

# LEZIONI DI EPIDEMIOLOGIA

*Dott. SIMONE ACCORDINI*

## *Lezione n. 2*

- Determinante*
- Misure di associazione*



*Sezione di Epidemiologia & Statistica Medica  
Università degli Studi di Verona*

**Determinante (D):** caratteristica in funzione della quale variano i valori del parametro di occorrenza P

**Esempi:**

- Il fumo è un determinante dell'incidenza del tumore al polmone  
(il rischio per un fumatore è > del rischio per un non fumatore)*
- L'atopia è un determinante dell'incidenza di asma  
(il rischio per un soggetto atopico è > del rischio per un soggetto non atopico)*



Esempi:

OUTCOME	PARAMETRO DI OCCORRENZA
<i>presenza di asma</i>	→ <i>prevalenza di asma</i>
<b>soggetti atopici</b>	<b>vs</b> <b>soggetti non atopici</b>
<i>morte per cancro</i>	→ <i>rischio di morte per cancro</i> <i>tasso di mortalità per cancro</i>
<b>fumatori</b>	<b>vs</b> <b>non fumatori</b>
<i>livello di glicemia</i>	→ <i>valore medio della glicemia</i>
<b>diabetici trattati con un nuovo farmaco ipoglicemizzante</b>	<b>vs</b> <b>diabetici trattati con un farmaco ipoglicemizzante standard</b>
<i>tempo di sopravvivenza</i>	→ <i>mediana del tempo di sopravvivenza</i>
<b>pazienti affetti da tumore al pancreas trattati con EPD</b>	<b>vs</b> <b>pazienti affetti da tumore al pancreas trattati con SPD</b>



Il determinante è una caratteristica misurata (generalmente) a **livello individuale** che ammette **almeno 2 valori**:

$D_0$  = assente (soggetto non esposto)

$D_1$  = presente (soggetto esposto)

Esempio:

*determinante = fumo*

• variabile qualitativa ⇒ 0 = non fumatore  
1 = fumatore

• variabile quantitativa ⇒ 0  
1  
2  
... *pack-years*



**Esercizio:**

*Nell'ECRHS è stata misurata la broncoreattività bronchiale (BHR), la presenza di sibili nel torace e l'abitudine al fumo in un campione casuale di adulti (20-44 anni) in tre città del Nord Italia.*

**Numero totale di soggetti (N), numero (%) di soggetti broncoreattivi e numero (%) di soggetti con sibili in varie categorie di fumatori.**

	N	BHR		SIBILI	
		N	%	N	%
fumatori	422	71	(16.8)	115	(27.3)
ex fumatori	209	38	(18.0)	34	(16.3)
non fumatori	472	83	(17.5)	48	(10.2)

**Il fumo è un determinante della broncoreattività?**

**Il fumo è un determinante dei sibili?**

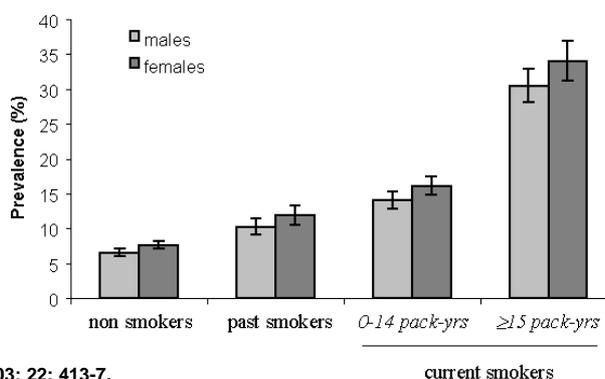


**RELAZIONE DI OCCORRENZA:**

Relazione tra un **parametro di occorrenza (P)** e

un **determinante (D)** o un insieme di determinanti:  **$P = f(D)$**

Esempio: Prevalenza di tosse e catarro cronici in Italia (soggetti di età 20-44 anni).



