

ISTITUTO SUPERIORE DI SANITÀ

Schemi di Epidemiologia di base

A cura di
Francesco Rosmini, Silvia Andreozzi e Luigina Ferrigno

Centro Nazionale di Epidemiologia, Sorveglianza e Promozione della Salute

Strumenti di riferimento
06/S2

Istituto Superiore di Sanità

Schemi di Epidemiologia di base.

A cura di Francesco Rosmini, Silvia Andreozzi e Luigina Ferrigno

2006, 100 p. (Strumenti di riferimento 06/S2)

Il presente materiale didattico è il risultato di una lunga esperienza d'insegnamento della Epidemiologia di base. Si compone di schemi ordinati secondo un programma basato sugli ottenimenti teorici della seconda metà del '900. Sono presi in rassegna i seguenti argomenti: misure di frequenza e follow up, misure di effetto, confronto tra associazione e causalità, sperimentazione clinica e studi osservazionali, cause di artefatti, fenomeni di confondimento, ed effetti di più fattori di rischio. Nel programma è anche presente la tradizione empirica maggiormente collegata, rispetto a quella teorica, ad attività di sanità pubblica, in particolare sono considerati i criteri per indagini di epidemie e per attività di screening. Il formato a schema può avere una rilevanza didattica nel caso di attività di formazione dove si richieda ai discenti un limitato impegno di memorizzazione.

Parole chiave: Epidemiologia di base, Materiale didattico

Istituto Superiore di Sanità

Outlines of Basic Epidemiology.

Edited by Francesco Rosmini, Silvia Andreozzi and Luigina Ferrigno

2006, 100 p. (Strumenti di riferimento 06/S2) (in Italian)

The report presents the educational material resulted from a long teaching experience on basic Epidemiology. It is organized into outlines which are ordered according to the theoretic achievements established in the second half of '900. The topics considered are: measures of frequency and follow up, measures of effect, comparison of association and causality, clinical experiments and observational studies, causes of bias, confounding, and effects of more risk factors. The empirical tradition linked to activities of public health is presented with elements for the study of epidemics and screening activities. The schematic format can have an educational relevance for teaching activities where students are requested of a moderated work of memorization.

Key words: Basic epidemiology, Teaching material

Per informazioni su questo documento scrivere a: francesco.rosmini@iss.it.

Il rapporto è disponibile online dal sito di questo Istituto: www.iss.it.

Presidente dell'Istituto Superiore di Sanità e Direttore responsabile: *Enrico Garaci*
Registro della Stampa - Tribunale di Roma n. 475/88 del 16 settembre 1988

Redazione: *Paola De Castro, Sara Modigliani e Sandra Salinetti*

La responsabilità dei dati scientifici e tecnici è dei singoli autori o curatori.

© 2006 Istituto Superiore di Sanità (Viale Regina Elena, 299 - 00161 Roma)

INDICE

Presentazione	1
1. Misure di frequenza e follow-up	3
2. Misure di effetto e causalità	29
3. Sperimentazioni e studi osservazionali	45
4. Definizioni di bias, effetto di più fattori di rischio	65
5. Studio di un'epidemia	79
6. Screening e validità di una misura	91
Bibliografia	100

Presentazione

Quando nel 1993 il professor Panà direttore della Scuola di Specializzazione di Igiene e Medicina Preventiva dell'Università di Tor Vergata mi propose di svolgere un corso di fondamenti di epidemiologia mi occupavo prevalentemente di ricerca, ma accettai con entusiasmo perché insegnare, oltre alla soddisfazione intrinseca che offre, è anche una sfida culturalmente proficua. La conduzione del corso avrebbe comportato la necessità di documentarsi, l'impegno di memoria e il ricorso a un eloquio privo di sottointesi, azioni che avrebbero fornito occasioni di aggiornamento e ripensamento delle mie conoscenze. Inoltre limiti di tempo e obiettivi didattici avrebbero richiesto l'esercizio in qualche misura creativo di scegliere i contenuti giusti, concatenarli con associazioni logiche, renderli comprensibili attraverso esempi illuminanti, individuare quesiti allo scopo di negare l'apparente banalità delle definizioni.

Il corso sarebbe consistito in lezioni ed esercitazioni a cadenza settimanale di una o due ore per un totale di dodici ore e avrebbe riguardato la definizione e l'uso di elementi di epidemiologia nel ragionamento quantitativo, mentre a un corso gemello sarebbero toccate le strategie e i metodi epidemiologici.

Sulla scelta dei contenuti del corso erano possibili varie soluzioni. Sostanzialmente erano praticabili due strade: presentare i concetti epidemiologici, così come erano applicati ai vari problemi sanitari secondo l'approccio empirico sviluppato nell'ottocento, o descrivere la disciplina delineata nella seconda metà del '900 quando uno sforzo di sintesi aveva portato a "conciliare idee in conflitto e a rendere unitaria la fondazione concettuale, omettendo particolari ininfluenti e a conciliare i principi epidemiologici con gli scopi e metodi più generali della ricerca scientifica" (Kenneth J.Rothman).

A prescindere dal piacere personale di esporre argomenti epidemiologici logicamente connessi, la seconda soluzione aveva il vantaggio di fornire allo studio una struttura schematizzabile nel tempo a disposizione e di essere più facilmente ricordata. Comunque nel programma preparato rimane traccia della tradizione empirica in due lezioni riguardanti lo studio delle epidemie, e i criteri sottesi alle attività di screening poiché sembrava opportuno completare il corso con esempi su specifiche attività di sanità pubblica.

Anche per la conduzione del corso non ero in una landa didattica. Mi ispirai allo stile americano perché offriva i maggiori vantaggi. Quelli americani, per mia esperienza sono corsi brevi che comportano per il discente un impegno di memorizzazione leggero, i contenuti sono fissati schematicamente in *hand-out*, ossia dispense distribuite prima delle lezioni e la verifica della comprensione dei concetti è fatta proponendo dei quiz. Libri sono consigliati per eventuali approfondimenti, ma non sono richiesti per familiarizzarsi con il linguaggio tecnico. Per parlare degli argomenti in studio ci sono i *lab* dove piccoli gruppi di studenti imparano dialogando, stimolati da un esercitatore. Questo ultimo aspetto non avrebbe potuto essere propriamente realizzato, comunque essendo prevista una media di quindici partecipanti, durante la lezione avrei provveduto a stimolare l'attenzione nonché il ricordo di argomenti connessi alla lezione trattati in precedenza con continue domande.

Le dispense preparate per il corso hanno un'importanza decisiva perché sono alla base della lezione, debbono essere adeguatamente articolate per dare risalto alle peculiarità degli argomenti trattati e allo stesso tempo schematiche per facilitare il ricordo dell'insieme. Per molti anni le dispense sono state fornite con una disattenzione un po' snobistica riguardo agli aspetti grafici (corsivo amanuense) e al supporto cartaceo (fogli più o meno sparsi), solo negli ultimi tempi ho chiesto a Silvia Andreozzi e a Luigina Ferrigno di rivedere le dispense e dare loro una efficace impostazione grafica. Dispense come queste non possono essere redatte in una sequenza di righe ad interlinea 2, le dimensioni ed il tipo dei caratteri tipografici, le sottolineature, le cornici, gli spazi bianchi a risalto della formula fondamentale sono espedienti grafici che hanno una rilevante importanza didattica.

Come accennato il corso è nato dodici anni fa, se solo ora ho deciso di dare una forma librerica a queste dispense è perché ne sono discretamente soddisfatto. Il merito oltre che al sottoscritto, a Silvia e Luigina va anche a tutti gli specializzandi che hanno segnalato errori, incongruenze, e hanno direttamente o indirettamente manifestato la loro opinione su come i vari argomenti erano trattati. Tutto ciò ha portato ad una certa personalizzazione del corso e alcuni argomenti sono proposti in maniera abbastanza originale. Alcune assunzioni o semplificazioni potrebbero apparire discutibili, ci tenevo a rendere evidenti aspetti che facilitassero l'intuizione degli argomenti trattati. Il programma proposto risponde agli obiettivi di un'introduzione all'Epidemiologia, la mia speranza è che questo approccio possa incoraggiare tutti coloro potenzialmente interessati a continuare con entusiasmo lo studio dell'Epidemiologia.

1. Misure di frequenza e follow-up

Premessa

Aprè il capitolo una lapidaria definizione dell'epidemiologia: "lo studio della malattia (e della salute) in popolazioni umane". Di seguito sono presentate due definizioni: tradizionali le prime, piú intelligenti le seconde, di incidenza e prevalenza, misurazioni base dell'epidemiologia (pag. 5).

L'esempio successivo chiarisce perché è necessario stabilire l'entità della popolazione ed il tempo di osservazione, elementi della definizione dei tassi (pag. 6).

Dopo un cenno agli aspetti matematici ed una rassegna dei principali tassi applicati alla misura di morbosità (o morbilità) e mortalità della popolazione si passa all'esposizione di alcune applicazioni (pag. 7-8).

I dati di notifica nazionali sono utilizzati per costruire gli andamenti dei tassi di incidenza, mortalità e letalità di due malattie infettive per individuare aspetti legati alla prevenzione e alla terapia (pag. 9-10).

Le impressionanti discrepanze nei risultati di uno studio di prevalenza e di incidenza sulla stessa popolazione sono utilizzate per concludere che il tasso di prevalenza dipende dal tasso d'incidenza e dalla durata della malattia (pag. 11).

I tassi possono essere calcolati su specifici sottogruppi, ad esempio se sono considerati solo i maschi i tassi calcolabili sono specifici per sesso (pag. 12).

I tassi specifici a confronto forniscono piú informazione del corrispondente tasso grezzo (Totale) (pag. 13).

Dopo l'esempio di calcolo di un tasso di mortalità età-specifico, vengono presentati i tassi di mortalità grezzi ed età-specifici di Messico e Svezia per stabilire dove è piú conveniente avviare un'attività di pompe funebri (pag. 14-15).

La struttura per età della popolazione messicana e svedese spiega il tipo di mortalità nei due Paesi (pag. 15).

Distribuzioni di frequenza diverse di una caratteristica (es. età) potrebbero non consentire il confronto dei tassi grezzi di due o piú popolazioni, ma possono essere corrette mediante la standardizzazione (pag. 16).

Dalle popolazioni naturali si passa a popolazioni costituite da soggetti reclutati dall'epidemiologo (pag. 17).

Il Tasso d'incidenza può essere stimato mediante il Rischio ed il Tasso medio d'incidenza.

Il calcolo delle due stime d'incidenza per il confronto di due popolazioni si basa su un esempio estremizzato per rendere evidente la differenza (pag. 18-19).

Il tempo di osservazione è uno degli elementi per il calcolo del tasso d'incidenza, risulta perciò critica la scelta della scala temporale (pag. 20).

"Alla fine saremo tutti morti" è la tetra osservazione di Keynes che stabilisce il limite temporale di un tasso di mortalità. L'esempio demografico in tabella introduce poi al concetto di suddivisione del tempo di osservazione e misurazione della sopravvivenza condizionata al tempo (pag. 21).

L'analisi di sopravvivenza riguarda le stesse popolazioni utilizzate per il calcolo del rischio e del tasso d'incidenza medio. Queste due stime d'incidenza sono meno informative delle curve ottenute con l'analisi di sopravvivenza (pag. 22-23).

La suddivisione per intervalli del tempo di osservazione è un metodo utile per:

- studiare gli andamenti temporali di un tasso (pag. 24);
- eseguire l'analisi di sopravvivenza quando una parte dei soggetti non sono stati seguiti per tutto il tempo di osservazione previsto dallo studio (pag. 25).

Un esempio mostra il razionale per l'uso della probabilità condizionata per il calcolo del tasso di sopravvivenza quando ci sono soggetti con tempi di osservazione incompleti (pag. 26-27).

L'andamento a scaletta è tipico delle curve di sopravvivenza calcolate secondo il metodo di Kaplan-Meier. Le curve di sopravvivenza in tabella mettono in evidenza il benefico effetto della supplementazione della dieta con niacina (pag. 28).

Epidemiologia

Studio della malattia (e della salute) in popolazioni umane (Kleinbaum, Kupper, Morgestein; 1982)

Prevalenza

Frequenza di una malattia ad un certo punto del tempo (MacMahon, Pugh; 1970)

Incidenza

Sviluppo della malattia durante un dato periodo (MacMahon, Pugh; 1970)

Altre definizioni:

Prevalenza

Misurazione dicotomica di uno stato di malattia (Kleinbaum, Kupper, Morgestein; 1982)

Incidenza

Misurazione dicotomica di un cambio dello stato di salute (Kleinbaum, Kupper, Morgestein; 1982)

Fallacia nel ragionamento numerico

Esempio:

Un rapporto sulla leucemia acuta riguardava i pazienti al di sopra dei 13 anni di età, ricoverati presso un ospedale universitario, lungo un periodo di 15 anni. La seguente tabella mostra la distribuzione di frequenza dei pazienti per quinquennio suddivisi in due classi di età.

Età (anni)	Quinquennio		
	Primo	Secondo	Terzo
>13-29	41	50	51
30 e oltre	49	76	109
Totale	90	126	160

(Derivato da: Colton; 1979)

Con questi dati gli autori conclusero che l'incidenza della leucemia acuta era aumentata, in particolare al di sopra dei 30 anni di età.

Conclusioni dubbie: casi prevalenti o incidenti da soli non sono sufficientemente informativi.

SOLUZIONE:

 **Insieme ai casi osservati bisogna definire la relativa popolazione e quanto tempo è stata osservata.**

Misura di frequenza

Premessa: quantità epidemiologiche

1) Rapporto (Ratio)

Il rapporto (r) è la misura della grandezza relativa di due numeri

$$r = a:b = c/d \quad \text{Esempio: } 1:7 = 10/70$$

Altra definizione: frazione in cui il numeratore non è incluso nel denominatore

Esempio:

numero di fumatori/numero di non fumatori

2) Proporzione

Una proporzione (pr) è una frazione in cui il valore del numeratore è una parte del valore del denominatore

Esempio:

La proporzione di bimbi con morbillo in una popolazione scolastica è:

$$pr = 25/200 = 0.125 \text{ ovvero } 12.5\% \quad \begin{array}{l} \text{Bimbi con morbillo} = 25 \\ \text{Bimbi nella scuola} = 200 \end{array}$$

3) Tasso (Rate)

Può essere sinonimo di rapporto o proporzione. Volendo essere più precisi, il *tasso* si riferisce a rapporti o proporzioni che rappresentano cambiamenti nel corso del tempo (Last, 1988)

Un tasso è costituito da:

- a) numero di eventi
- b) popolazione
- c) tempo
- d) un fattore, potenza di 10

Esempio:

$$\text{Tasso annuale} = \frac{10 \text{ eventi in un anno}}{\text{Pop. di 400 soggetti}} \times 1000 = 25 \text{ per 1000}$$

Tassi principali

Morbosità o Morbilità

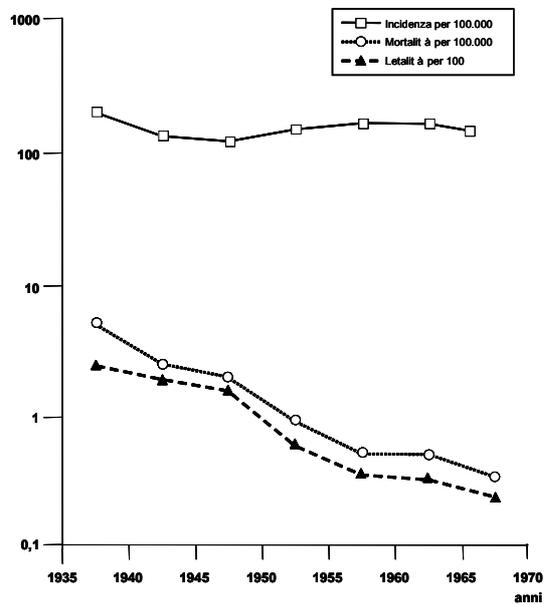
Tasso di incidenza =	N. di nuovi casi di una malattia nella popolazione durante un dato periodo di tempo
	N. di persone a rischio di sviluppare la malattia nella stessa popolazione (popolazione a rischio) durante quel periodo di tempo
Tasso d'attacco =	N. di casi epidemici nella popolazione durante il periodo epidemico
	N. di persone a rischio di sviluppare la malattia all'inizio dell'epidemia
Tasso di prevalenza =	N. di casi di malattia presenti nella popolazione ad un certo momento
	N. di persone nella stessa popolazione nello stesso momento

Mortalità

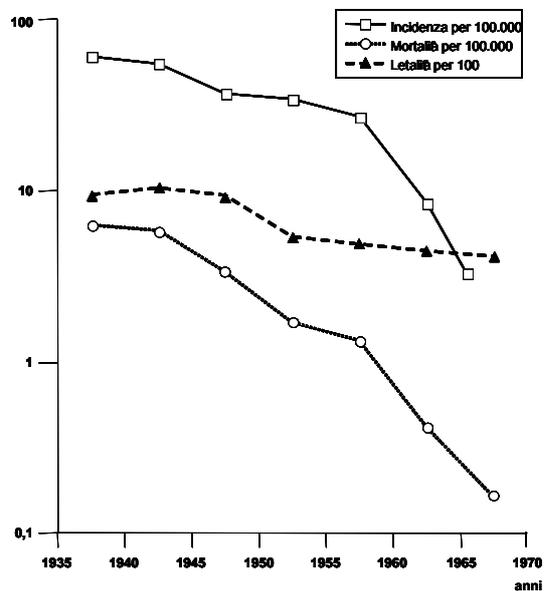
Tasso di letalità = (case fatality rate)	N. di morti per una malattia in una popolazione in un certo periodo
	N. di nuovi casi della malattia nella stessa popolazione nello stesso periodo
Tasso di mortalità =	N. di morti (in totale o per causa) nella popolazione durante un certo periodo
	N. di soggetti della stessa popolazione durante lo stesso periodo di tempo

Casi e morti dovuti a due malattie trasmissibili. Dati notificati in Italia dal 1935-1970, aggregati per quinquennio

