	X
Intervistato 5	F	45	13/12/97	08:15		X	X	X	X
Intervistato 6	F	48	13/12/97	15:00		X			X
Intervistato 7	F	10			X		X	X	X
Intervistato 8	M	25				X			X
Intervistato 9	M	41			X		X	X	
Intervistato 10	F	10			X	X	X	X	X
Intervistato 11	F	5				X			X
Intervistato 12	M	49			X		X	X	X

coorte	$RR_{\text{primo}} = \frac{5/8}{1/4} = 2,5$	$RR_{\text{secondo}} = \frac{4/8}{2/4} = 1,0$
caso controllo	$OR_{\text{primo}} = \frac{5/1}{3/3} = 5,0$	$OR_{\text{secondo}} = \frac{4/2}{4/2} = 1,0$

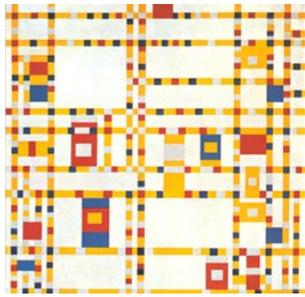


# Rischio attribuibile individuale

*(rischio attribuibile per gli esposti)*

$$RA = I_{Exp+} - I_{Exp-}$$

*quantità di rischio supplementare  
(eccesso di rischio) attribuibile  
al fattore di rischio considerato  
**è una misura di incidenza***

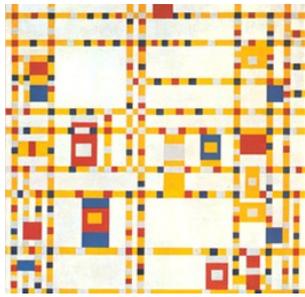


# Frazione attribuibile

*(rischio attribuibile per gli esposti)*

$$FA = \frac{I_{Exp+} - I_{Exp-}}{I_{Exp+}}$$

*la proporzione di malati fra gli esposti  
che eviterebbe la malattia se fosse rimosso  
il fattore di rischio  
**è una proporzione***

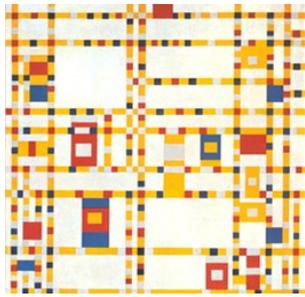


# Rischio attribuibile di popolazione

$$RAP = RA \times P$$

dove P rappresenta la prevalenza del fattore di rischio nella popolazione

*la quantità di casi nell'intera popolazione  
che non ammalerebbero se fosse rimosso  
il fattore di rischio considerato  
**è una misura di incidenza***