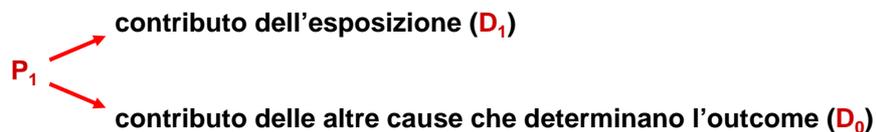
Cerveri I, et al.

Eur Respir J 2003; 22: 413-7.



**Lo scopo della ricerca epidemiologica è in generale lo studio e la misurazione della relazione di occorrenza**

- La relazione d'occorrenza è anche nota come "ASSOCIAZIONE tra il determinante e la malattia"
  - Se il determinante ha solo due livelli:  $D_0$  = esposizione assente  
 $D_1$  = esposizione presente
- ⇒ la relazione di occorrenza si misura confrontando i parametri di occorrenza valutati separatamente nei due gruppi:  $P_0 \leftrightarrow P_1$
- ( $P_0$  = parametro di occorrenza - gruppo non esposto  
 $P_1$  = parametro di occorrenza - gruppo esposto)



In generale:

- $P_1 > P_0$  → associazione positiva (esposizione = **fattore di rischio**)
- $P_1 < P_0$  → associazione negativa (esposizione = **fattore protettivo**)
- $P_1 \sim P_0$  → non esiste associazione (non è un determinante)

Esempio: Prevalenza di rinite allergica in Italia (soggetti di età 20-44 anni).  
de Marco R, et al. *Clin Exp Allergy* 2002;32:1405-12.

area sub-continentale ( $D_0$ ) → 18.0%  
area mediterranea ( $D_1$ ) → 20.2% }  $P_1 > P_0$



... talvolta:

$P_1 < P_0$  → **associazione negativa** (esposizione = **fattore di rischio**)

$P_1 > P_0$  → **associazione positiva** (esposizione = **fattore protettivo**)

$P_1 \sim P_0$  → **non esiste associazione** (**non** è un determinante)

Esempio:

Studio sulla sopravvivenza dopo Pancreaticoduodenectomia Standard (SPD) o Extended (EPD) per adenocarcinoma duttale pancreatico in stadio I-III.

Iacono C, et al. *World J Surg* 2002;26:1309-14.

Probabilità di sopravvivenza dopo 1 anno dal trattamento:

SPD (D <sub>0</sub> )	→	31%	} P <sub>1</sub> > P <sub>0</sub>
EPD (D <sub>1</sub> )	→	76%	



La relazione di occorrenza viene espressa mediante **misure sintetiche di effetto**, che esprimono il grado di relazione tra D e P