Anno a metà quin.	Popolazione a metà quin. a	Casi		Morti		Tassi annuali per 100.000 ab.		Rapporti morti/casi per 100
		<i>totale registrato</i> b	<i>media annuale</i> c=b/5	<i>totale registrato</i> d	<i>media annuale</i> e=d/5	<i>casi</i> c/a x 100000	<i>morti</i> e/a x 100000	<i>tasso di letalità</i> e/c x 100
MALATTIA A								
1937	43.477.000	413.955	82.791	10.677	2.135	190,4	4,91	2,56
1942	45.156.000	279.476	55.895	5.504	1.101	123,8	2,44	1,97
1947	46.054.000	258.845	51.769	4.203	841	112,4	1,82	1,62
1952	47.399.000	334.476	66.911	2.184	437	141,2	0,92	0,65
1957	49.054.000	368.724	73.745	1.402	280	150,3	0,57	0,38
1962	50.693.000	382.737	76.597	1.391	270	151,0	0,55	0,36
1967	52.491.000	355.012	71.002	912	182	135,27	0,33	0,26
MALATTIA B								
1937	43.477.000	136.289	27.258	13.392	2.678	62,7	6,16	9,83
1942	45.156.000	131.834	26.367	13.472	2.729	58,4	6,04	10,35
1947	46.054.000	88.234	17.647	8.472	1.694	32,3	3,68	9,6
1952	47.399.000	70.647	14.130	3.833	766	29,8	1,62	5,43
1957	49.054.000	52.999	10.599	2.782	556	21,6	1,13	5,25
1962	50.693.000	19.30	3.926	957	191	7,74	0,38	4,88
1967	52.491.000	8.417	1.683	383	77	3,21	0,15	4,55



Incidenza, Mortalità e Letalità della Malattia A dal 1935 al 1970, nella popolazione italiana



Incidenza, Mortalità e Letalità della Malattia B dal 1935 al 1970, nella popolazione italiana

Tabella 1. Prevalenza di malattie cardiocircolatorie (MCC) ad un primo esame su 4469 persone tra i 30 e i 62 anni. Studio di Framingham

Età (anni)	Maschi			Femmine			Rapporto dei tassi di prevalenza
	Numero esaminati	Casi MCC	Tasso per 1000	Numero esaminate	Casi MCC	Tasso per 1000	
30-44	1083	5	5	1317	7	5	1,0
45-62	941	43	46	1128	21	19	2,4
Total	2024	48		2445	28		

Tabella 2. Incidenza di malattie cardiocircolatorie (MCC) in un periodo di otto anni su 4995 persone tra i 30 e i 59 anni che non avevano la malattia al primo esame. Studio di Framingham

Età (anni)	Maschi			Femmine			Rapporto dei tassi di incidenza
	Numero esaminati	Casi MCC	Tasso per 1000	Numero esaminate	Casi MCC	Tasso per 1000	
30-39	825	20	24,2	1036	1	1,0	24,2
40-49	770	51	66,3	955	19	19,9	3,3
50-59	617	81	131,3	792	53	66,9	2,0
Total	2212	152		2783	73		

(Derivato da: Mausner and Bahn; 1985)

Relazione tra tasso di prevalenza e tasso di incidenza

La prevalenza di una malattia in una popolazione è una funzione dell'incidenza e della durata media della malattia.

$$P \cong I \cdot D$$

P = Tasso di prevalenza

I = Tasso di incidenza annuale

D = Tempo medio della durata della malattia in anni

Questo è vero se:

- l'incidenza e la durata della malattia sono stabili nel periodo osservato
- non ci sono cambiamenti nella popolazione riguardo a individui suscettibili, malati e immuni

Oltre alla numerosità, può essere utile considerare altre caratteristiche della popolazione come:

SESSO
ETÀ
RAZZA
ECC.

TASSO SPECIFICO: la popolazione è selezionata per una certa caratteristica

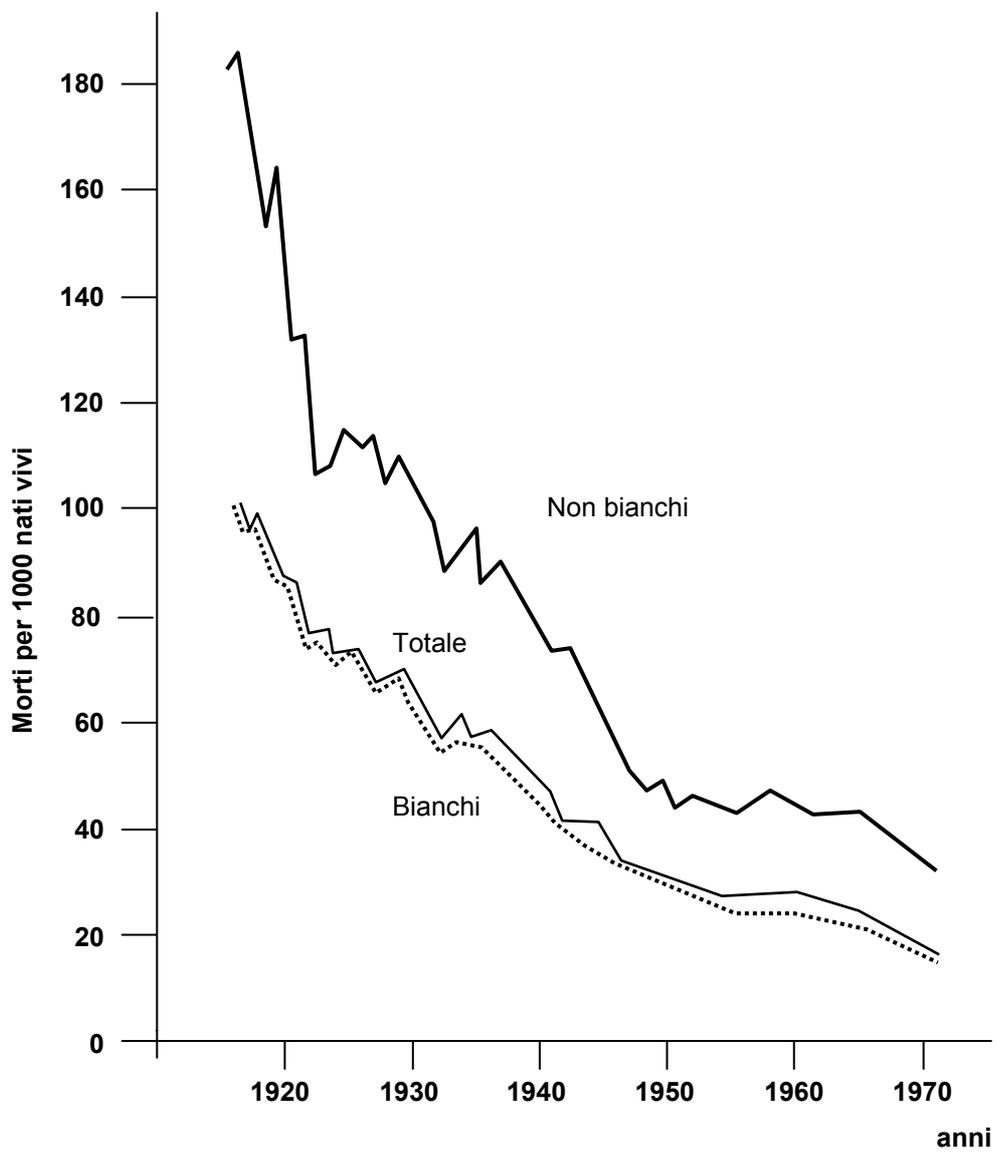
TASSO GREZZO: popolazione generale

Esempio:

nella scuola “G. Verdi”, 133 dei 271 studenti erano maschi. Dei 71 casi di morbillo, avvenuti in ottobre e novembre, 36 riguardavano i maschi.

Calcolare i tassi d’attacco, sesso-specifici per morbillo durante l’epidemia dei mesi di ottobre e novembre.

	Maschi	Femmine
Tasso d’attacco: specifico per sesso	$\frac{36}{133} \times 100 = 27\%$	$\frac{35}{138} \times 100 = 25\%$



Tassi di mortalità infantile per razza, Stati Uniti 1915-1971

Tasso di mortalità età specifico (TMES)

È il numero di morti (D) che avvengono in una fascia di età (X) di una popolazione (P) in un periodo di tempo (t).

Esempio:

X = 5-9 anni

t = 1 anno

D_x = 5451

P_x = 811363 (Popolazione tra 5-9 anni stimata a metà anno)

K = 1000 (Fattore di comodo)

$$\text{TMES}_{(5-9)} = \frac{5451}{811363} \times 1000 = 6.7 \text{ per } 1000$$

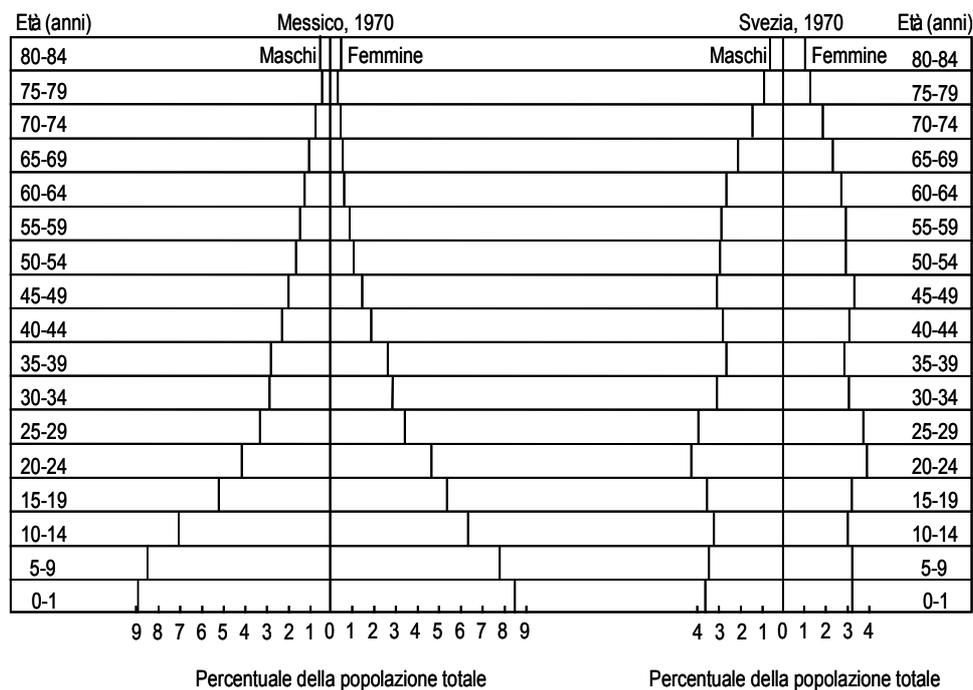
Nota:

La stima a metà anno (1 luglio) è usata perché la popolazione potrebbe cambiare durante l'anno a causa di sbilanciamento tra nascite/morti o immigrazioni/emigrazioni.

Tassi di Mortalità età-specifici (per 1000) in Messico e Svezia (1969)		
Età (anni)	Messico	Svezia
0-1	68,4	11,6
1-4	9,0	0,6
5-14	106	0,3
15-24	2,1	0,7
25-34	3,5	1,0
35-44	5,9	1,9
45-54	8,7	4,4
55-65	18,1	11,3
65-74	39,7	31,1
≥75	119,2	105,6
Tassi di mortalità grezzi	9,4	10,5

(Derivato da: WHO Statistical Annual 1969, Vol. 1, Geneva; 1972)

Distribuzione per età e sesso



(Derivato da: Freedman and Berelson; 1974)

Confrontare tassi grezzi è spesso fuorviante a causa della diversa struttura per età delle popolazioni.

SOLUZIONI:

- 1) **Confrontare i tassi specifici per età**
- 2) **Calcolare i tassi “standardizzati” ovvero tassi grezzi resi confrontabili perché sono state appianate le differenze per età**

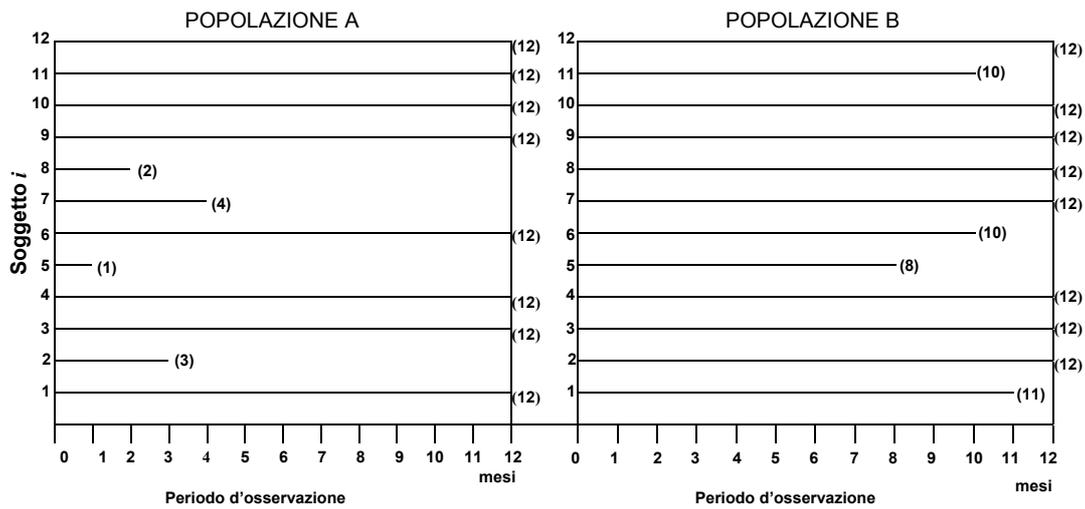
Nota:

La “standardizzazione” è una tecnica che consente di rendere confrontabili più tassi non solo per l’età, ma se necessario, per altre caratteristiche della popolazione.

Popolazione, morti, tassi di mortalità grezzi e standardizzati* in due contee della Florida, 1960				
Contea	Popolazione nel 1960	Morti nel 1960	Tasso di mortalità grezzo (per 1000)	Tasso di mortalità standardizzato per età (per 1000)
Pinellas	347.665	5.726	15,3	10,4
Dade	93.5047	8.322	8,9	10,9
* standardizzazione diretta, effettuata usando come popolazione standard l'insieme delle due popolazioni.				
Caratteristiche della distribuzione per età della popolazione nelle contee di Pinellas e Dade, 1960				
Popolazione	Pinellas		Dade	
< 5 anni	7,4%		10,0%	
≥ 65	24,9%		10,1%	
(Derivato da: Vital Statistics of the US; 1960)				

Esempio:

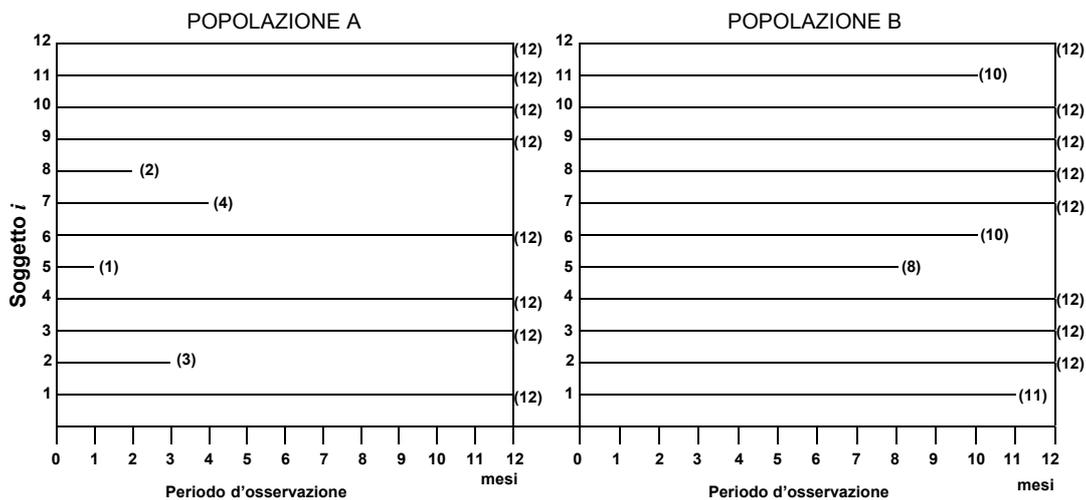
Malattia X in due popolazioni, ciascuna costituita da 12 soggetti seguiti per un anno.



Nota:

- Le popolazioni A e B si dicono “fisse” perché non sono ammesse nuove entrate di soggetti dopo l’inizio dell’osservazione.
- I soggetti colpiti dalla malattia X escono dalla popolazione osservata perché non sono più a rischio.
- In questa situazione sono possibili due diverse misure d’incidenza: il “Rischio” e il “Tasso medio d’incidenza”.

Stima del rischio di malattia



- 1) n. di casi: 4
- 2) popolazione a rischio all'inizio: 12
- 3) tempo d'osservazione: 1 anno

- 1) n. di casi: 4
- 2) popolazione a rischio all'inizio: 12
- 3) tempo d'osservazione: 1 anno

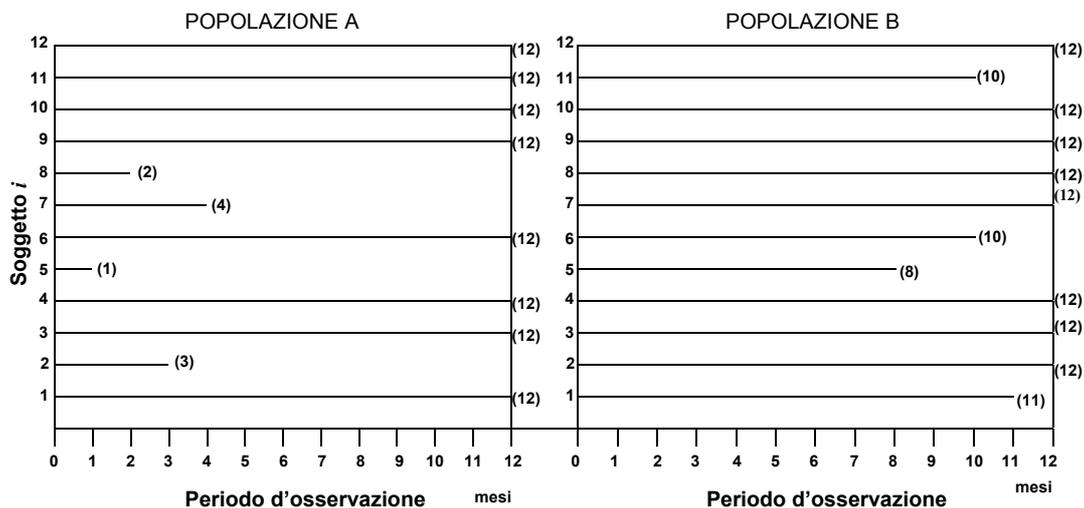
$$\text{Rischio annuale} = \frac{4}{12} \times 100 = 33 \text{ per } 100$$

$$\text{Rischio annuale} = \frac{4}{12} \times 100 = 33 \text{ per } 100$$

Nota:

- Il rischio è una proporzione: valori possibili da 0 a 1 senza dimensione (numeri puri).
- Il tempo non è nella formula, ma deve essere specificato. Si interpreta come una probabilità condizionale: probabilità di malattia nell'individuo in un dato periodo, ammesso che non muoia prima.
- Nelle popolazioni A e B il rischio è lo stesso, ma i casi in A sono avvenuti prima dei casi in B per tener conto di ciò si può stimare il "Tasso medio d'incidenza".