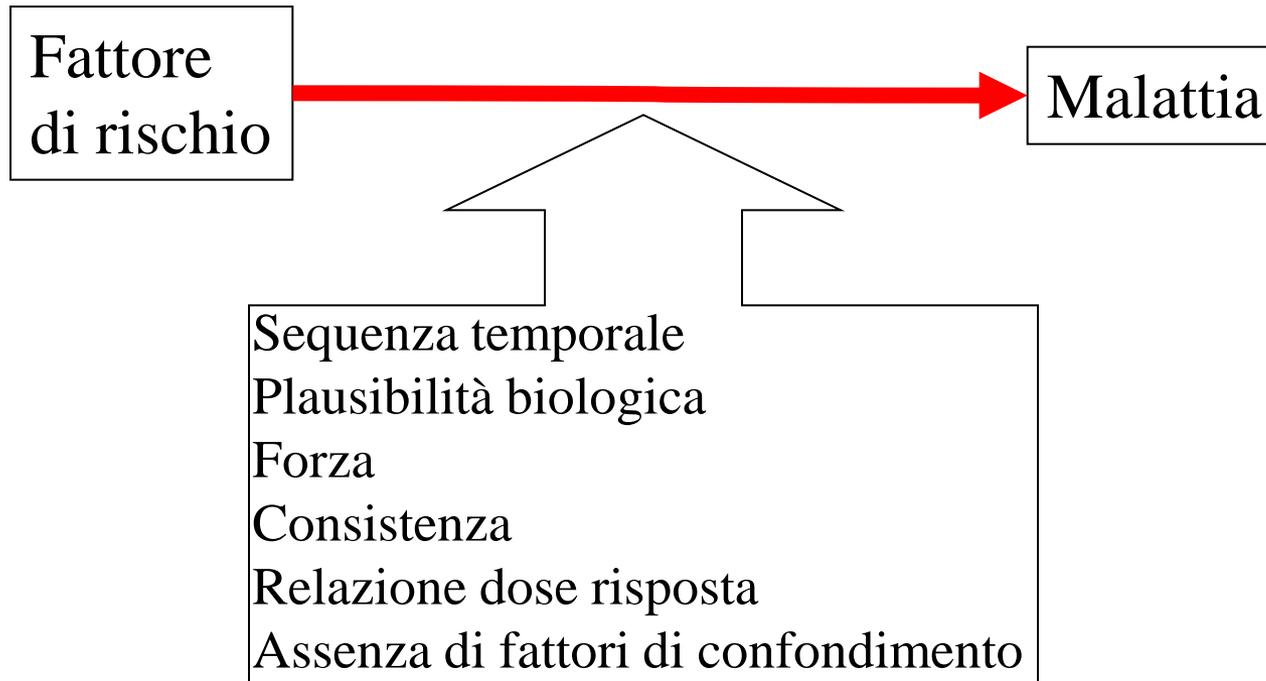
# Frazione eziologica nella popolazione

$$FE = \frac{RAP}{I_{Tot}} = \frac{I_{Tot} - I_{Exp-}}{I_{Tot}}$$

*la proporzione di casi nell'intera popolazione  
dovuti al fattore di rischio  
è una proporzione*

# L'obiettivo

## *Lo studio delle associazioni*



# Il chi-quadro di Mantel-Haenszel

⊕ **Obiettivo: confrontare due proporzioni, studiarne il legame in**  
➤ **presenza di un fattore di stratificazione**

➤ **Assunzioni:**

⊕ **Campioni indipendenti**

⊕ **Distribuzione binomiale**

**In uno studio caso – controllo sono state rilevate informazioni per valutare l’associazione tra l’assunzione di alcool e tumore all’esofago in soggetti fumatori e non fumatori.**

**Come cambia il rischio associato al consumo di alcool considerando che un individuo può essere fumatore?**

**309 casi sono fumatori di questi 265 assumono alcool**

**208 controlli sono fumatori di questi 151 assumono alcool**

**Sono anche bevitori:**

**63 casi non fumatori**

**107 controlli non fumatori**

Fumatori (f)			
	Casi	Controlli	Totale
Alcool si	265	151	416
Alcool no	44	57	101
Totale	309	208	517
Non fumatori (nf)			
Alcool si	63	107	170
Alcool no	63	136	243
Totale	126	243	369

$$\text{OR f} = \text{ad/bc} = \frac{265 \cdot 57}{151 \cdot 44} = 2,27$$

$$\text{OR nf} = \text{ad/bc} = \frac{63 \cdot 136}{107 \cdot 63} = 1,27$$

**E' possibile determinare un unico OR che ci dia informazioni sul rischio di essere bevitori e anche fumatori?**



**Con il metodo di Mantel-Haenszel si “aggiusta” il confronto tra i due gruppi per una terza variabile.**



**L’aggiustamento è ottenuto mediante la determinazione del chi-quadro delle tavole 2x2 di ogni strato, pesato per la dimensione del campione di ciascuno strato.**

**La statistica test valuta contemporaneamente i risultati di tutti gli strati**

$$X^2_{\text{CMH}} = \frac{(|\sum \mathbf{a}_i - \sum \mathbf{E}(\mathbf{a}_i)| - 1/2)^2}{\sum \text{Var}(\mathbf{a}_i)}$$