### EFFETTO

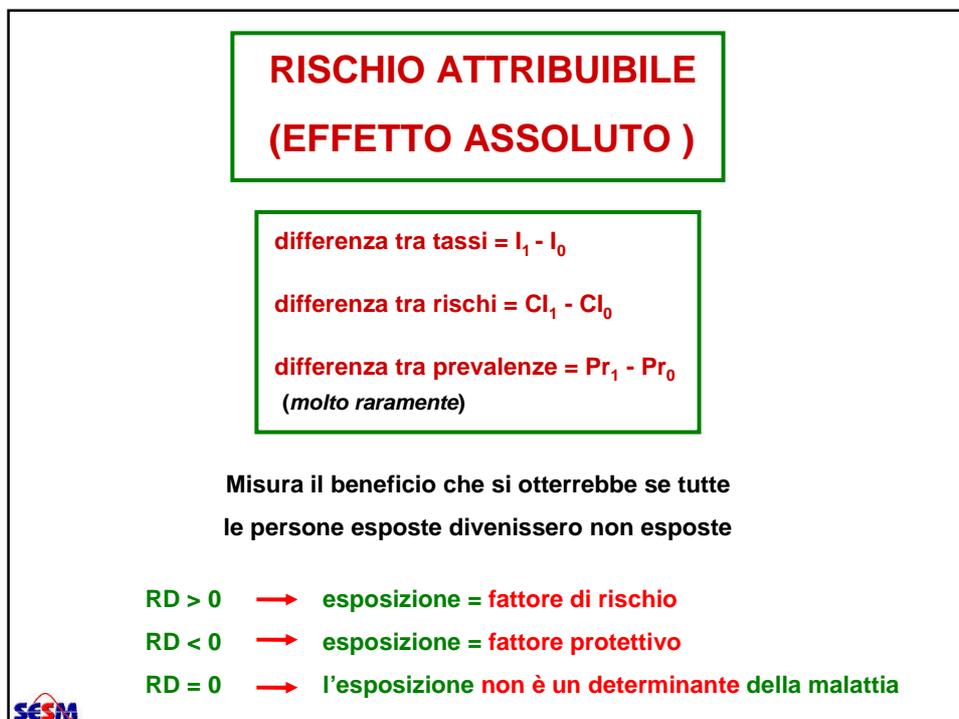
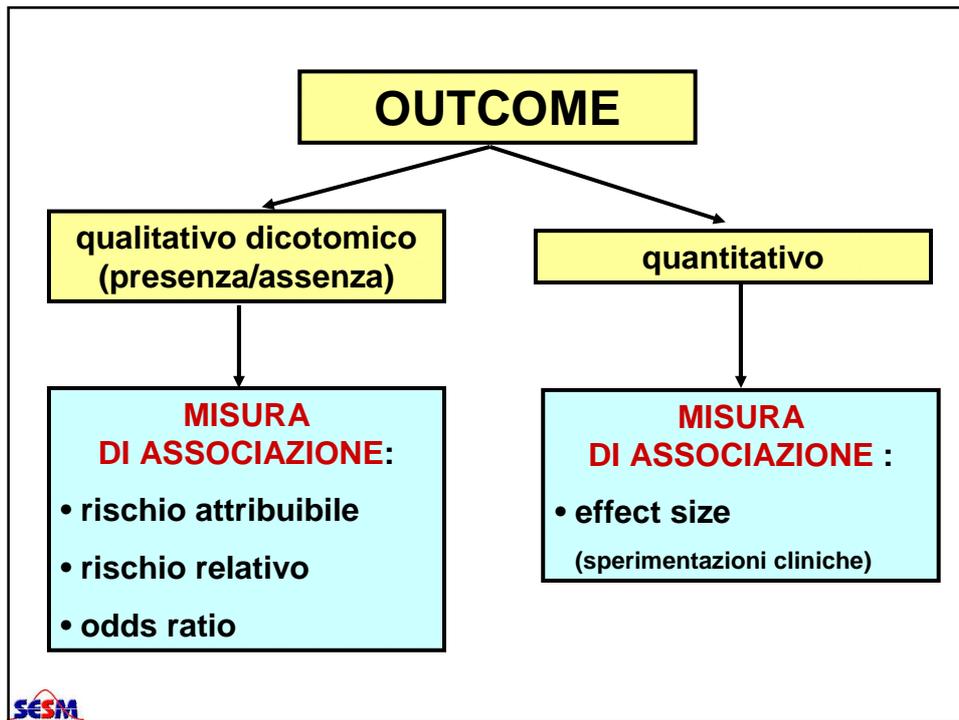
Quantità di cambiamento nella frequenza della malattia causata da uno specifico fattore

Esempi:

\*Ogni anno l'allattamento al seno dei bambini nati da madri infette da HIV è responsabile della trasmissione della malattia a 200 per 1000 bambini-anno nati da tali madri

\*I forti fumatori (>45 pack-yrs) maschi hanno un rischio di avere bronchite cronica 17 volte maggiore rispetto ai non fumatori (Cerveri et al. *ERJ* 2001;18:85-92).