Stima del tasso d'incidenza medio



n. casi: 4
 mesi-persona: 106
 anni-persona: $106/12 = 8.83$

n. casi: 4
 mesi-persona: 135
 anni-persona: 11.25

Tasso d'incidenza medio	$= \frac{4}{8.83} \times 100 = 45.3$ per 100 anni ⁻¹
-------------------------	---

Tasso d'incidenza medio	$= 35.6$ per 100 anni ⁻¹
-------------------------	-------------------------------------

Nota:

- Valori possibili da 0 a ∞ .
- Unità di misura: tempo⁻¹ (tempo: anni, mesi, ecc.), la potenza è negativa perché il tempo è al denominatore.
- Nella formula è considerata sia la popolazione che il tempo d'osservazione.
- Se la malattia è rara (frequenza < 10%) e se l'unità di tempo è la stessa, il valore del Tasso d'incidenza medio è simile a quello del Rischio.

Dimensione temporale

Dipende dal periodo di latenza della malattia in studio

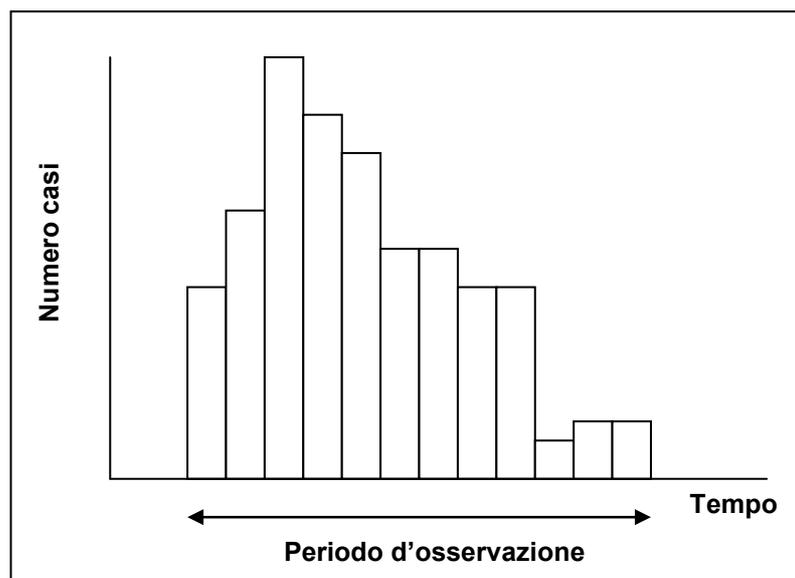
Esempio:

Periodi di osservazione degli effetti della bomba atomica

1° giorno	effetto diretto dell'esplosione
2° giorno – 3° mese	effetto della distruzione
15° anno – 70° anno	effetto delle radiazioni
?	effetti teratogeni

Il Tasso d'attacco è riferito al periodo d'osservazione dell'epidemia

CURVA EPIDEMICA



Mortalità e periodo di osservazione

“Alla fine saremo tutti morti” John M. Keynes citato da Harold A. Kahn

Esempio:

Ricostruzione della mortalità di una popolazione maschile, India, 1931

Età (x)	l_x	d_x	q_x
0	100.000	24.874	0,2487
1	75.126	6.896	0,0918
2	68.230	3.850	0,0564
3	64.380	2.524	0,0392
4	61.856	1.695	0,0274
5	60.161	1.159	0,0193
6	59.002	853	0,0145
7	58.149	669	0,0115
8	57.480	539	0,0094
9	56.941	474	0,0083
10	56.467	447	0,0079
11	56.020	452	0,0081
12	55.568	465	0,0084
13	55.103	484	0,0088
14	54.619	507	0,0093
15	54.112	532	0,0098
16	53.580	558	0,0104
17	53.022	583	0,0110
18	52.439	607	0,0116
19	51.832	629	0,0121
20	51.203	649	0,0127

88	69	33	0,4783
89	36	19	0,5278
90	17	10	0,5882
91	7	4	0,5714
92	3	2	0,6667
93	1	1	1,0000

Età(x) = anni compiuti

l_x = n. di soggetti viventi all'età X

d_x = n. di morti all'età X

q_x = probabilità condizionata di morire all'età X

Confronto di dati di sopravvivenza

Tempo inizio intervallo i (mesi)	POPOLAZIONE A				POPOLAZIONE B			
	N_i	E_i	S_i	$S(t)$	N_i	E_i	S_i	$S(t)$
0 -	12	1	0,92	1	12	0	1	1
2 -	11	2	0,82	0,92	12	0	1	1
4 -	9	1	0,89	$0,92 \cdot 0,82 = 0,75$	12	0	1	1
6 -	8	0	1	$0,75 \cdot 0,89 = 0,67$	12	0	1	1
8 -	8	0	1	0,67	12	1	0,92	1
10 -	8	0	1	0,67	11	3	0,73	0,92
12 -	8	---	---	0,67	8	---	---	0,67

N_i = n. persone vive all'inizio dell'intervallo i

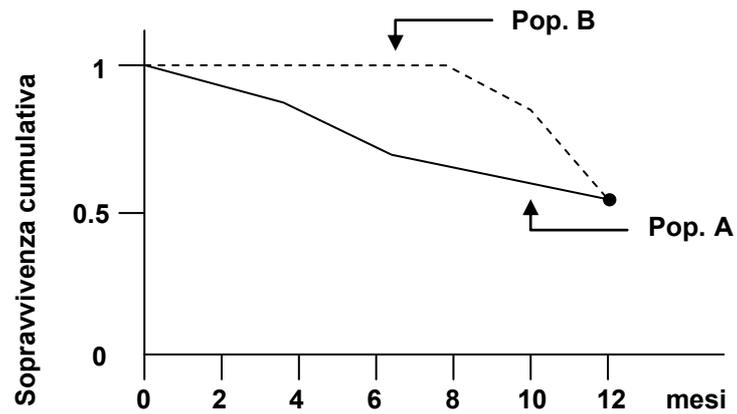
E_i = n. casi durante l'intervallo i

$S_i = (N_i - E_i)/N_i$

$S(t)$ = Tasso di sopravvivenza cumulativa all'inizio t dell'intervallo

Nella pagina successiva i Tassi di sopravvivenza $S(t)$ sono utilizzati per disegnare le curve di sopravvivenza.

Confronto tra curve di sopravvivenza



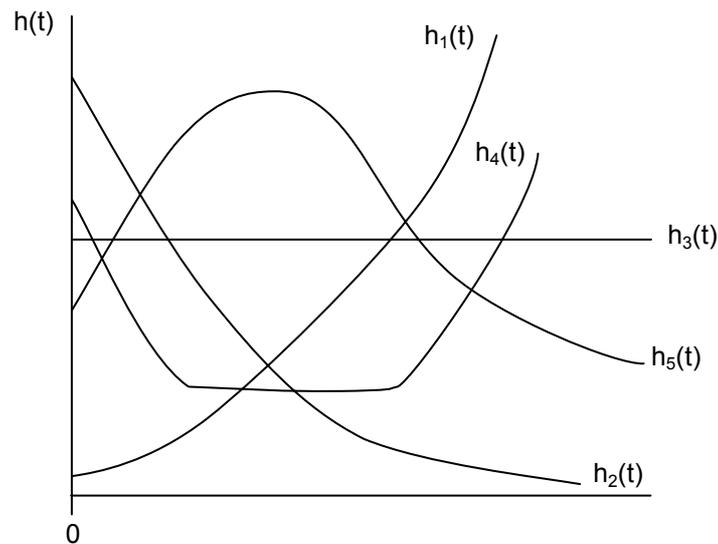
Stessa sopravvivenza dopo 12 mesi, ma prima le curve non coincidono

Nota:

- Invece della mortalità si stima la sopravvivenza per ragioni “psicologiche” e perché il calcolo della stima è più diretto.
- Sopravvivenza non necessariamente da morte. L’evento evitato può essere una malattia X.

A volte conviene suddividere il tempo di osservazione in intervalli più piccoli:

1. *quando il tasso d'incidenza calcolato su periodi lunghi potrebbe appiattire andamenti interessanti*



andamento di:

- $h_1(t)$: leucemia acuta
- $h_2(t)$: operazione dopo incidente
- $h_3(t)$: individui sani tra 18 e 40 anni
- $h_4(t)$: vita umana
- $h_5(t)$: tubercolosi

Esempi di tassi di mortalità condizionata

$$\hat{h}(t) = \frac{\text{n. morti nell'intervallo che comincia al tempo } t}{\text{n. di sopravvissuti al tempo } t}$$

A volte conviene suddividere il tempo di osservazione in intervalli più piccoli:

2. *quando ci sono soggetti osservati per periodi parziali*

Ragioni

Persi all'osservazione

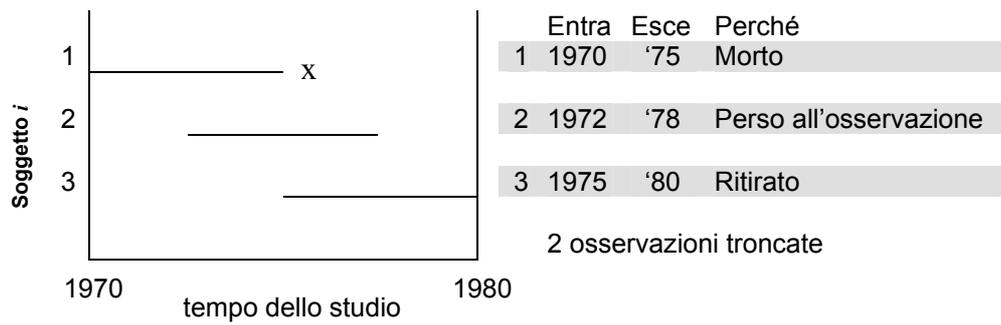
- Trasferito
- Tossicità del trattamento
- Problemi organizzativi

Ritirati

- Fine dello studio

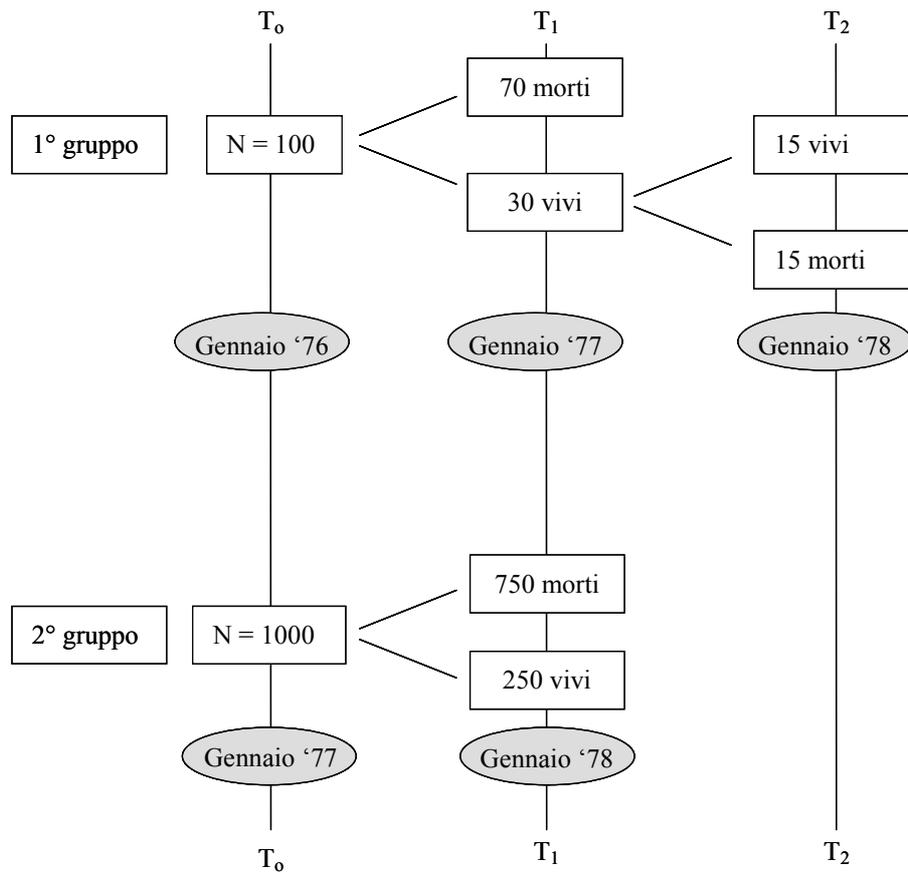
Esempio:

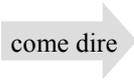
Clinical Trial (*evento – morte*), 1970-80



Esempio:

Due gruppi di un Clinical Trial trattati in tempi diversi



Abbiamo riportato i due gruppi allo stesso punto d'inizio T_0  come dire Assumiamo stessa sopravvivenza tra i 2 gruppi

Come stimare P (sopravvivenza > 2 anni)?

1) Considero solo quelli seguiti per 2 anni

$$\hat{S}_{(2)} = \frac{15}{100} = 0,15 \quad \Rightarrow \text{va bene, ma butto via dei dati!}$$

2) Approccio sbagliato

assumo che i 250 ritirati sono sopravvissuti

$$\hat{S}_{(2)} = \frac{15+250}{1100} = 0,24 \quad \Rightarrow \text{è una stima in eccesso}$$

assumo che tutti i 250 ritirati sono morti

$$\hat{S}_{(2)} = \frac{15}{1100} = 0,014 \quad \Rightarrow \text{è una stima per difetto}$$

ignoro i 250

$$\hat{S}_{(2)} = \frac{15}{1100-250} = 0,018 \quad \Rightarrow \text{correzione insufficiente}$$

3) Approccio corretto

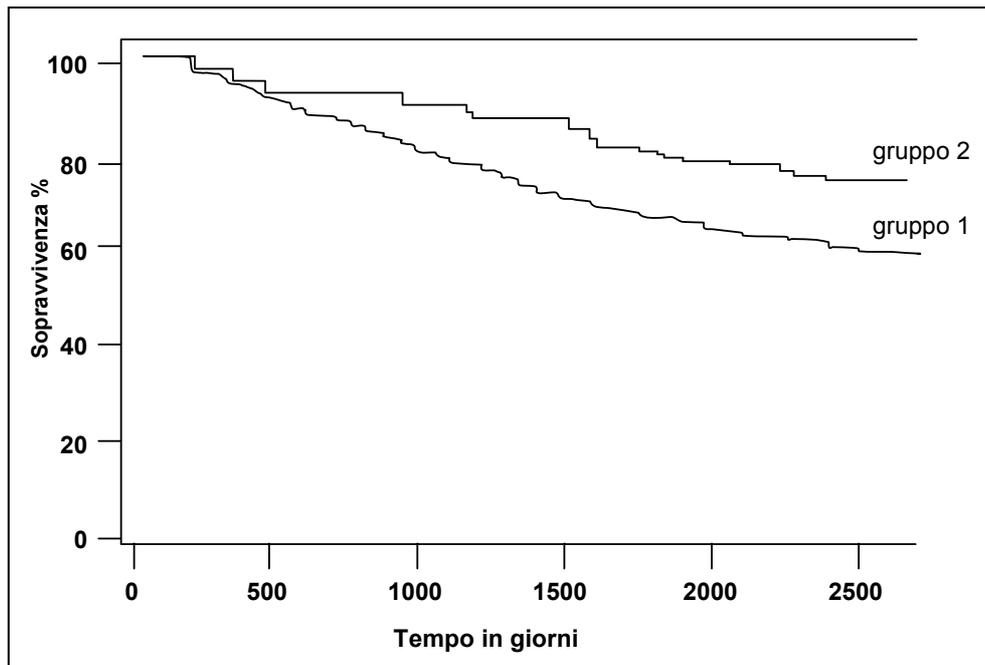
scompongo in probabilità condizionate

$$\begin{aligned} \hat{S}_{(2)} &= \hat{S}_{(1)} \times \hat{S}_{(2|1)} \\ &= \frac{30+250}{1100} \cdot \frac{15}{30} \\ &= \frac{280}{1100} \cdot \frac{1}{2} = \frac{140}{1100} = 0,13 \end{aligned}$$

Nota:

- Il periodo di osservazione è stato diviso in due intervalli di 1 anno.
- Non è stata considerata la differenza di calendario tra i due gruppi ovvero entrambi partono da un artificiale tempo T_0 di inizio osservazione.

Esempio dell'uso delle curve di sopravvivenza secondo Kaplan-Meier*



Proporzione di soggetti sieropositivi senza insorgenza di AIDS trattati con differenti livelli di Niacina nella dieta: gruppo 1 ≤ 61.0 mg/g; gruppo 2 > 61.0 mg/g.

(Derivato da: Tang *et al.*;1993)

* Metodo del prodotto limite di Kaplan-Meier: calcola la sopravvivenza cumulativa, dividendo il tempo d'osservazione in intervalli in modo che non ci siano individui con tempi di osservazione diversi nello stesso intervallo

☐ Notare l'andamento scalare del grafico, i casi incidenti causano i dislivelli

2. Misure di effetto e causalità

Premessa

La misura dell'effetto di una esposizione si ottiene mediante il confronto di due o più tassi (pag. 30).

Il metodo più semplice per il calcolo di una misura d'effetto utilizza le tabelle di contingenza (2x2, 2xn) (pag. 30-32).

L'effetto stimato come rischio relativo si ottiene dal rapporto dei tassi (pag. 33).

L'effetto stimato come rischio attribuibile si ottiene dalla differenza dei tassi (pag. 33).

Il rischio attribuibile può essere calcolato rispetto alla popolazione generale o al gruppo dei soli esposti, il primo fornisce informazioni utili alle strategie di prevenzione, il secondo per stabilire il grado di causalità (pag. 34-35).

I rischi non si possono addizionare o sottrarre direttamente se non riguardano lo stesso gruppo di popolazione (pag. 36).

L'Odds ratio è un'altra misura di effetto. È simile al rischio relativo se la malattia non è troppo frequente (tasso < 10%) (pag. 37).

Per il calcolo di un gradiente d'effetto la tabella di contingenza 2xn deve essere scomposta in tabelle 2x2 con lo stesso gruppo di riferimento (es. gruppo non esposto) (pag. 37).

Le stesse stime riguardano le misure di effetto in Epidemiologia e misure di associazione in Statistica. La Statistica è interessata a stabilire quanto coincida l'occorrenza di esposizione (televisione) con l'occorrenza di malattia (omicidio). L'Epidemiologia se la coincidenza è dovuta a un nesso causale (pag. 38).