**La distribuzione della statistica test è  $X^2$  con 1 grado di libertà**

**L'odds ratio di Mantel-Haenszel è dato da**

$$\mathbf{OR}_{MH} = \frac{\Sigma(a_i d_i / N_i)}{\Sigma(b_i c_i / N_i)}$$

**Nel nostro esempio:**

$$\mathbf{OR}_{MH} = \frac{[(265*57)/517] + [(63*136)/369]}{[(44*151)/517] + [(63*107)/369]} = \frac{52,44}{31,12} = \mathbf{1,69}$$

**Il test di Mantel-Haenszel può essere  
utilizzato anche per la verifica dell'ipotesi  
sul odds-ratio**

$$\mathbf{H_0 : OR = 1}$$

$$\mathbf{H_1 : OR \neq 1}$$

**Il test di significatività per verificare l'ipotesi  $H_0$  OR=1  
è sempre un test  $X^2$**

$$X^2 = \frac{[\sum a_i - \sum E(a_i)]^2}{\sum \text{Var}(a_i)}$$

**I valori attesi e la varianza sono:**

$$\begin{aligned} \sum E(a_i) &= E(a_1) + E(a_2) = \\ &= [(a_1 + b_1)(a_1 + c_1) / N_1] + [(a_2 + b_2)(a_2 + c_2) / N_2] = \\ &= 248,63 + 58,05 = 306,68 \end{aligned}$$

$$\begin{aligned} \sum \text{Var}(a_i) &= \text{Var}(a_1) + \text{Var}(a_2) = \\ &= \{[(a_1 + b_1)(c_1 + d_1)(a_1 + c_1)(b_1 + d_1)] / N_1^2 (N_1 - 1)\} + \\ &+ \{[(a_2 + b_2)(c_2 + d_2)(a_2 + c_2)(b_2 + d_2)] / N_2^2 (N_2 - 1)\} = \\ &= 19,58 + 20,67 = 40,25 \end{aligned}$$

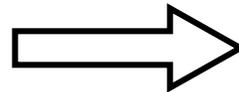
**Applicando ai dati del nostro esempio il metodo di Mantel-Haenszel si ha:**

$$\chi^2 = \frac{[\sum a_i - \sum E(a_i)]^2}{\sum \text{Var}(a_i)} = \frac{(328-306,68)^2}{40,25} = 10,77$$

**p<0,001; l'OR è statisticamente significativo**

**L'intervallo di confidenza può essere calcolato con il metodo di Miettinen:**

$$OR \left( 1 \pm \frac{z_{1-\alpha/2}}{\sqrt{\chi^2}} \right)$$



**Limiti di confidenza al 95%**

$$1,69^{(1 - 1,96 / 3,28)} = 1,23$$

$$1,69^{(1 + 1,96 / 3,28)} = 2,32$$

**La probabilità di tumore all'esofago può anche essere predetta attraverso la regressione logistica**

$$\ln(p \text{ avere tumore} / q \text{ non avere tumore}) = \\ b_0(\text{parametro che indica il rischio di tumore senza altri fattori}) + \\ +b_1(\text{parametro che "pesa" la variabile}) * \text{fumo} + \\ +b_2 * \text{alcohol}$$

**Alcohol e fumo non sono variabili quantitative!**

**Si trasformano in variabili apparentemente quantitative assegnando arbitrariamente *valore 1* alla *presenza* del fattore di rischio e *valore 0* all'*assenza* del fattore di rischio. Queste variabili vengono comunemente chiamate variabili "dummy".**

**Il rischio di tumore in relazione all'abitudine all'alcool e al fumo può essere valutata con la *regressione logistica***

$$y = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_n x_n$$

**Dove**

$$y = \ln\left(\frac{p}{q}\right)$$

**Trasformazione logit**

**Questo rapporto è l'*odd***

**Semplificando con una sola variabile  $y = b_0 + b_1 x$**

$$p = \frac{\exp(b_0 + b_1 x_1)}{1 + \exp(b_0 + b_1 x_1)} \quad \Rightarrow \quad \text{OR} = \exp b_1$$