- RD misura l'effetto di un'esposizione sulla popolazione
- RD permette di valutare il **beneficio atteso** da un intervento di prevenzione



utile per stabilire a quale intervento sanitario dare la priorità

**Esempio:**

*L'incidenza di HIV in bimbi nati da madri infette non allattati al seno è pari a 180 casi per 1000 bambini-anno e a 380 casi per 1000 bambini-anno tra quelli allattati:*

$$RD = 380 \times 1000 \text{ anni}^{-1} - 180 \times 1000 \text{ anni}^{-1}$$

$$= 200 \text{ per } 1000 \text{ bambini-anno}$$

l'effetto dell'allattamento al seno da madri affette da HIV è di produrre 200 casi ogni 1000 bambini per anno

RD > 0 ⇒ esposizione = **fattore di rischio**



**RISCHIO RELATIVO  
(EFFETTO RELATIVO)**

rapporto tra tassi =  $I_1 / I_0$

rapporto tra rischi =  $CI_1 / CI_0$

rapporto tra prevalenze =  $Pr_1 / Pr_0$

Misura la forza dell'associazione causale  
tra il determinate e la malattia

RR > 1 → esposizione = **fattore di rischio**

RR < 1 → esposizione = **fattore protettivo**

RR = 1 → l'esposizione **non è un determinante della malattia**



- RR è la più utilizzata misura della relazione di occorrenza
- RR misura la **forza dell'associazione tra D e P**
- RR è la più importante **misura eziologica**

**Esempio:**

*La mortalità per ulcera negli USA nel 1950 era di 93 casi per 1000000 persone-anno, mentre nel 1990 era di 31 casi per 1000000 persone-anno.*

$$RR = 93 \times 1000000 \text{ anni}^{-1} / 31 \times 1000000 \text{ anni}^{-1} = 3$$

Un americano nel 1950 aveva un rischio di morire per ulcera 3 volte superiore rispetto ad un americano nel 1990.

$RR > 1 \Rightarrow$  esposizione = **fattore di rischio**



**Esempio:**

*Numero di cardiopatie ischemiche (CHD) in funzione del tipo di personalità (Western Collaborative Group Study):*

	CHD	NO CHD	
Tipo A	178	1.411	1589
Tipo B	79	1.486	1565
	257	2.897	3154

**tipo A:** competitivo, apprensivo

**tipo B:** rilassato e non competitivo

Coorte di individui tra i 34 e i 59 anni seguiti per un periodo di 8 anni



## DISTRIBUZIONE BIVARIATA (CROSS-TABULATION)

- Permette la **rappresentazione congiunta** della distribuzione di frequenza di due variabili qualitative
- Permette di capire la **relazione** tra due variabili qualitative

*esempio: distribuzione dell'abitudine al fumo e della tosse e catarro cronici (TC) in adulti di età 20-44 anni a Verona (indagine ISAYA).*

Fumo	$n_i$	$p_i(\%)$	TC	$n_i$	$p_i(\%)$
<b>non fumatore</b>	1066	50.1%	<b>assente</b>	1897	89.1%
<b>ex fumatore</b>	379	17.8%	<b>presente</b>	232	10.9%
<b>fumatore</b>	684	32.1%	<b>Totale</b>	2129	100.0%
<b>Totale</b>	2129	100.0%			



### DISTRIBUZIONE CONGIUNTA ASSOLUTA

distribuzione congiunta del fumo e della TC ( $n_{ij}$ )

FUMO	TC		TOTALE
	assente	presente	
<b>non fumatore</b>	997	69	1066
<b>ex fumatore</b>	335	44	379
<b>fumatore</b>	565	119	684
<b>TOTALE</b>	1897	232	2129

distribuzione marginale del fumo ( $n_{i\cdot}$ )

distribuzione marginale della TC ( $n_{\cdot j}$ )

dimensione campionaria ( $n$ )



## DISTRIBUZIONE CONGIUNTA RELATIVA (%)

FUMO	TC		TOTALE
	assente	presente	
non fumatore	997	69	1066
ex fumatore	335	44	379
fumatore	565	119	684
TOTALE	1897	232	2129

non fumatori con TC ( $n_{12}$ ) (arrow to 69)  
dimensione campionaria ( $n$ ) (arrow to 2129)

$(n_{ij} / n) * 100$

FUMO	TC		TOTALE
	assente	presente	
non fumatore	46.8%	3.2%	50.0%
ex fumatore	15.7%	2.1%	17.8%
fumatore	26.5%	5.6%	32.1%
TOTALE	89.1%	10.9%	100.0%