La causalità epidemiologica è complessa e i criteri per riconoscerla non sono conclusivi. Non riusciamo a individuare le cause di una malattia, ma solo i suoi fattori di rischio (pag. 39).

Kenneth J. Rothman spiega la causalità debole epidemiologica assumendo che l'insorgenza di una malattia risulti dall'azione di più cause componenti (pag. 40).

L'Epidemiologia valuta il livello di causalità di un'associazione tra fattore di rischio e malattia mediante criteri logici, epidemiologici, biologici (pag. 41).

La verifica di un gradiente di effetto (es. dose-risposta) è uno dei criteri epidemiologici di causalità più persuasivi (pag. 41).

Se l'esposizione non riguarda direttamente il fattore di rischio si parla di esposizione a un determinante o a un proxy (pag. 42).

Avere contatti con gatti rispetto all'infezione da *Leptospira* è più probabile che sia un fattore di rischio, un determinante o un proxy? (pag. 43).

Confronto tra due tassi o rischi

Scopo: verificare se l'incidenza di una malattia tra due gruppi di persone che differiscono riguardo a una caratteristica (esposizione) è differente.

Prima cosa da fare: si costruisce una tabellina 2 x 2.

In una tabellina 2 x 2 sono ordinate le frequenze assolute (a, b, c, d) di quattro sottogruppi di una popolazione classificati secondo due caratteristiche che in epidemiologia sono la malattia e l'esposizione.

		Caratteristica I			
		+	-		
Caratteristica II	+	a	b	a+b	} (tot parziali)
	-	c	d	c+d	
		a+c	b+d	a+b+c+d	(totale)
		} (tot parziali)			

Importanza delle tabelline 2 x 2

“Okay the next principle, and this is one I really want you to keep in mind, is that every study ultimately can be reduced to a 2 x 2 table. If you can't reduce it to a 2 x 2 table that means you don't know your exposure well enough, or you don't know your case definition well enough, or you didn't ask your question clearly enough. If you've asked your question clearly and defined your cases and controls and your exposure well, then you should be able to reduce the study to a 2 x 2 table”

Howard Ory

In sostanza se riusciamo a ordinare i nostri dati in una tabellina 2 x 2 vuol dire che abbiamo chiara l'ipotesi di studio.

Nota:

- Una tabellina 2 x 2 va bene quando malattia ed esposizione sono definite come “presenza/assenza”.
- il numero dei positivi per esposizione e malattia si deve scrivere nella cella in alto a sinistra.
- tabelline 2 x n, dove n=3, 4, ..., quando l'esposizione è definita da più di due categorie.

Esempio di tabelline 2 x 3:

		esposizione		
		++	+	-
malattia	+			
	-			

Esempio:

1000 individui a cui è stata misurata la pressione sistolica seguiti per un anno per la malattia X.

		Pressione sistolica (mmHg)		
		> 170	≤ 170	
Malattia X	+	45	65	110
	-	200	690	890
		245	755	1000

Ipotesi: se la misura del rischio di malattia X è diversa tra i due gruppi allora la pressione sistolica è associata alla malattia X.

$$\text{Rischio}_{>170} = \frac{45}{245} \times 100 = 18\%$$

$$\text{Rischio}_{\leq 170} = \frac{65}{755} \times 100 = 9\%$$

Un rischio è il doppio dell'altro, tale risultato può essere espresso come "effetto", ossia la malattia X è un effetto della pressione sistolica elevata

Misure di effetto

Effetto: è la differenza nel verificarsi di una malattia in due gruppi di persone che differiscono riguardo a una caratteristica *causale* (esposizione) (Rothman, 1986)

Esempio:

Rischio relativo (RR) di malattia X in soggetti con pressione sistolica >170 mmHg (gruppo di riferimento soggetti con pressione ≤170 mmHg)

$$RR = \frac{18 \times 100}{9 \times 100} = 2$$

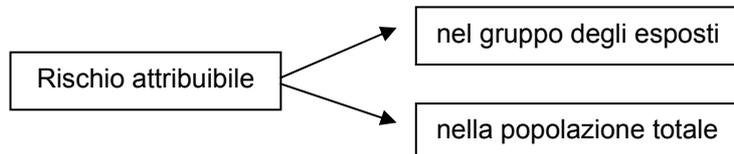
Effetti relativi → Rapporti di tassi → **Rischio relativo**

RR > 1 l'esposizione è un fattore di rischio

RR < 1 l'esposizione è un fattore protettivo

RR = 1 l'esposizione non è un fattore né di rischio né protettivo

Effetti assoluti → Differenze di tassi → **Rischio attribuibile**



Rischio attribuibile negli esposti

Esempio:

Uno studio prospettivo di malattie cardiovascolari (MCV)

Fattore in studio	MCV	Anni/persona	Incidenza per 1000 anni ⁻¹
Essere fumatore	84	2916	28,8
Essere non fumatore	87	4913	17,7
Popolazione totale	171	7829	21,8

Rischio attribuibile negli esposti:

(incidenza negli esposti – incidenza nei non esposti) x 1000
(0.0288 – 0.0177) x 1000 = 11.1 per 1000 anni⁻¹

Proporzione dell'incidenza negli esposti che è attribuibile all'esposizione

$$\frac{0,0288 - 0,0177}{0,0288} = \frac{0,0111}{0,0288} = 0,38 \text{ (38\%)}$$

Nota:

- Solo il 38% dei casi di MCV nei fumatori è attribuibile all'abitudine al fumo.
- Il rischio attribuibile implica che non tutta l'incidenza della malattia è dovuta all'esposizione perché anche persone non esposte sviluppano la malattia.

Rischio attribuibile nella popolazione totale

(incidenza nella popolazione totale – incidenza nei non esposti) x 1000
(0,0218 – 0,0177) x 1000 = 4,1 per 1000 anni⁻¹

Proporzione dell'incidenza nella popolazione totale che è "attribuibile" all'esposizione.

$$\frac{0,0041}{0,0218} = 0,19 \text{ (19\%)}$$

La cosiddetta Formula di Levin è una maniera alternativa per ottenere lo stesso risultato

$$\frac{p (r-1)}{p (r-1) + 1} = \frac{0,372 (1,63 - 1)}{0,372 (1,63 - 1) + 1} = 0,19 \text{ (19\%)}$$

p = proporzione della popolazione esposta = 37,2%

r = rischio relativo = 0,0288/0,0177 = 1,63

Nota:

- Il rischio attribuibile di popolazione risponde alla domanda: se nella popolazione fosse eliminata l'abitudine al fumo, che proporzione dell'incidenza della malattia sarebbe prevenuta? Il 19% dei casi di MCV.
- La formula di Levin si usa quando non sono noti i tassi d'incidenza.

Come si sommano i rischi?

Popolazione: 700 soggetti

Evento: malattia X

Esposizione: obesità

	Obesi	Non obesi
N	200	500
Rischio annuale	15 per 1000	10 per 1000

$$\text{Rischio annuale}_{\text{Totale}} = R_e P_e + R_o P_o$$

R_e = Rischio annuale della malattia negli esposti

P_e = Proporzione di esposti nella popolazione

R_o = Rischio annuale della malattia tra i non esposti

P_o = Proporzione dei non esposti nella popolazione

$$\text{Rischio annuale}_{\text{Totale}} = \frac{15}{1000} \cdot \frac{200}{700} + \frac{10}{1000} \cdot \frac{500}{700} = \frac{8}{700} = 0,0114$$

$$\text{Rischio annuale}_{\text{Totale}} = 0,0114 \times 1000 = 11,4 \text{ per } 1000$$

Nota:

- La somma dei tassi medi d'incidenza si esegue con un procedimento analogo.

Un'altra misura d'effetto: ODDS RATIO (OR)

Calcolo:

		Esposizione	
		+	-
Malattia X	+	a	b
	-	c	d

$$OR = \frac{a \cdot d}{b \cdot c}$$

Nota:

- L'OR è una buona stima del RR quando la malattia X è rara (< 10%).
- L'OR, a differenza del RR, è calcolabile negli studi caso-controllo (studi molto importanti per determinare i fattori di rischio delle malattie).

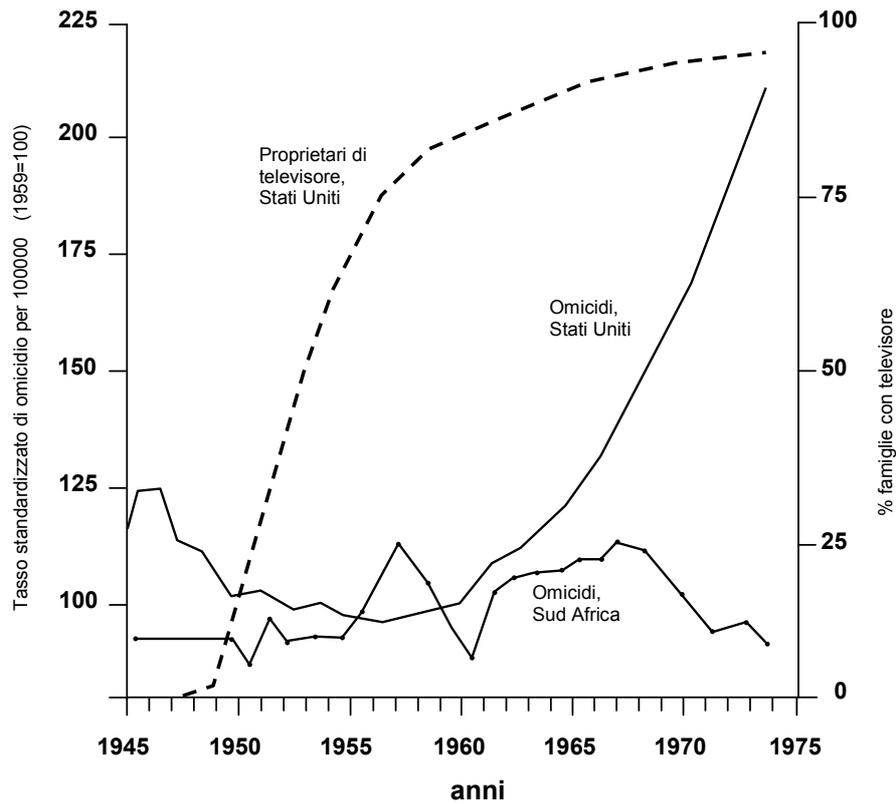
Come calcolare l'OR quando ci sono più di due categorie nel fattore in studio

Madre fumatrice n. sigarette	Casi*	Controlli*	OR?
< 10	71	2414	1,00
11-20	203	420	
21-30	55	86	
≥ 31	40	48	

* Casi = bimbi malformati. Controlli = bimbi normali.

(Derivato da: Kelsey, 1978)

Esempio di associazione non causale



Tasso di omicidio nella popolazione dei bianchi negli USA ed in Sud Africa in rapporto alla diffusione della televisione, 1945-1973 (in Sud Africa nel periodo considerato non c'era il servizio televisivo).

(Derivato da: Centerwall, 1989)

Definizione di causalità

La causa di un evento è qualcosa nei suoi antecedenti temporali che se non presente, impedirebbe il verificarsi dell'evento

Postulati di causalità delle malattie infettive (dalla Enciclopedia Britannica)

Microrganismo:

- 1) presente in ogni caso della malattia
- 2) non presente in altre malattie o in individui sani
- 3) isolabile in coltura pura in grado di riprodurre la malattia

(Henle-Kock)

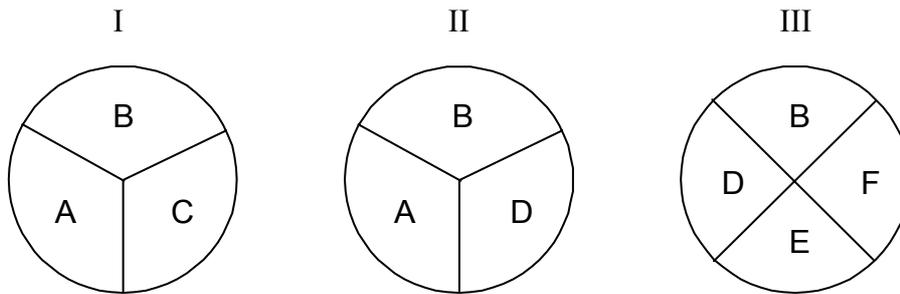
Attualmente “I criteri per il riconoscimento della relazione causa-effetto in medicina si sono al contempo complicati e indeboliti, e tale relazione ha assunto un carattere inequivocabilmente stocastico (probabilistico)”

(Vineis, 1990)

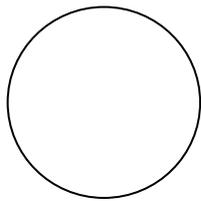
Le *Esposizioni* quando aumentano la frequenza relativa di malattia sono definite *Fattori di rischio* e non cause

Modello causale di una malattia

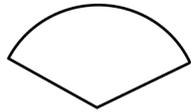
(da Rothman, 1986)



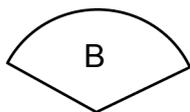
Schema delle tre cause sufficienti di una malattia in una certa popolazione



Causa sufficiente



Causa componente (fattore di rischio)



Causa necessaria (es. presenza dell'agente microbico in una malattia infettiva)

Una causa componente poco frequente nella popolazione tende ad essere maggiormente associata alla malattia.

Una causa componente tende ad essere tanto meno associata quanto più sono presenti nella popolazione cause sufficienti a cui non partecipa.

Il periodo necessario per riunire le cause componenti a formare le cause sufficienti definisce il tempo di induzione della malattia.

Criteria per stabilire una relazione causale

(Gordis, 1996)

- 1) Relazione temporale
- 2) Forza dell'associazione (RR, RA, OR)
- 3) Relazione dose-risposta
- 4) Replicazione dei risultati
- 5) Plausibilità biologica
- 6) Spiegazioni alternative (confondimento)
- 7) Cessazione dell'esposizione
- 8) Specificità dell'associazione
- 9) Concordanza con altre conoscenze

Relazione dose-risposta

Esempio:

Casi di cancro alla mammella e anni-persona a rischio secondo la dose di radiazioni in rad tra le donne di Hiroshima e Nagasaki che al momento dello scoppio della bomba atomica, avevano 10 o più anni di età.

	Radiazioni	Casi	Anni/persona	Incidenza (per 10⁶ anni⁻¹)	RR
	non esposti	38	208515	182	1
	0-9	105	463086	227	1,25
dose (rad)	10-99	48	164639	292	1,60
	100+	34	52185	652	3,58

(da Rothman, 1986)

Tipi di esposizioni a rischio di malattia

- a) Fattore di rischio: esposizione direttamente connessa col meccanismo patogenetico della malattia.

■ *Esempio*: l'asbesto per il mesotelioma

- b) Determinante: esposizione genericamente associata alla malattia in quanto variamente collegata con alcuni fattori di rischio.

■ *Esempio*: età, sesso, stato sociale, ecc.

- c) Proxi: esposizione facilmente rilevabile usata per valutare un fattore di rischio difficilmente misurabile.

■ *Esempio*: gli occhiali scuri come proxi della luce solare.

■ *Nota*:

- Più un'esposizione può essere controllata e più la sua conoscenza riveste importanza per la sanità pubblica.

Associazione tra sieropositività per Leptospira e rilevanti esposizioni per altitudine di residenza, Italia 1988

		Pianura* Odds Ratio corretti*** (IC 95%)	Collina/Montagna** Odds Ratio corretti*** (IC 95%)
Sesso	Femmina	1,52 (1,11-2,08)	0,90 (0,50-1,59)
	Maschio	1,00	1,00
Residenza	Campagna	1,11 (0,77-1,60)	5,07 (2,54-10,1)
	Area urbana	1,00	1,00
Allevamento di animali	Si	1,52 (0,72-3,31)	1,83 (0,54-6,21)
	No	1,00	1,00
Attività agricola	Si	0,79 (0,46-1,38)	0,94 (0,48-1,86)
	No	1,00	1,00
Corsi d'acqua	Si	1,24 (0,81-1,89)	2,00 (0,52-7,77)
	No	1,00	1,00
Contatti con gatti	Si	1,26 (0,88-1,80)	3,06 (1,51-6,22)
	No	1,00	1,00
Contatti con cani	Si	0,86 (0,60-1,24)	1,03 (0,59-1,82)
	No	1,00	1,00
Topi/ratti	Si	1,40 (0,86-2,27)	2,00 (0,57-7,01)
	No	1,00	1,00

* Sieroprevalenza per Leptospira = 17,0%

** Sieroprevalenza per Leptospira = 6,2%

*** Corretti per tutte le variabili in tabella

3. Sperimentazioni e studi osservazionali

Premessa

Le misure di frequenza e di effetto di una popolazione sono determinate mediante studi epidemiologici. Gli studi epidemiologici: osservazionali e sperimentali, comportano confronti tra gruppi di popolazione (pag. 46).

Gli studi osservazionali a differenza di quelli sperimentali possono portare a risultati distorti a causa dell'autoselezione. Nell'esempio i soggetti dei gruppi a confronto hanno deciso con l'aiuto di un medico dove allocarsi (autoselezione) a differenza di quelli della sperimentazione che sono stati allocati casualmente dai ricercatori (randomizzazione) (pag. 47).

Solo i risultati distorti da cause note (confondenti), nell'esempio l'aritmia, possono essere corretti (pag. 47).

Solo gli studi osservazionali sono eticamente leciti per lo studio delle cause di malattia (pag. 48).

Tra gli studi osservazionali quello di coorte è il più vicino alla sperimentazione. Comincia con soggetti senza la malattia (pag. 49-50).

Lo studio caso-controllo a differenza di quello di coorte è sempre retrospettivo rispetto all'occorrenza dell'evento (es. malattia), ossia lo studio comincia a partire dal verificarsi dei casi di malattia (pag. 51-52).

Lo studio trasversale è il migliore per descrivere lo stato di salute di una popolazione (pag. 53).

Lo studio ecologico non recluta soggetti da studiare, utilizza dati d'archivio. È molto economico, ma poco conclusivo (pag. 54).

Analogamente al rischio relativo che è basato sul rapporto di rischi, l'odds ratio è basato sul rapporto degli odds di malattia (pag. 55).

La formula: $(ad)/(bc)$ per il calcolo dell'odds ratio è solo uno schema mnemonico (pag. 56).

La stima dell'odds ratio è sempre superiore alla corrispondente del rischio relativo (pag. 57).

La differenza di stima tra odds ratio e rischio relativo diminuisce quanto più è minore l'incidenza della malattia nelle popolazioni in studio (pag. 58-59).

Gli studi retrospettivi e alcuni metodi di analisi consentono la stima di un'associazione solo mediante l'odds ratio (pag. 60).

Nello studio caso-controllo il diverso campionamento di casi e controlli porta ad una stima corretta dell'odds ratio di esposizione a patto che non si sia verificato un bias di selezione (pag. 61).

In uno studio caso-controllo è possibile solo il calcolo dell'odds ratio di esposizione, fortunatamente il suo valore è uguale a quello dell'odds ratio di malattia (pag. 62).

Struttura degli studi epidemiologici

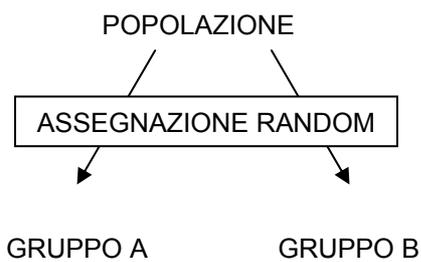
1) Obiettivi

- a. Identificare i fattori di rischio di malattie
- b. Valutare aspetti preventivi e terapeutici della pratica medica
- c. Valutare nuove attività di assistenza sanitaria

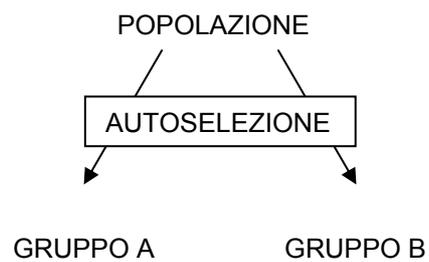
2) Tipi di studi

- a. Studi sperimentali (esperimento clinico controllato randomizz.)
- b. Studi osservazionali (coorte, caso-controllo, ecc.)

STUDIO SPERIMENTALE



STUDIO OSSERVAZIONALE



PROBLEMA

Una popolazione di 2000 pazienti con patologia coronarica, riceve per metà un nuovo trattamento. Dei 2000 pazienti, 700 hanno aritmia "X".

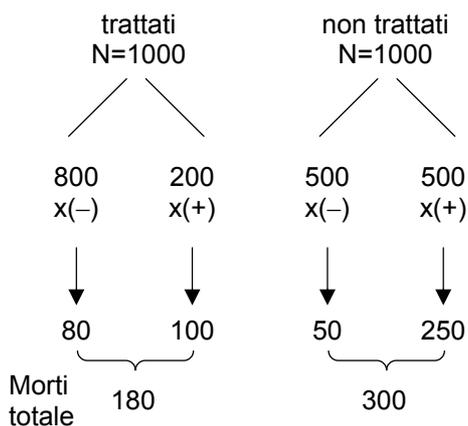
X (+) = pazienti con aritmia "X", hanno una mortalità del 50%

X (-) = pazienti senza aritmia "X", hanno una mortalità del 10%

Il trattamento non influisce sulla mortalità

STUDIO OSSERVAZIONALE

A causa dell'autoselezione* la proporzione di pazienti con aritmia nei due gruppi "può" essere differente



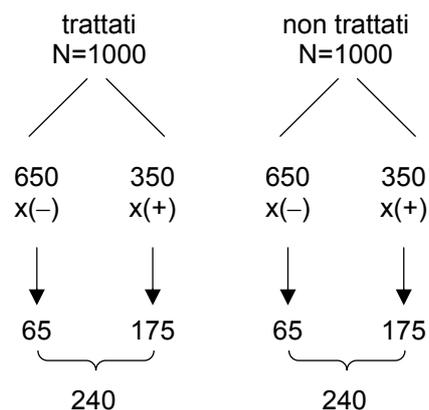
Mortalità

$$\frac{180 \times 100}{1000} = 18\% \quad \frac{300 \times 100}{1000} = 30\%$$

RISULTATO DISTORTO

STUDIO SPERIMENTALE

Per merito della randomizzazione, la proporzione di pazienti con aritmia nei due gruppi è simile



$$\frac{240 \times 100}{1000} = 24\% \quad \frac{240 \times 100}{1000} = 24\%$$

RISULTATO VALIDO

* In questo esempio i soggetti con aritmia tendono a non accettare un nuovo trattamento.

Principale limite degli studi sperimentali

Si possono studiare solo trattamenti potenzialmente benefici

PROBLEMA

☐ Come studiare le cause delle malattie?

SOLUZIONE

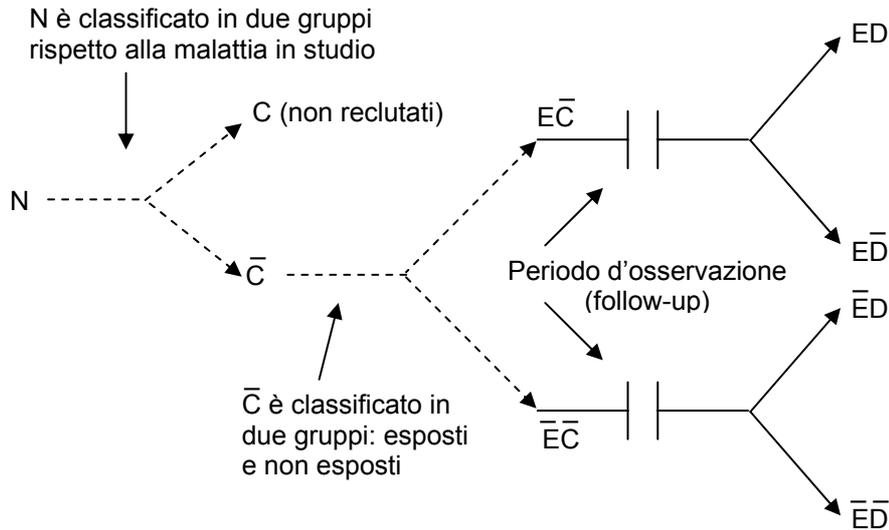
☐ Con gli studi osservazionali:
Studio di coorte
Studio caso-controllo
Studio trasversale
Studio ecologico
Ecc.

☐ Nota:

- Gli studi osservazionali possono dare risultati validi se si adottano le giuste precauzioni e correzioni.

Studio di coorte

(Schema derivato da: Kleinbaum, Kupper, Morgestein; 1982)



Legenda

- N = popolazione
- C = casi prevalenti
- D = casi incidenti
- \bar{C} o \bar{D} = non casi
- E = soggetti esposti
- \bar{E} = soggetti non esposti

		Esposizione		
		+	-	
Malattia	+	ED	$\bar{E}D$	D
	-	$\bar{E}D$	$\bar{E}\bar{D}$	\bar{D}
		E	\bar{E}	

Nota:

- Sinonimi: studio di follow-up, studio prospettivo.
- I casi di malattia possono essere avvenuti prima dell'inizio dello studio (studio retrospettivo di coorte, coorte storica) o dopo (studio prospettivo di coorte).
- Studio prospettivo di coorte: il più vicino all'esperimento.
- Studio retrospettivo di coorte: adatto per malattie rare o con un lungo periodo di latenza.
- Problema principale: perdita di soggetti a causa di migrazioni, mancanza di partecipazione, morte.
- La popolazione può essere fissa o dinamica.

Studio di coorte

Condizioni:

- Quando la malattia in studio è comune ($> 20\%$)
- Quando il periodo di latenza è relativamente corto
- Quando i “persi al follow-up” possono essere minimizzati

Vantaggi:

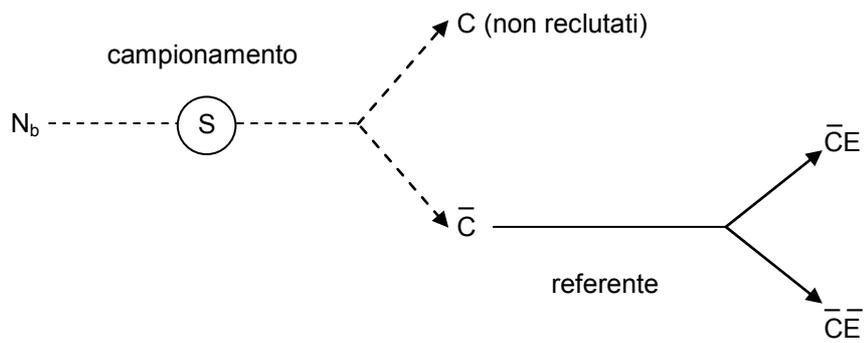
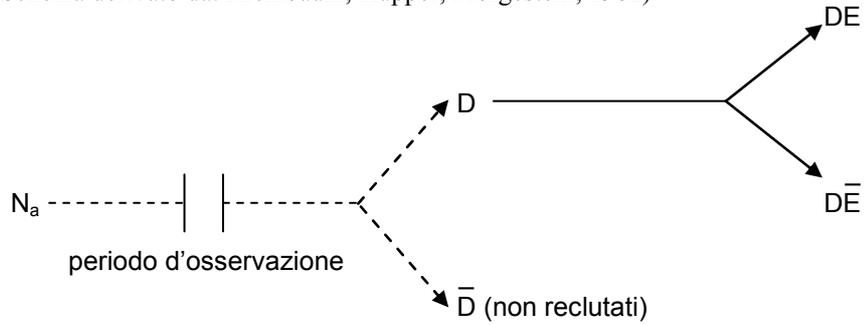
- Permette il calcolo di tassi e di tutte le misure di associazione
- Meno soggetto a certi bias riguardo alla sequenza temporale, accertamento dell'esposizione
- Utile se l'esposizione è rara

Svantaggi:

- Raccolta dati solitamente lunga e complessa
- Non adatto a verificare rapidamente nuove ipotesi
- Inadatto a studiare malattie rare

Studio caso-controllo

(Schema derivato da: Kleinbaum, Kupper, Morgestein; 1982)



Legenda

- N = popolazione
- C = casi prevalenti
- D = casi incidenti
- \bar{C} = controlli
- \bar{D} = non casi
- E = soggetti esposti
- \bar{E} = soggetti non esposti

		Esposizione		
		+	-	
Casi	DE	$\bar{D}\bar{E}$	D	
	$\bar{C}E$	$\bar{C}\bar{E}$	\bar{C}	
		E	\bar{E}	

Nota:

- Sinonimi: studio retrospettivo.
- N_a potrebbero essere i nuovi ricoverati presso una divisione ospedaliera, N_b i ricoverati di un'altra divisione dello stesso ospedale.

Studio caso-controllo

Condizioni:

- Questo studio è valido se:
 1. I controlli sono rappresentativi dei sani/non casi nella stessa popolazione dalla quale si sono sviluppati i casi
 2. L'esposizione non condiziona il campionamento dei controlli

Vantaggi:

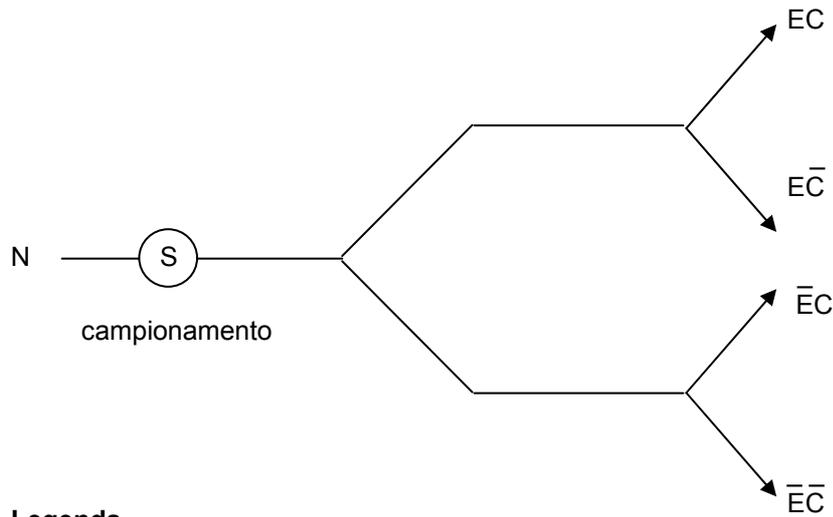
- adatto per malattie rare e con qualsiasi periodo di latenza
- relativamente rapido ed economico

Svantaggi:

- ad eccezione dell'Odds ratio, non produce misure di frequenza o d'associazione
- più complesso stabilire la sequenza temporale, problematico il reclutamento e l'accertamento dell'esposizione
- inadatto se l'esposizione è rara

Studio trasversale

(Schema derivato da: Kleinbaum, Kupper, Morgestein; 1982)



Legenda

N = popolazione

C = casi prevalenti

C̄ = non casi

E = soggetti esposti

Ē = soggetti non esposti

		Esposizione		
		+	-	
Malattia	+	EC	ĒC	C
	-	E C̄	Ē C̄	C̄
		E	Ē	

Nota:

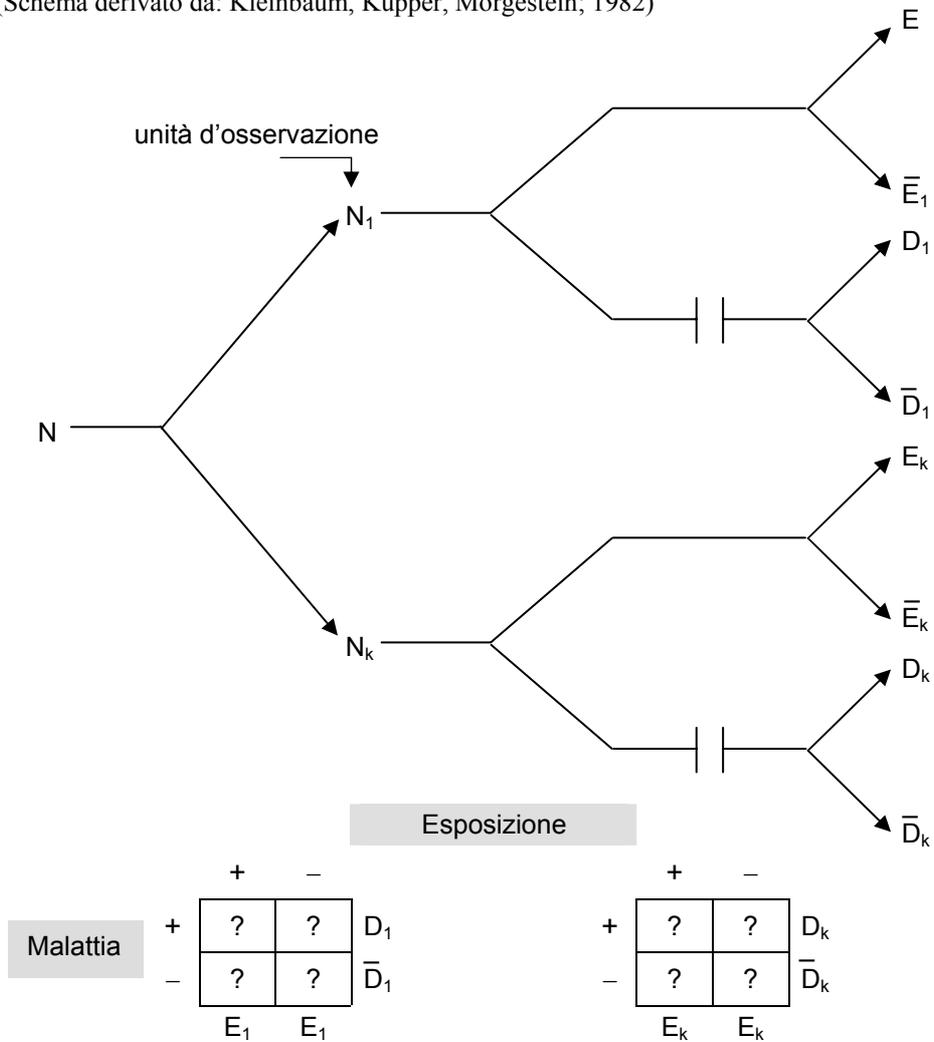
- Sinonimi: studio di prevalenza, survey, cross-sectional study.
- Scopi:
 - 1) descrivere lo stato della popolazione
 - 2) valutare ipotesi eziologiche

Svantaggi:

- inadatto per malattie rare o di breve durata
- non distingue tra fattori di rischio e fattori prognostici

Studio ecologico

(Schema derivato da: Kleinbaum, Kupper, Morgestein; 1982)



Nota:

- Studio a struttura incompleta.
- Sinonimo: studio di correlazione.
- L'unità d'osservazione non è un individuo, ma un gruppo: villaggio, provincia, Stato, ecc.

Vantaggi: economico, si basa su dati secondari

Svantaggi: "fallacia ecologica"

Relazione tra RR e OR

In uno studio prospettivo è possibile calcolare entrambi

Esempio:

		malattia		
		N	+	-
Fattore in studio (FS)	+	1000	100	900
	-	1000	25	975



 Follow-up

Probabilità (rischio di malattia)	= P	$\left\{ \begin{array}{l} \text{FS (+)} = 100/1000 = 0,100 \\ \text{FS (-)} = 25/1000 = 0,025 \end{array} \right.$	$\text{RR} = \frac{0,100}{0,025} = 4,0$
---	-----	--	---

Odds di malattia	= $\frac{P}{1-P}$	$\left\{ \begin{array}{l} \text{FS (+)} = 0,100/(1-0,100) = 0,1111 \\ \text{FS (-)} = 0,025/(1-0,025) = 0,0256 \end{array} \right.$	$\text{OR} = \frac{0,1111}{0,0256} = 4,33$
---------------------	-------------------	---	--

Odds (scommessa?):

Il rapporto della probabilità di un evento sulla probabilità del non evento.