$(69 / 2129) * 100$  (arrow to 3.2%)



## DISTRIBUZIONI CONDIZIONALI

- Rappresentano la distribuzione di una variabile **all'interno delle modalità dell'altra variabile**
- Se le distribuzioni condizionali sono differenti, si può supporre che esista una **relazione tra le due variabili**



**DISTRIBUZIONI CONDIZIONALI AI MARGINALI DI RIGA:  
DISTRIBUZIONE DELLA TC PER LIVELLO DI FUMO**

FUMO	TC		TOTALE
	assente	presente	
non fumatore	997	69	1066
ex fumatore	335	44	379
fumatore	565	119	684
TOTALE	1897	232	2129

marginali di riga ( $n_i$ )

$(n_{ij} / n_i) * 100$

FUMO	TC		TOTALE
	assente	presente	
non fumatore	93.5%	6.5%	100.0%
ex fumatore	88.4%	11.6%	100.0%
fumatore	82.6%	17.4%	100.0%
TOTALE	89.1%	10.9%	100.0%

$(69 / 1066) * 100$



**DISTRIBUZIONI CONDIZIONALI AI MARGINALI DI COLONNA:  
DISTRIBUZIONE DEL FUMO PER LIVELLO DELLA TC**

FUMO	TC		TOTALE
	assente	presente	
non fumatore	997	69	1066
ex fumatore	335	44	379
fumatore	565	119	684
TOTALE	1897	232	2129

marginali di colonna ( $n_j$ )

$(n_{ij} / n_j) * 100$

FUMO	TC		TOTALE
	assente	presente	
non fumatore	52.5%	29.7%	50.1%
ex fumatore	17.7%	19.0%	17.8%
fumatore	29.8%	51.3%	32.1%
TOTALE	100.0%	100.0%	100.0%

$(69 / 232) * 100$



**Esempio:**

Numero di cardiopatie ischemiche (CHD) in funzione del tipo di personalità (Western Collaborative Group Study):

	CHD	NO CHD	
Tipo A	178	1.411	1589
Tipo B	79	1.486	1565
	257	2.897	3154

tipo A: competitivo, apprensivo  
 tipo B: rilassato e non competitivo

Coorte di individui tra i 34 e i 59 anni seguiti per un periodo di 8 anni

$$CI_A = \text{prob}(\text{CHD} | \text{Tipo A}) = 178/1589 = 11.2\%$$

$$CI_B = \text{prob}(\text{CHD} | \text{Tipo B}) = 79/1565 = 5.0\%$$

$$RR = \frac{11.2}{5.0} = 2.24$$

Un soggetto di tipo A ha un rischio di sviluppare una cardiopatia ischemica in un periodo di 8 anni pari a più del doppio del rischio per un soggetto di tipo B.

RR > 1 ⇒ esposizione = **fattore di rischio**



**Esempio:**

Tassi di mortalità (x 100000 anni<sup>-1</sup>) per cancro al polmone e cardiopatie ischemiche (CHD) nei medici maschi britannici (Doll & Peto. BMJ 1976;2:1525-36).

	cancro polmonare	CHD
fumatori	140	669
non fumatori	10	413

RR 14.0 1.6

RD 130 256 x 100000 anni<sup>-1</sup>

Il fumo è un fattore di rischio più 'forte' per il cancro al polmone (outlook eziologico)

Il fumo ha un impatto sulla salute pubblica maggiore per le cardiopatie ischemiche (outlook salute pubblica)



**Esercizio:**

*Incidenza di tumore al polmone in soggetti di età 45-54 anni,  
tra i fumatori e i non fumatori (x 100000 anni<sup>-1</sup>)*

fumatori	non fumatori
67.0	5.8

**RD = ???**

**RR = ???**



**Quali misure di rischio utilizzare per valutare la relazione di occorrenza: *tasso di incidenza (I)*, *incidenza cumulativa (CI)* o *prevalenza (Pr)*?**

- Le misure della relazione tra D e P dipendono dal tipo di misura di rischio utilizzata.
- Se il tempo di osservazione è molto breve / e *CI* danno risultati simili. **Per malattie acute e di breve durata talvolta può essere usata anche *Pr*.**
- Se il tempo di osservazione di una comunità è relativamente lungo / è la migliore stima del rischio da utilizzare.