Altro modo per calcolare l'odds di malattia tra gli esposti

$$\frac{D_e}{N_e} : \frac{\bar{D}_e}{N_e} = \frac{100}{1000} : \frac{900}{1000} = \frac{100}{900} = 0,1111$$

D_e = N casi esposti
 \bar{D}_e = N sani esposti
 N_e = Totale esposti

Analogamente tra i non esposti

$$\text{odds}_{\text{malattia}} = \frac{25}{975} = 0,0256$$

L'OR può essere calcolato anche col rapporto dei prodotti incrociati

		malattia	
		+	--
Fattore in studio (FS)	+	1000	100 (A) 900 (B)
	-	1000	25 (C) 975 (D)

$$\text{OR} = \frac{A}{B} : \frac{C}{D} = \frac{A \times D}{B \times C} = \frac{100 \times 975}{900 \times 25} = 4,33$$

Quando l'OR viene usato per stimare il RR c'è un errore sistematico, ossia un bias (si pronuncia baias)

$$OR = \frac{P_1 / (1 - P_1)}{P_2 / (1 - P_2)} = \frac{P_1}{1 - P_1} \times \frac{1 - P_2}{P_2} = \frac{P_1}{P_2} \times \frac{1 - P_2}{1 - P_1}$$



RR



Bias

Esempio precedente:

$$OR = 4,0 \times \frac{1 - 0,025}{1 - 0,100} = 4,33$$

OR > RR

La stima dell'OR è sempre superiore a quella del RR quando si misura l'associazione tra malattia e fattore di rischio

Entità dell'errore sistematico quando si stima il RR mediante OR

1. La probabilità della malattia tra gli esposti è bassa

Esempio:

		malattia		
		N	+	-
Fattore in studio (FS)	+	1000	10	990
	-	1000	5	995
		Follow-up →		

$\frac{P_1}{1-P_1} \approx P$	→	$\frac{10}{990} \approx \frac{10}{1000}$	Odds di malattia quasi uguale al rischio
-------------------------------	---	--	--

$$RR = \frac{P_1}{P_2} = \frac{10/1000}{5/1000} = 2,000$$

$$OR = \frac{P_1/(1-P_1)}{P_2/(1-P_2)} = \frac{10/990}{5/995} = 2,009$$

Conclusione

RR e OR quasi uguali

2. La probabilità della malattia tra gli esposti è alta

Esempio:

		malattia	
		+	-
Fattore in studio (FS)	+	1000	500
	-	1000	250
		Follow-up →	

$$\frac{P_1}{1-P_1} \neq P \longrightarrow \frac{500}{500} \neq \frac{500}{1000} \quad \text{Odds di malattia diverso dal rischio}$$

$$RR = \frac{500/1000}{250/1000} = 2,000$$

$$OR = \frac{500/500}{250/750} = 3,000$$

Conclusione

OR non è una buona stima di RR

Vantaggi di OR su RR

1. Unica stima possibile negli studi caso-controllo
2. Direttamente stimato dal modello logistico*

Svantaggi di OR su RR

- Il significato di OR è meno intuitivo

* modello matematico per una rappresentazione di un sistema, processo o relazione in forma matematica in cui sono usate equazioni per simulare il comportamento del sistema o del processo sotto studio

(Derivato da: Last, 1988)

In uno studio caso-controllo, un campione di casi e soggetti di controllo con diversa proporzione di campionamento porta comunque ad una stima corretta dell'OR di esposizione.

Esempio:

		Proporzione di campionamento	
		casi	controlli
Fattore in studio (FS)	+	100 (0,2) = 20	900 (0,1) = 90
	-	25 (0,2) = 5	975 (0,1) = 97,5

OR:

$$1. \text{ Popolazione totale} = \frac{100 \times 975}{900 \times 25} = 4,3$$

$$2. \text{ Campione} = \frac{20 \times 97,5}{90 \times 5} = 4,3$$

Nota:

- I risultati campionari sono validi perché la proporzione di campionamento è la stessa all'interno di ciascun gruppo, se fosse diversa i risultati sarebbero distorti a causa di un bias di selezione (vedi Lezione 4).

OR di malattia = OR di esposizione

(Derivato da: Kahn, Sempos; 1989)

Nelle tabelle seguenti sono ordinati prospettivamente e retrospettivamente le frequenze di uno studio di coorte* con 2000 partecipanti

		malattia					malattia		
		N	+	-	N	+	-		
Fattore in studio (FS)	+	1000	100	900	125	1875		Lettura retrospettiva (dalla malattia al fattore in studio)	
	-	1000	25	975	100	975			
		<div style="border-top: 1px solid black; margin-top: 5px;"></div> Lettura prospettiva (dal fattore in studio alla malattia)							

Prospettivamente

$$\text{OR di malattia} = \frac{100}{900} : \frac{25}{975} = \frac{100 \times 975}{900 \times 25} = 4,3$$

Retrospettivamente

$$\text{OR di esposizione} = \frac{100}{25} : \frac{900}{975} = \frac{100 \times 975}{900 \times 25} = 4,3$$

* L'esempio riguarda uno studio prospettivo, per il quale è possibile il calcolo di entrambi gli OR di malattia e di esposizione, ma l'uguaglianza verificata è vera anche per gli studi caso-controllo.

Riassumendo:

Quando l'incidenza annuale della malattia in studio tra gli esposti è $< 10\%$ abbiamo visto che:

$$RR \approx \text{OR di malattia}$$

Poiché $\text{OR di malattia} = \text{OR di esposizione}$

ne deriva che:

$$RR \approx \text{OR di esposizione}$$

Conclusion

L'unica misura di associazione calcolabile negli studi caso-controllo, cioè l'OR di esposizione, consente la stima del Rischio Relativo, quando la malattia in studio è sufficientemente rara.

4. Definizioni di bias, effetto di più fattori di rischio

Premessa

Quando si vuole criticare uno studio epidemiologico il pensiero passa in rassegna i bias di selezione, i bias d'informazione, fenomeni di confondimento e l'imprecisione delle stime, quest'ultimo connesso alla variabilità statistica è trattato nel Corso della dott.ssa Stazi (pag. 67).

Il bias d'informazione si ha quando la qualità della raccolta dati è diversa nei due gruppi a confronto (pag. 67).

Il bias di selezione porta a scegliere un gruppo in studio che non riflette le caratteristiche d'interesse della popolazione che dovrebbe rappresentare (pag. 67).

Secondo il tipo di situazione che ha portato a un bias d'informazione si distinguono bias da ricordo, da intervistatore, da osservatore (pag. 68).

Secondo il tipo di situazione che ha portato a un bias di selezione si distinguono bias da individuazione selettiva, bias da selezione di casi prevalenti (pag. 68).

Il fenomeno del confondimento non è solitamente legato a scelte di progettazione dello studio, ma alle caratteristiche della popolazione (pag. 69).

Lo schema più semplice di confondimento considera due fattori di rischio della stessa malattia, uno in studio, l'altro noto (confondente) (pag. 69).

In Tabella 1, l'associazione tra l'esposizione al fattore in studio e la malattia è stimata da un odds ratio grezzo pari a 1,95 (pag. 70).

In Tabella 2 si vede che i controlli tendono ad essere più giovani dei casi, ossia l'età è un fattore di rischio della malattia (pag. 70).

In Tabella 3 si vede che il gruppo con meno di quarant'anni è stato meno esposto, rispetto al gruppo più anziano, ossia l'esposizione è associata all'età (potete calcolare l'odds ratio, è diverso da 1) (pag. 70).

L'età è un fattore di confondimento dell'associazione tra la malattia e l'esposizione al fattore in studio. In Tabella 4 è stato eliminato il confondimento mediante stratificazione per età. L'associazione stimata dall'odds ratio grezzo (vedi Tabella 1) era dovuta unicamente al confondimento dovuto all'età. Tolto il confondimento non c'è associazione (odds ratio corretto = 1) (pag. 71).

Ci sono vari metodi per eliminare il confondimento dovuto a fattori di rischio noti, uno solo per eliminare il confondimento di fattori di rischio non noti: l'allocatione random dei soggetti nei gruppi a confronto degli studi sperimentali (pag. 72).

La standardizzazione è un metodo per l'eliminazione del confondimento nel confronto tra tassi (pag. 73).

Solitamente le pubblicazioni scientifiche riportano sia le misure grezze che le corrispondenti corrette dal confondimento (pag. 74).

Immaginiamo che una malattia sia associata a due fattori di rischio e di conoscere il rischio per l'esposizione a ciascuno dei due fattori, ma non il rischio complessiva dell'esposizione ad entrambi (pag. 75-76).

Esistono due modi teorici per la stima complessiva, basati su un modello additivo, o su un modello moltiplicativo. Se il rischio teorico si rivelassero diversi dal rischio complessivo reale si parlerebbe di "interazione", ossia si direbbe che l'effetto di un

fattore modifica quello dell'altro, rendendo teoricamente imprevedibile il livello dell'effetto complessivo.

Anche se le condizioni che producono confondimento e interazione hanno delle somiglianze i due fenomeni sono diversi e indipendenti. Sapere della presenza dell'uno non rende più o meno probabile l'altro (pag. 77).

La stima dell'associazione tra esposizione al fattore di rischio e malattia può essere distorta a causa di:

- Errori sistematici ➡ (*bias*)
- Coincidenza con altre esposizioni ➡ (*confondimento*)
- Errori non sistematici ➡ (*variabilità statistica*)

Errori sistematici

Distorsione da bias di informazione

Difetto nella misura dell'esposizione o dell'evento che porta a *differente* qualità di informazione tra i gruppi a confronto.

(Last, 1988)

Distorsione da bias di selezione

Errore dovuto a differenza sistematica nelle caratteristiche (di esposizione o malattia) tra coloro che sono selezionati per lo studio e coloro che non lo sono.

(Last, 1988)

Esempio:

Associazione spuria tra caffè e cancro del pancreas

Tipi di distorsione da bias d'informazione

Distorsione da ricordo:

differente capacità nei casi e nei controlli di ricordare un'esposizione passata.

Esempio:

l'associazione tra rosolia nella madre e malformazione congenita nel figlio era più grande in uno studio retrospettivo che prospettivo.

Distorsione da intervistatore:

la conoscenza dello stato di caso o controllo potrebbe influenzare la maniera in cui sono poste le domande sull'esposizione.

Distorsione da osservatore:

quando lo stato di esposizione, noto all'osservatore, influenza la diagnosi.

Esempio:

la conoscenza di esposizione ad alcool potrebbe aumentare la probabilità di diagnosi di "cirrosi epatica".

Tipi di distorsione da bias di selezione

Distorsione da individuazione selettiva:

se la malattia tende ad essere subclinica, saranno maggiormente selezionate persone sotto sorveglianza medica, quest'ultima potrebbe essere dovuta all'esposizione.

Esempio:

flebite e contraccettione orale: i medici potrebbero visitare più spesso e attentamente le donne che fanno uso della pillola.

Distorsione da selezione di casi prevalenti

- a) L'esposizione è causata dalla malattia
Esempi: - Estrogeni e cancro dell'endometrio
- Parametri chimico-clinici negli studi di prevalenza
- b) L'esposizione non è solo associata all'insorgenza, ma anche alla durata della malattia.

Distorsione da non partecipazione o da abbandono dello studio

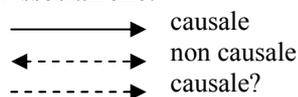
Confondimento

Il confondimento è una distorsione della misura di effetto del fattore in studio dovuta all'effetto di un altro fattore (confondente).

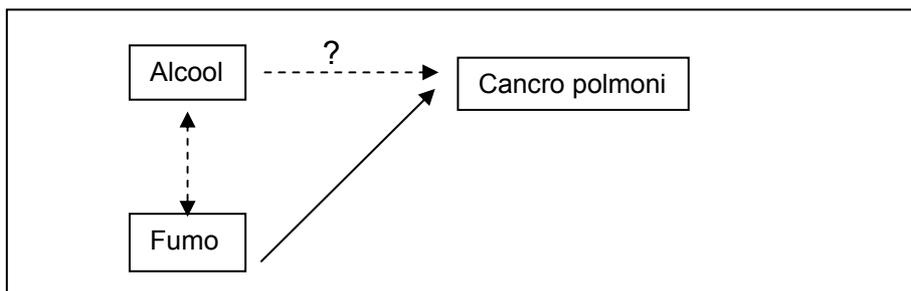
Un fattore per essere confondente deve:

- 1) essere un fattore di rischio per la malattia fra i non esposti
- 2) essere associato con l'esposizione nella popolazione da cui derivano i casi
- 3) non essere un anello intermedio della catena causale che lega l'esposizione alla malattia

Associazione:



Esempio 1: Presenza di confondimento del Fumo sul possibile effetto dell'Alcool



Esempio 2: Assenza di confondimento perché "Basso peso alla nascita" è nella catena causale

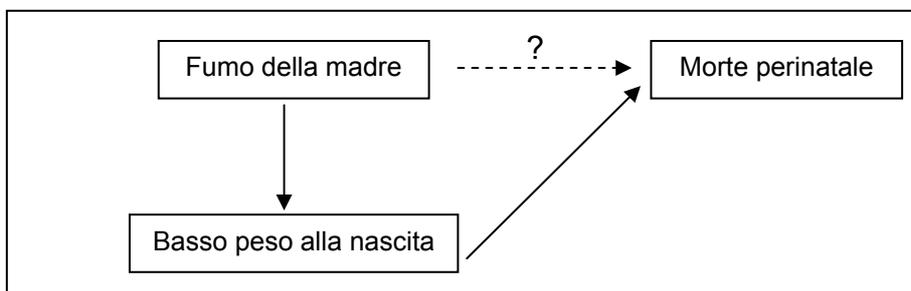


Tabella 1

Esempio ipotetico di confondimento da età in uno studio caso-controllo

Esposizione di casi e controlli al fattore in studio

Esposti	Casi	Controlli
Sì	30	18
No	70	82
Totale	100	100

$$OR = \frac{30 \times 82}{18 \times 70} = 1,95$$

Tabella 2

Esempio ipotetico di confondimento da età in uno studio caso-controllo

Distribuzione di casi e controlli per età

Età (anni)	Casi	Controlli
< 40	50	80
≥ 40	50	20
Totale	100	100

Nota: L'età è un fattore di rischio

Tabella 3

Esempio ipotetico di confondimento da età in uno studio caso-controllo

Relazione tra esposizione ed età

Età (anni)	Esposti	Non esposti	% esposti
< 40	13	117	10%
≥ 40	35	35	50%

Nota: L'età è associata all'esposizione

Tabella 4

Esempio ipotetico di confondimento da età in uno studio caso-controllo

Esposizione di casi e controlli al fattore in studio per strati di età

Età (anni)	Esposti	Casi	Controlli
< 40	Si	5	8
	No	45	72
	Totale	50	80
$OR = \frac{5 \times 72}{8 \times 45} = 1,0$			
≥ 40	Si	25	10
	No	25	10
	Totale	50	20
$OR = \frac{25 \times 10}{10 \times 25} = 1,0$			

Nota:

- L'OR corretto per l'effetto dell'età è diverso dall'OR grezzo calcolato in Tabella 1.
- La correzione è consistita nella stratificazione secondo le categorie della variabile confondente.

Come prevenire il confondimento (all'inizio dello studio):

- restrizione a un gruppo omogeneo per il fattore confondente della popolazione in studio
- appaiamento di casi e controlli per il fattore di confondimento
- randomizzazione (sperimentazione clinica)

Come rimuoverlo (alla fine dello studio):

- stratificazione dei dati
- standardizzazione dei dati
- analisi con modelli matematici (es. regressione logistica)

Esempio di confronto di tassi grezzi, specifici per età, e standardizzati

	Morti per 1000 per anno	
	California	Maine
Tasso grezzo	8,3	11,1
Tassi specifici		
Età (anni)		
< 15	1,6	1,9
15- 24	1,3	1,1
25- 34	1,5	1,4
35 -44	2,8	2,9
45- 54	6,7	6,9
55- 64	15,4	17,3
65-74	32,8	39,0
75+	91,7	104,1
Tasso standardizzato per età	8,8	9,9

Nota:

L'età è un confondente della associazione tra mortalità e residenza in quanto tasso grezzo e standardizzato sono diversi.

Esempio di tabelle che riportano Odds ratio grezzi e corretti

Abitudine al fumo durante la gravidanza in donne che hanno avuto un parto spontaneo pretermine e in controlli. Boston, 1977-80

Livello di esposizione	% soggetti controllo (n=2252)	Casi (n. 307)		
		%	OR*	IC 95%*
Mai fumato	48,7	37,8	1,0	
Smesso: prima del concepimento	22,8	19,5	1,1	0,8-1,5
durante il primo trimestre	5,2	6,2	1,5	0,9-2,6
Fumato durante tutta la gravidanza	23,3	36,5	2,0	1,5-2,7

* Odds ratio grezzi (OR) e intervalli di confidenza (IC) sono stati stimati confrontando ciascuna categoria di esposizione col gruppo di non esposti.

Parto spontaneo pretermine in relazione all'abitudine al fumo. Boston, 1977-80

Livello di esposizione	OR*	IC 95%
Mai fumato	1,0	
Smesso: prima del concepimento	1,4	0,9-2,0
durante il primo trimestre	1,6	0,8-2,9
Fumato durante tutta la gravidanza	2,2	1,4-3,5

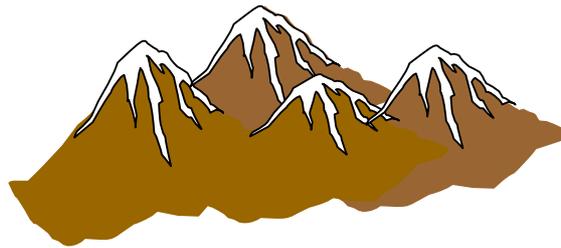
* Odds ratio corretti per razza, scolarità, età, condizioni socio-economiche, stato civile, uso di alcol, marijuana, parità, aborti indotti e spontanei, incompetenza cervicale, emorragie, body mass index, consumo di caffè

(Derivato da: Williams, 1992)

Se un soggetto è esposto a due fattori di rischio della stessa malattia?

Esempio:

Malattia: Steatosi delle Ande*



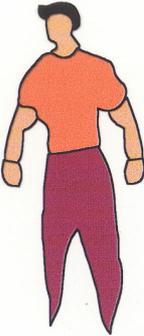
**malattia degenerativa probabilmente dovuta all'uso di tapapa (distillato alcolico) e/o burro di lama*



Popolazione in studio: abitanti del villaggio di Tutzuotam



Tabù alimentari:

Uomini		Donne		Soldati		Servi/e	
							
Tapapa	Sì	Tapapa	No	Tapapa	Sì	Tapapa	No
Burro	No	Burro	Sì	Burro	Sì	Burro	No

Tasso di Steatosi delle Ande a Tutzuotam:

Servi/e: 10%
 Donne: 30%
 Uomini: 25%
 Soldati: segreto militare

Tasso teorico nei soldati:

Assunzioni: burro e tapapa sono fattori di rischio indipendenti della steatosi andina.
Il sesso non è un fattore di rischio della steatosi andina

1) Possibile calcolo:

$$\begin{aligned} \text{Effetto assoluto della TAPAPA} &= (0,25 - 0,10) \times 100 = 15\% \\ \text{Effetto assoluto del BURRO} &= (0,30 - 0,10) \times 100 = 20\% \end{aligned}$$

↑
Tasso di base (Servi/e)

$$\text{Tasso nei SOLDATI} = (0,15 + 0,20 + 0,10) \times 100 = 45\%$$

↑ ↑ ↑
Attribuibile Attribuibile Tasso di base
alla TAPATA al BURRO stimato sui servi

2) Altro possibile calcolo:

$$\begin{aligned} \text{Effetto relativo della TAPAPA} &= 0,25 : 0,10 = 2,5 \\ \text{Effetto relativo del BURRO} &= 0,30 : 0,10 = 3,0 \end{aligned}$$

← ←

$$\text{Tasso nei SOLDATI} = [(2,5 \times 3,0) \times 0,10] \times 100 = 75\% \quad \text{RR RR}$$

↑
7,5 volte l'incidenza di base

Conclusion

- Se il valore vero del tasso dei soldati è simile a quello ottenuto dal calcolo 1) si dice che l'unione dei due effetti segue un "modello additivo".
- Se è simile a quello del calcolo 2), segue un modello moltiplicativo.
- Altrimenti si dice che c'è "interazione", ossia non è vero che i due effetti siano tra loro indipendenti.

5. Studio di un'epidemia

Premessa

Tra epidemie ed Epidemiologia esiste un ovvio legame, ma il concetto di epidemia si è andato evolvendo sulle base delle necessità della sanità pubblica, ad esempio ora anche un unico caso potrebbe costituire un'epidemia (pag. 80-81).

Si studiano le epidemie per ragioni di sanità pubblica o di ricerca (pag. 82).

La ragione dell'interesse è che i casi epidemici, a differenza di quelli endemici, sono tutti dovuti ad una stesso fattore di rischio a cui in precedenza la popolazione non era esposta (pag. 82).

Esistono false epidemie, i casi in eccesso potrebbero essere dovuti a nuovi sistemi di diagnosi o a nuove definizioni della malattia (pag. 82).

Un epidemia si studia considerando dapprima elementi comuni della storia dei casi e le loro caratteristiche (pag. 83-87). Dalle informazioni disponibili si definiscono le caratteristiche della popolazione a rischio (pag. 83).

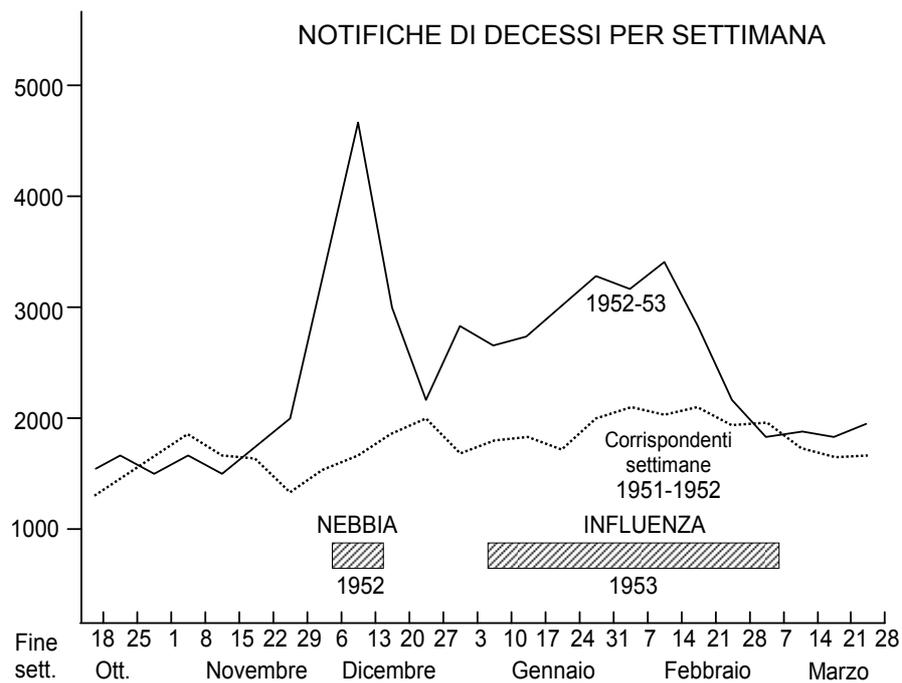
Una buona definizione della popolazione a rischio consente di calcolare tassi di attacco elevati. Il tasso d'attacco ideale sarebbe uguale a 1, ossia la l'esposizione ai fattori di rischio scoperti ha fatto ammalare tutte le persone a rischio (pag. 83).

Se la definizione della popolazione a rischio è troppo generica, ma il valore del tasso d'attacco è elevato, si possono calcolare tassi d'attacco specifici per le esposizione a diversi potenziali fattori di rischio (approccio di coorte) (pag. 88-89).

Nel caso il valore del tasso d'attacco sia troppo piccolo, si rinuncia al calcolo di tassi specifici, preferendo un approccio caso-controllo (pag. 88).

Studio di un'epidemia

EPIDEMIA: il verificarsi in una comunità o area geografica di un'affezione chiaramente in eccesso rispetto all'attesa normale.



Effetti a Londra della grande nebbia del dicembre 1952 e dell'epidemia di influenza avvenuta i primi mesi del 1953.

(Derivato da: Mausner, Kramer; 1985)

Gli anglofoni oltre al termine "Epidemic" usano il termine "outbreak" quando l'epidemia è circoscritta.

**Se l'eccesso di casi non è sicuro (fluttuazioni casuali?)
si parla di un "CLUSTER" (it.: gruppo) di casi.**

Quanti casi in eccesso?

**Anche un caso se la malattia non era precedentemente
presente**

Esempio:

1 caso di poliomielite in Italia, ora

**Le epidemie sono dovute solo a malattie trasmissibili acute?
No**

Esempio:

13 casi di mesotelioma in provincia di Firenze tra il 1979-1984 mai osservati in precedenza

(Paci E. *et al.*, 1987)

Si dice che gli epidemiologi cavalcano la parte discendente della curva epidemica, cioè arrivano a cose fatte.

È solitamente vero.

Allora perché studiare un'epidemia?

- Per limitare o prevenire nuovi casi di malattia
- Per soddisfare a una richiesta politica e del pubblico
- Per scoprire nuove malattie (AIDS, malattia del legionario) o nuove cause di malattie (Talidomide)

Come si scopre un'epidemia?

- dai dati di attività di sorveglianza
- dai registri d'ospedale o di laboratorio
- dalle assenze da scuola o dal lavoro
- dai certificati di morte
- dalla lettura mattutina del giornale

ATTENZIONE!

- Nuova diagnosi più sensibile \Rightarrow falsa epidemia
- Nuova definizione più ampia di una malattia soggetta a notifica \Rightarrow falsa epidemia

Come si studia un'epidemia?

Solitamente ci sono due fasi:

- 1) descrittiva (indiziaria)
- 2) analitica (probatoria)

1. Fase descrittiva

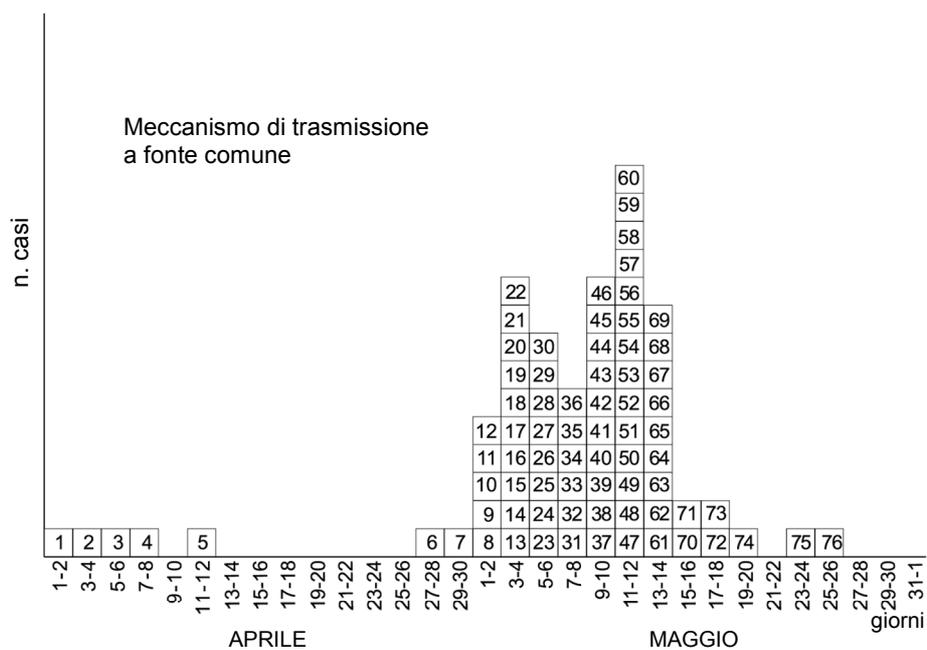
- rintracciare il massimo numero di casi (definire il caso attraverso segni e sintomi, e non tramite una specifica diagnosi)
- descrivere i casi in termini di tempo, luogo, persona

Nota bene:

La descrizione è principalmente quantitativa, attraverso:

- enumerazione dei casi
- tasso d'attacco (N di casi/ N persone a rischio di sviluppare la malattia all'inizio dell'epidemia)
- tasso d'attacco specifico (N di casi con una certa caratteristica/ N di persone a rischio con la stessa caratteristica all'inizio dell'epidemia)

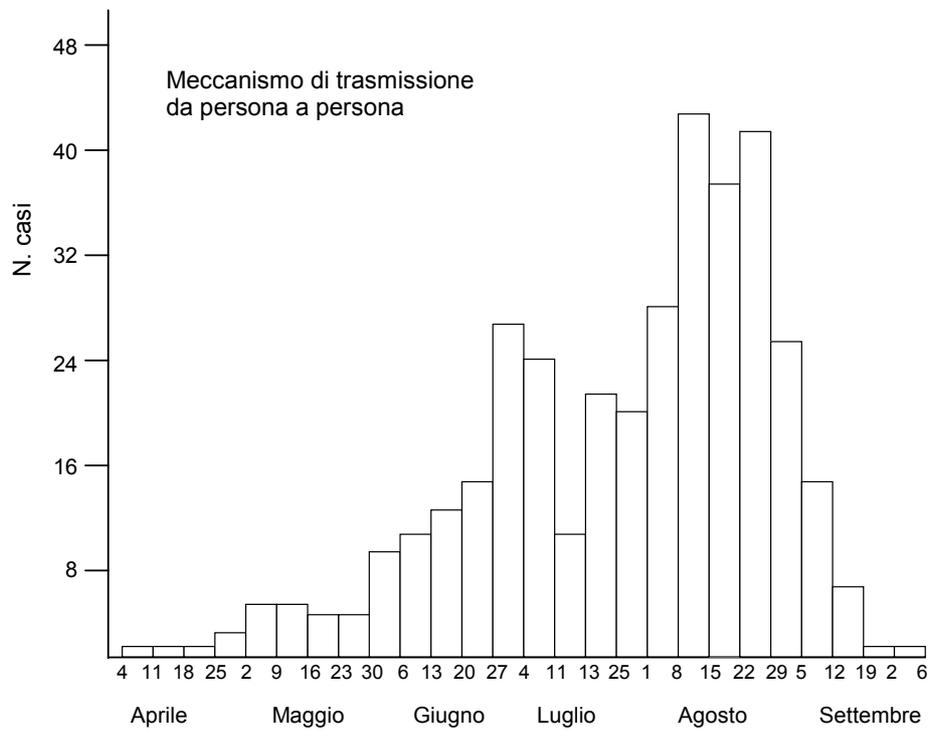
Descrizione temporale dei casi: alcuni esempi di “curve epidemiche”



Curva epidemica dei casi di epatite virale A, USL Monti del Matese, Aprile-Maggio 1982

Nota:

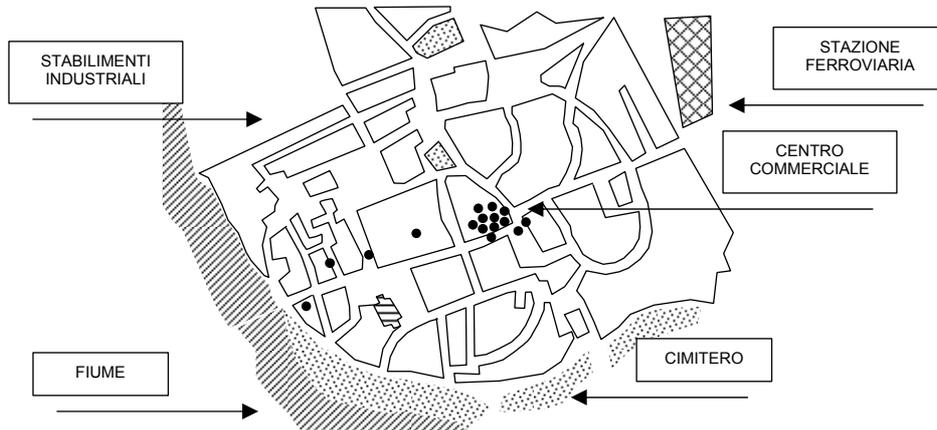
All'interno delle celle è riportato il codice identificativo del caso.



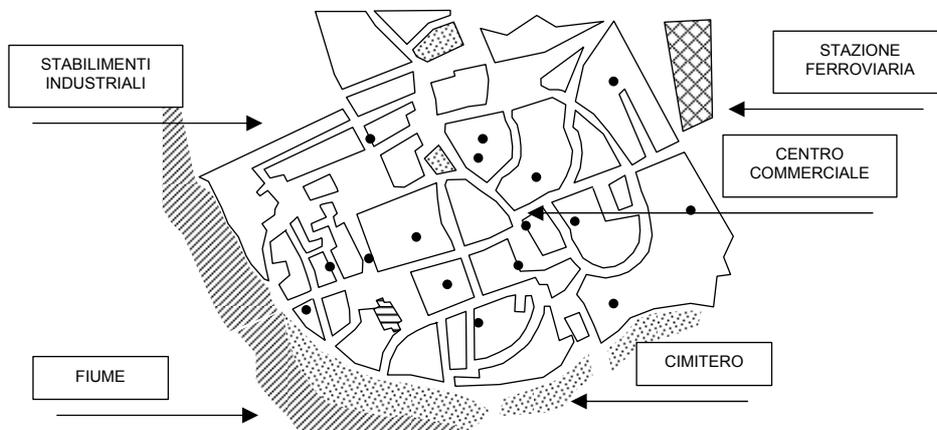
Epidemia di epatite infettiva in un istituto per soggetti ritardati di mente per settimana d'insorgenza

(Derivato da: Matthew EB. *et al.*, 1973)

Descrizione per luogo dei casi



Distribuzione dei casi di salmonellosi per luogo lavorativo, o per luogo di residenza nel caso di disoccupati.



Distribuzione dei casi di salmonellosi per luogo di residenza nello stesso periodo.

Descrizione per caratteristiche personali

Esempio:

Casi di epatite A e percentuale di ammalati per grado di istruzione nelle scuole di West Branch e St. Joseph, nella provincia di Ogemaw, nell'aprile-maggio 1968.

	Scuola di West Branch			Scuola di St. Joseph		
	N. casi	Alunni	Proporzione di ammalati	N. casi	Alunni	Proporzione di ammalati
Asilo	2	126	1,6			
1 ^a elementare	0	128	0	1	37	2,7
2 ^a elementare	0	121	0	1	41	2,4
3 ^a elementare	0	107	0	0	37	0
4 ^a elementare	2	106	1,9	0	26	0
5 ^a elementare	1	120	0,8	0	30	0
1 ^a media inf.	1	111	0,9	0	32	0
2 ^a media inf.	3	110	2,7	0	26	0
3 ^a media inf.	6	120	5,0	0	21	0
1 ^a media sup.	7	143	4,9			
2 ^a media sup.	10	128	7,8			
3 ^a media sup.	2	112	1,8			
4 ^a media sup.	2	93	2,1			
Totale	36	1525	2,4	2	250	0,8

Proporzione di ammalati = tasso d'attacco specifico per grado di istruzione

2. Fase analitica

Scopo: verificare l'associazione tra il fattore di rischio indiziato e la malattia

Metodo: dividere la popolazione a rischio in due gruppi e confrontarla

- Confronto rispetto alla malattia tra esposti al fattore di rischio vs non esposti \Rightarrow studio di coorte (storica)

oppure

- Confronto rispetto all'esposizione tra colpiti e non colpiti dall'epidemia \Rightarrow Studio caso-controllo

Quale tra i due scegliere:

- popolazione a rischio ben definita (tasso d'attacco $>30\%$)



Studio di coorte (storica)

- popolazione a rischio non ben definita (tasso d'attacco $<10\%$)



Studio caso – controllo

Tassi d'attacco per salmonellosi, specifici per alimento, tra i partecipanti ad un banchetto di nozze

	GRUPPO A Persone che mangiarono				GRUPPO B Persone che non mangiarono			
	Malati	Sani	Totale	T.A.%	Malati	Sani	Totale	T.A.%
Prosciutto	20	35	55	36	1	---	1	100
Melone	18	32	50	36	3	3	6	50
Crespelle	20	25	45	44*	1	10	11	9*
Tagliolini	20	30	50	40	1	5	6	16
Spaghetti	10	15	25	40	11	20	31	35
Cernia	15	25	40	37	5	10	15	33
Patate	8	11	19	42	13	24	37	35
Filetto	18	31	49	36	3	4	7	42
Funghi	14	22	36	38	7	13	20	35
Insalata	19	27	46	41	2	8	10	20
Pesca melba	18	31	49	36	3	4	7	42
Torta	19	27	46	41	2	8	10	20
Ghiaccio	15	26	41	36	6	9	15	40

* test esatto di Fisher (2 code): p=0,04

6. Screening e validità di una misura

Premessa

Anche se le attività di prevenzione appaiono desiderabili in realtà pongono problemi etici (es. limitazione della libertà individuale), di opportunità, e di costo-beneficio a livello individuale e della popolazione. Chi propone uno screening, strategia alla base della prevenzione secondaria, deve prepararsi a rispondere a parecchie obiezioni (pag. 92-93).

Il test di screening deve avere precise caratteristiche. I due diagrammi mostrano come si distribuiscono i risultati di cinque repliche di due test quantitativi (pag. 94).

Un esame di laboratorio o strumentale, quantitativo alla base di un test di screening è reso dicotomico stabilendo un valore (*cutpoint*) che definisce due intervalli che indicano positività o negatività circa il sospetto diagnostico. Dato che il risultato di un test di screening deve essere dicotomico la sua accuratezza si valuta come Sensibilità e Specificità (pag. 95).

Il disegno di Brown è una brillante metafora di uno studio di valutazione inteso a stimare l'accuratezza di un test che separa le sfere dai cubi (pag. 95).

Le quattro proporzioni: a, b, c, d per il calcolo della sensibilità e specificità si possono utilmente ordinare in una tabellina 2x2. Notare i totali in basso, (a+c) diviso (a+b+c+d) forniscono la stima del tasso di prevalenza della popolazione a cui si applica lo screening (pag. 96).

Sensibilità e Specificità sono caratteristiche utili in quanto costanti del test di screening, ma un'altra misura ricavabile dalla tabellina 2x2 è il valore predittivo positivo che è molto più informativo (pag. 96).