Un'altra misura di associazione utilizzata spesso è l'**ODDS RATIO**.

Se E è un evento di interesse, definiamo come **ODDS**:

$$\text{Odds}(E) = \frac{\text{Prob}(E)}{1-\text{Prob}(E)}$$

	Outcome (M <sub>1</sub> )	Non outcome (M <sub>0</sub> )	Totale
Esposti (D <sub>1</sub> )	r <sub>1</sub>	n <sub>1</sub> - r <sub>1</sub>	n <sub>1</sub>
Non esposti (D <sub>0</sub> )	r <sub>0</sub>	n <sub>0</sub> - r <sub>0</sub>	n <sub>0</sub>
Totale	r	n - r	n

**Odds di malattia**

$$\text{Odds}(M_1 | D_1) = \frac{\text{prob}(M_1 | D_1)}{\text{prob}(M_0 | D_1)} = \frac{r_1/n_1}{(n_1 - r_1)/n_1} = \frac{r_1}{n_1 - r_1}$$

negli esposti

$$\text{Odds}(M_1 | D_0) = \frac{\text{prob}(M_1 | D_0)}{\text{prob}(M_0 | D_0)} = \frac{r_0/n_0}{(n_0 - r_0)/n_0} = \frac{r_0}{n_0 - r_0}$$

nei non esposti



**ODDS RATIO**

$$OR = \frac{\text{Odds}(M_1 | D_1)}{\text{Odds}(M_1 | D_0)} = \frac{r_1}{n_1 - r_1} \cdot \frac{n_0 - r_0}{r_0}$$

	Outcome (M <sub>1</sub> )	Non outcome (M <sub>0</sub> )	Totale
Esposti (D <sub>1</sub> )	r <sub>1</sub>	n <sub>1</sub> - r <sub>1</sub>	n <sub>1</sub>
Non esposti (D <sub>0</sub> )	r <sub>0</sub>	n <sub>0</sub> - r <sub>0</sub>	n <sub>0</sub>
Totale	r	n - r	n

- è una **stima del RR**
- è l'**unica misura di associazione** calcolabile negli **studi caso-controllo**

OR > 1 → l'esposizione è un **fattore di rischio**

OR < 1 → l'esposizione è un **fattore protettivo**

OR = 1 → l'esposizione **non è un determinante della malattia**



Esempio: Numero di cardiopatie ischemiche (CHD) in funzione del tipo di personalità.

	CHD	NO CHD	
Tipo A	178	1.411	1589
Tipo B	79	1.486	1565
	257	2.897	3154

$$CI_1 = \text{pr}(\text{CHD} | \text{Tipo A}) = 178/1589 = 0.112$$

$$CI_0 = \text{pr}(\text{CHD} | \text{Tipo B}) = 79/1565 = 0.050$$

$$RR = \frac{0.112}{0.050} = 2.24$$

$$OR = \frac{178 \cdot 1486}{1411 \cdot 79} = 2.37$$

Un soggetto di tipo A ha un rischio di sviluppare una cardiopatia ischemica in un periodo di 8 anni pari a più del doppio del rischio per un soggetto di tipo B.

RR (OR) > 1 ⇒ **fattore di rischio**



### Esercizio:

Liddell ed altri (1984) fornirono i risultati di uno studio sull'associazione tra carcinoma bronchiale ed esposizione all'asbesto nelle miniere e negli stabilimenti di crisolito in Canada. Tra i 520 soggetti esposti ad asbesto, 148 presentavano carcinoma ai polmoni; mentre tra i 418 non esposti, 75 presentavano carcinoma ai polmoni.