Il valore predittivo positivo dà una valutazione contestuale rispetto alle varie popolazioni a cui potrebbe essere applicato lo screening circa la sua capacità di diagnosi (pag. 97).

Lo scenario dell'esempio a pagina 89 potrebbe essere il seguente. All'inizio il clinico per mettere alla prova il nuovo test seleziona casi dubbi che hanno quindi un elevato tasso (vero) di prevalenza e un elevato valore predittivo positivo, in seguito applica il test ad una popolazione "naturale" il valore predittivo positivo risulta deludente (pag. 98).

L'esercizio mostra come solitamente alla base di uno screening ci sia un test presuntivo conveniente seguito da una conferma diagnostica (pag. 99).

Considerazioni preliminari allo screening

- È disponibile un test diagnostico?
- È applicabile a un problema sanitario serio?
- Qual è la prevalenza della fase preclinica della malattia?
- Qual è la percentuale di quelli positivi al test che hanno la malattia (valore predittivo positivo)?
- Lo stress fisico e psichico di chi si sottopone al test è compensato dalla diminuzione del rischio di mancata diagnosi?
- Diagnosi precoce e terapia tempestiva migliorano sostanzialmente l'esito (studio di valutazione)?
- Qual è il costo di ciascun caso vero diagnosticato dallo screening?
- Il costo è giustificato dal miglioramento dell'esito della malattia (studio di valutazione)?

IL CASO ETICO

Un neo due volte sospetto

Il caso

Diverse associazioni di categoria dei medici dermatologi italiani hanno dichiarato «Skin Cancer Day» il 6 maggio 2000 e hanno chiesto alle Aziende sanitarie dei capoluoghi di provincia di mettere a disposizione dell'iniziativa un certo numero di ambulatori e un'attrezzatura minima. In questo modo è stata offerta alla popolazione la possibilità di farsi visitare per identificare melanomi e carcinomi cutanei precedentemente non diagnosticati. Una semplice visita dermatologica può identificare lesioni sospette, da rendere oggetto di un successivo iter

diagnostico, inteso a discriminare tra «veri» e «falsi» positivi. Di questi ultimi non è possibile stimare *a priori* la rilevanza numerica. Quel che è certo è che essi daranno il via a procedimenti diagnostici che altrimenti non sarebbero stati necessari, con conseguente costo economico per il Servizio sanitario nazionale e ansia per i portatori. È plausibile che l'anticipo diagnostico di melanomi e carcinomi cutanei ne diminuisca la letalità, ma di ciò non vi è una prova scientifica. Neppure si è in grado di ipotizzare l'entità dell'eventuale riduzione.

- ❖ LA PROPOSTA DEI DERMATOLOGI ITALIANI CONSENTE UN'ANALISI ETICA IN TERMINI DI BENEFICENZA, NON-MALEFICENZA, AUTONOMIA E GIUSTIZIA? SE CIÒ NON È POSSIBILE, QUALI ULTERIORI ELEMENTI SAREBBERO NECESSARI?
- ❖ È IL CASO DI AVVIARE UN INTERVENTO DESTINATO FORSE A BENEFICIARE UNA PARTE LIMITATA DELLA POPOLAZIONE?
- ❖ SONO ACCETTABILI PROCEDURE DIAGNOSTICHE IN SOGGETTI ASINTOMATICI, DESTINATE DA UN LATO A RASSICURARE UNA GRAN PARTE DI CHI VI SI SOTTOPONE MA, DALL'ALTRO, A CREARE ANSIA NEI RESTANTI SOGGETTI?
- ❖ È ACCETTABILE CHE UN GRUPPO DI SPECIALISTI AVVII UN'INIZIATIVA, FORSE CON ALCUNI RISULTATI POSITIVI, DESTINATA AD AUMENTARE LA SPESA DEL SERVIZIO SANITARIO NAZIONALE, VUOI ATTRAVERSO GLI ESAMI DI SECONDO LIVELLO, VUOI ATTRAVERSO L'UTILIZZO DI STRUTTURE AMBULATORIALI PUBBLICHE FUORI ORARIO?
- ❖ DATO CHE INTERVENTI PROGRAMMATI DI SCREENING ONCOLOGICI DI EFFICACIA DIMOSTRATA NON VENGONO ANCORA FRUITI DA TUTTA LA POPOLAZIONE BERSAGLIO SONO OPPORTUNE INIZIATIVE COME QUELLA DEI DERMATOLOGI ITALIANI?

Il quesito

L'offerta di visite gratuite aperte a tutti per diagnosticare tempestivamente lesioni cutanee sospette si presenta in una luce favorevole. Se mancano prove scientifiche di una diminuzione della letalità di melanomi e carcinomi cutanei diagnosticati in anticipo, non è escluso che un effetto del genere ci sia. Dunque l'impresa sembra rispondere alle condizioni spesso menzionate in bioetica: rispetta l'autonomia dei fruitori, può forse produrre qualche beneficio, non è discriminante, è gratuita.

Invece sono proprio le iniziative come questa che mostrano come le valutazioni etiche siano infide quando vengono formulate su progetti non sufficientemente articolati, dei quali non vengano messi in luce aspetti che a prima vista possono sfuggire. L'effetto benefico della diagnosi tempestiva è incerto e una diagnosi non è una cosa buona in sé, ma va valutata alla luce degli effetti dei processi nei quali si inserisce. Tra questi, si possono cogliere subito alcuni aspetti negativi: la visita

dermatologica avvierà un certo numero di persone ad accertamenti successivi, creando in esse apprensioni. Queste in parte verranno dissipate, ma intanto quelle persone avranno sofferto inutilmente. Le sofferenze potrebbero risultare accettabili se il risultato complessivo dello screening fosse buono in misura rilevante, cioè permettesse di giovare a chi risulta malato, prolungandone la vita in buone condizioni od ottenendo addirittura un numero significativo di guarigioni. Ma, come abbiamo detto, non si dispone di dati del genere. Altre persone saranno rassicurate fin dalla prima visita, ma non è detto che esse non siano malate: se lo fossero, sarebbero indotte a non fare ulteriori accertamenti e cure e, se fosse vero che la diagnosi e la terapia precoci sono importanti, ne avrebbero un danno gravissimo. In parte le apprensioni suscitate dalla prima visita sommaria risulteranno giustificate e le persone affette da un tumore saranno curate. Ma neppure questo può essere considerato un be-

ne in sé, perché anticipare la cura di un tumore asintomatico non è un atto benefico se prolunga il periodo di sofferenza psicologica e forse anche fisica delle persone, senza allungarne la vita.

Il fatto che l'offerta dei dermatologi sia aperta a tutti non è necessariamente un tratto positivo: significa semplicemente che non si è fatto nessun tentativo di individuare una popolazione a rischio, alla quale indirizzare gli esami. Prestare un servizio a chi presumibilmente non ne ha bisogno è uno spreco, rappresenta cioè un uso inefficace di risorse che potrebbero essere utilizzate meglio in altre iniziative sanitarie, per esempio in screening la cui efficacia sia stata già dimostrata. Del resto il fatto che le visite siano gratuite non annulla il costo dell'operazione, costituito dalla mobilitazione degli ambulatori nei giorni festivi e dalla coda di esami successivi ai quali la prima visita sommaria avvierà molte persone.

Carlo Augusto Viano
Università di Torino

Il commento

140

EP anno 24 (3) maggio-giugno 2000

Da: Viano CA. Un neo due volte sospetto. *Epidemiol Prev* 2000;24:102. Riproduzione autorizzata.

Caratteristiche di un buon test

- 1) Economico
- 2) Semplice
- 3) Limitatamente scomodo o spiacevole
- 4) Affidabile*
- 5) Accurato**

*Fattori che condizionano l'affidabilità di un test

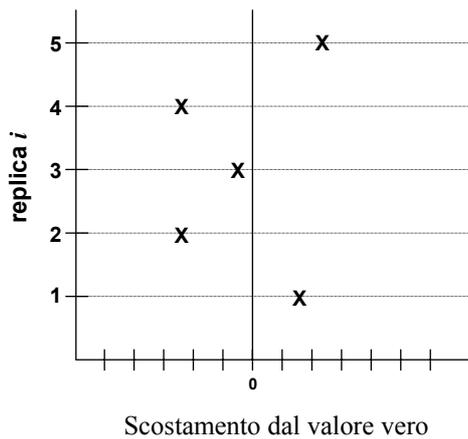
- 1) Variabilità biologica nel tempo
- 2) Errore dello strumento di misurazione
- 3) Variabilità del giudizio dell'osservatore
- 4) Variabilità nei giudizi di più osservatori

**Accuratezza:

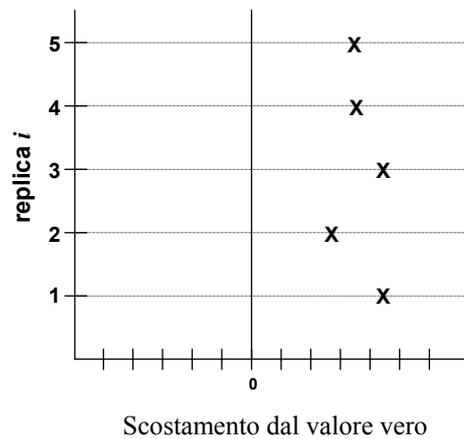
Il grado con cui una misurazione, o una stima basata su misurazioni, rappresenta il vero valore della caratteristica.

Differenza tra il risultato di due test e il valore vero in cinque verifiche indipendenti

Test 1: più accurato che affidabile



Test 2: più affidabile che accurato



PROBLEMA

Stima dell'accuratezza di una misurazione quando la caratteristica è dicotomica.

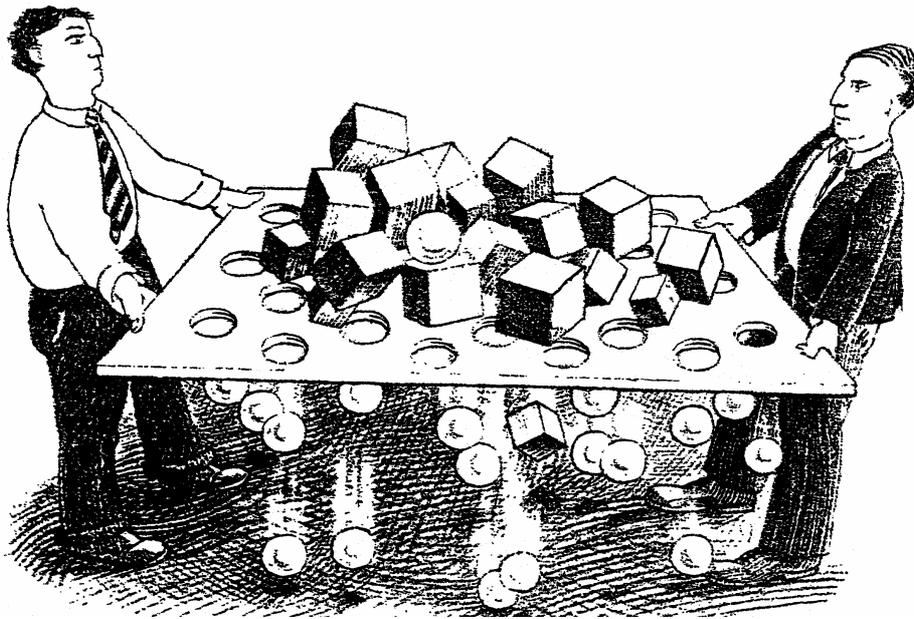
SOLUZIONE

Si misura la caratteristica in soggetti per i quali si dispone del *vero valore* della caratteristica e si calcola la SENSIBILITÀ e SPECIFICITÀ della misurazione.

Nota:

Per *vero valore* si intende una stima ottenuta con un sistema di misurazione più accurato, ovvero tradizionale.

Il *vero valore* è solitamente chiamato *golden standard*.



Disegno di Peter Brown. Riproduzione autorizzata.

Nota:

I due personaggi sono in grado di distinguere i cubi dalle sfere, più che fare uno screening stanno misurando l'accuratezza del sistema.

Esempi di misurazioni dicotomiche sono i test diagnostici

		Golden Standard (vero valore)		
		+	-	
Test in esame	+	veri positivi a	falsi positivi b	a+b
	-	falsi negativi c	veri negativi d	c+d
		a+c	b+d	a+b+c+d

Sensibilità = $\frac{a}{a+c}$

Specificità = $\frac{d}{b+d}$

Sensibilità di un test diagnostico: probabilità di un risultato positivo nei “malati”.
 Specificità di un test diagnostico: probabilità di un risultato negativo nei “non malati”.

Sensibilità e specificità sono caratteristiche intrinseche del test diagnostico, ma del tutto irrilevanti per il soggetto sottoposto al test,

egli infatti non vuole sapere:

qual è la probabilità che ha un malato di essere positivo al test (sensibilità)

ma vuole sapere:

qual è la probabilità di un positivo al test di essere ammalato (valore predittivo positivo)

		Golden Standard (vero valore)		
		+	-	
Test in esame	+	veri positivi a	falsi positivi b	a+b
	-	falsi negativi c	veri negativi d	c+d
		a+c	b+d	a+b+c+d

Valore predittivo positivo = $\frac{a}{a+b}$

Valore predittivo negativo = $\frac{d}{c+d}$

Prevalenza della condizione = $\frac{a+c}{a+b+c+d}$

Il valore predittivo dipende oltre che da sensibilità e specificità anche dalla prevalenza della malattia nella popolazione sottoposta al test

Esempio 1:

Sensibilità =82%
 Specificità =91%
 Prevalenza della malattia 10%

		Golden Standard (vero valore)		
		+	-	
Test in esame	+	41	40	81
	-	9	410	419
		50	450	500

Valore predittivo + = 51%
 Valore predittivo - = 98%

Esempio 2:

Sensibilità =82%
 Specificità =91%
 Prevalenza della malattia 60%

		Golden Standard (vero valore)		
		+	-	
Test in esame	+	246	18	264
	-	54	182	236
		300	200	500

Valore predittivo + = 93%
 Valore predittivo - = 77%

Diverso valore predittivo di uno screening effettuato in popolazioni con alta e bassa prevalenza di glaucoma

		Glaucoma		
		Presente	Assente	
Nuovo test	+	215	16	231
	-	15	114	129
		230	130	360

$$\text{Prevalenza} = \frac{230}{360} \times 100 = 64\%$$

$$\text{Sensibilità} = (215/230) \times 100 = 93\%$$

$$\text{Specificità} = (114/130) \times 100 = 87\%$$

$$\text{Valore predittivo positivo} = (215/231) \times 100 = 93\%$$

$$\text{Valore predittivo negativo} = (114/129) \times 100 = 88\%$$

		Glaucoma		
		Presente	Assente	
Nuovo test	+	215	248	463
	-	15	1822	1837
		230	2070	2300

$$\text{Prevalenza} = \frac{230}{2300} \times 100 = 10\%$$

$$\text{Sensibilità} = (215/230) \times 100 = 93\%$$

$$\text{Specificità} = (1822/2070) \times 100 = 88\%$$

$$\text{Valore predittivo positivo} = (215/463) \times 100 = 46\%$$

$$\text{Valore predittivo negativo} = (1822/1837) \times 100 = 99\%$$

Esercizio:

Un medico generico esegue uno screening su 2000 bambini per identificare casi di una affezione cardiaca la cui vera prevalenza è 10%; la sua capacità di diagnosi ha una sensibilità e una specificità pari all'80%.

I bambini con test positivo vengono rivisitati da un cardiologo la cui sensibilità diagnostica è pari al 90%.

a) Quanti bambini vengono considerati cardiopatici dal medico generico?

1. 1.800
2. 1.600
3. 520
4. 360
5. 160

b) Quanti bambini vengono considerati cardiopatici dal cardiologo?

1. 520
2. 468
3. 360
4. 180
5. 144

c) Qual è la sensibilità dell'intero programma di screening (cioè il risultato combinato dei due test)?

1. 90%
2. 80%
3. 72%
4. 60%
5. Altro

Bibliografia

- Centrewall BS. Exposure to television as a risk factor for violence. *Am J Epidemiol* 1989;129:643-52.
- Colton T. *Statistica in Medicina*. Padova: Piccin; 1979.
- Freedman R, Berelson B. The Human Population. *Scientific American* 1974;Sept:31-9.
- Gordis L. *Epidemiology*. Philadelphia: W.B. Saunders; 1996.
- Kahn HA, Sempos CT. *Statistical methods in epidemiology*. New York: Oxford University Press; 1989.
- Kleinbaum DG, Kupper LL, Morgenstein H. *Epidemiologic research*. New York: Van Nostrand Reinhold; 1982.
- Last JM. *A Dictionary of Epidemiology*. New York: Oxford University Press. 1988.
- Lilienfeld AM, Lilienfeld DE. *Foundations of Epidemiology*. New York: Oxford University Press; 1980.
- MacMahon B, Pugh TF. 1970. *Epidemiology: Principles and methods*. Boston: Little, Brown.
- Matthew EB, Dietzman DE, Madden DL, Newman SJ, Sever JL, Nagler B, Bouton SM, Rostafinski M. A major epidemic of infectious hepatitis in an institution for the mentally retarded. *Am J Epidemiol* 1973;98(3):199-215.
- Mausner JS, Kramer S. (Ed.). *Mausner & Bahn. Epidemiology - an introductory text*. Philadelphia: WB Saunders Company; 1985.
- Paci E, Dini S, Buiatti E, Seniori Costantini A, Lenzi S, Zappa M. Malignant mesothelioma in non-asbestos textile workers in Florence. *Am J Ind Med* 1987;11(3):249-54.
- Rothman KJ. *Modern Epidemiology*. Boston: Little, Brown & Company; 1986.
- Tang AM, Graham NM, Kirby AJ, McCall LD, Willett WC, Saah AJ. Dietary micronutrient intake and risk of progression to acquired immunodeficiency syndrome (AIDS) in human immunodeficiency virus type 1 (HIV-1)-infected homosexual men. *Am J Epidemiol*. 1993;138(11):937-51.
- Vineis P. *Modelli di rischio. Epidemiologia e causalità*. Torino: Giulio Einaudi; 1990.
- Williams MA, Mittendorf R, Stubblefield PG, Lieberman E, Schoenbaum SC, Monson RR. Cigarettes, coffee, and preterm premature rupture of the membranes. *Am J Epidemiol* 1992;135:895-903.

*Stampato da Tipografia Facciotti srl
Vicolo Pian Due Torri 74, 00146 Roma*

*Supplemento 1, al n. 3 vol. 19 (2006)
del Notiziario dell'Istituto Superiore di Sanità
ISSN 0394-9303*