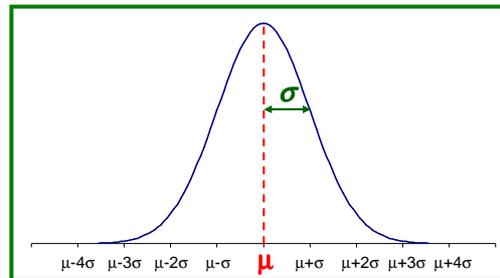
	Malati (M <sub>1</sub> )	Sani (M <sub>0</sub> )	Totale
Esposti (D <sub>1</sub> )			
Non esposti (D <sub>0</sub> )			
Totale			

OR = ???



## OUTCOME QUANTITATIVO (DISTRIBUZIONE SIMMETRICA)

Molte variabili biologiche ( $X$ ) hanno **DISTRIBUZIONE NORMALE** o approssimativamente normale ( $\rightarrow$  **MODELLO TEORICO**)



### STATISTICHE CAMPIONARIE:

$\bar{x}$  = media nel campione  
 $s$  = dev. std. nel campione



### CARATTERISTICHE DELLA POPOLAZIONE:

$\mu$  = media nella popolazione  
 $\sigma$  = dev. std. nella popolazione



## EFFECT SIZE



SPERIMENTAZIONI CLINICHE

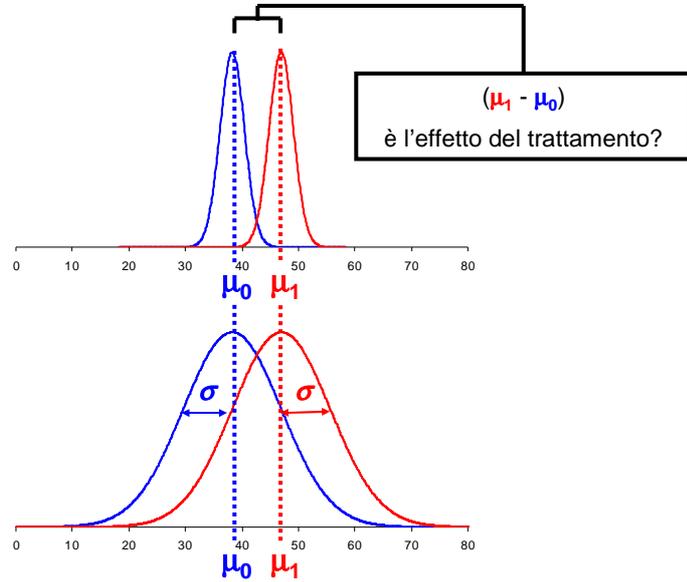


## OUTCOME QUANTITATIVO (distribuito normalmente)

- parametro biologico o patologico (*calcemia, dimensione della massa tumorale, ...*)
- differenza/rapporto tra il valore di un parametro biologico o patologico misurato **prima e dopo** la somministrazione del trattamento (*variazione della calcemia, ...*)



**CONFRONTO DELLA MEDIA  
TRA GRUPPO SPERIMENTALE E GRUPPO DI CONTROLLO**



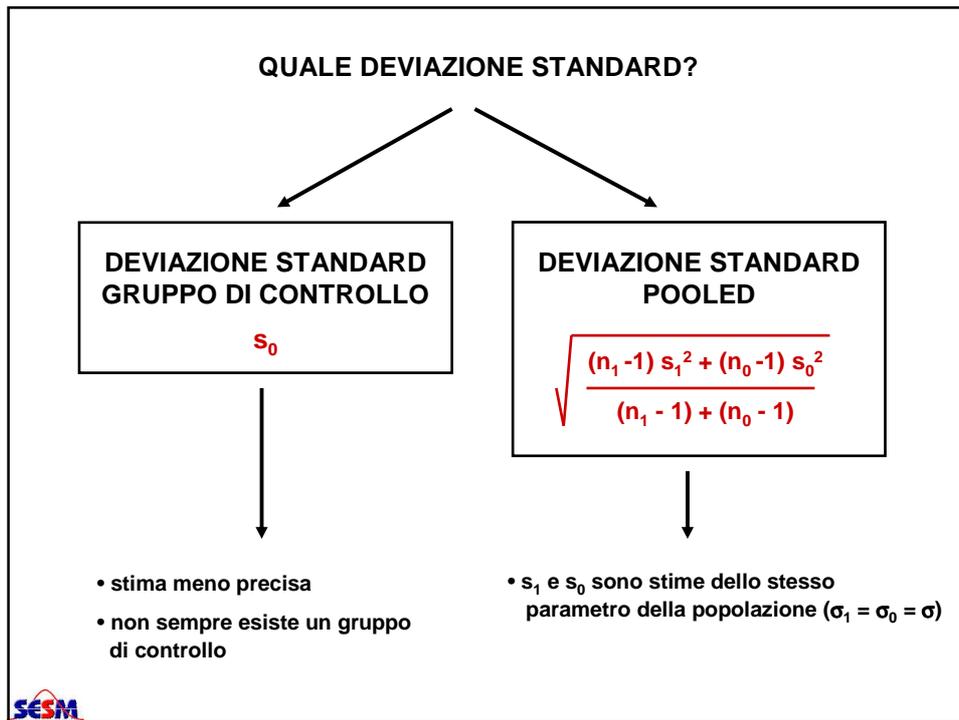
	Numero di pazienti	Valore medio	Dev. standard
Gruppo sperimentale ( $D_1$ )	$n_1$	$\bar{x}_1$	$s_1$
Gruppo di controllo ( $D_0$ )	$n_0$	$\bar{x}_0$	$s_0$

**MISURA DI EFFICACIA DEL TRATT. SPERIMENTALE:**

$$\text{effect size} = \frac{\bar{x}_1 - \bar{x}_0}{\text{dev. std.}}$$

- misura a-dimensionale della distanza tra le due medie
- permette il confronto tra effetti non misurati sulla stessa scala





**COME INTERPRETARE IL VALORE DELL'EFFECT SIZE: LIMITI DI COHEN**

Entità dell'effetto	Effect Size	Percentile
	2,0	97,7
	1,9	97,1
	1,8	96,4
	1,7	95,5
	1,6	94,5
	1,5	93,3
	1,4	91,9
	1,3	90
	1,2	88
	1,1	86
	1,0	84
	0,9	82
<b>GRANDE</b>	<b>0,8</b>	79
	0,7	76
<b>MEDIO</b>	<b>0,5</b>	69
	0,4	66
	0,3	62
<b>PICCOLO</b>	<b>0,2</b>	58
	0,1	54
	0,0	50

media del gruppo sperimentale come percentile della distribuzione del gruppo di controllo

NB: i limiti di Cohen riflettono il valore degli effect size ottenuti nelle scienze comportamentali (→ grandi effetti)



**Esempio:**

*Cockburn et al (1980) riportano una ricerca clinica per la prevenzione dell'ipocalcemia infantile, nella quale donne in gravidanza che ricevevano un supplemento di vitamina D venivano messe a confronto con donne non trattate.*

Calcemia del bambino misurata 6 giorni dopo la nascita

	Numero di pazienti	Media (mg/100 ml)	DS (mg/100 ml)
Vitamina D (D <sub>1</sub> )	233	9.36	1.15
Controllo (D <sub>0</sub> )	394	9.01	1.33

La vitamina D determina un aumento rilevante della calcemia nei neonati?



	Numero di pazienti	Media (mg/100 ml)	DS (mg/100 ml)
Vitamina D (D <sub>1</sub> )	233	9.36	1.15
Controllo (D <sub>0</sub> )	394	9.01	1.33

$$\text{dev. std. pooled} = \sqrt{\frac{(233 - 1) 1.15^2 + (394 - 1) 1.33^2}{(233 - 1) + (394 - 1)}} = 1.27$$

$$\text{effect size} = \frac{\bar{x}_1 - \bar{x}_0}{\text{dev. std.}} = \frac{9.36 - 9.01}{\text{dev. std.}} = \begin{cases} \frac{0.35}{1.33} = 0.26 \\ \frac{0.35}{1.27} = 0.28 \end{cases}$$

La vitamina D determina un aumento limitato della calcemia nei neonati.

