

Statistica ed Epidemiologia

L'epidemiologia è il metodo più semplice e diretto di studiare le cause di malattia negli esseri umani e molti contributi importanti sono stati portati da studi che non hanno richiesto niente più della capacità di contare, di pensare in modo logico e di avere un'idea originale.

Hennekens

- Perché questa persona è sana e quest'altra malata ?
- Perché questa persona è guarita da questa malattia in questo momento ?
- La salute e la malattia non sono distribuite a caso tra gli esseri umani
- Per ogni malattia esistono fattori causali e fattori protettivi

Chi ha peccato, lui o i suoi genitori, perché nascesse cieco ?

dal Vangelo secondo Giovanni

Vi sono nei corpi molti atomi che sono per noi salutari: ma se ne debbono d'altra parte diffondere molti che apportano le malattie e danno la morte quando corrompono il cielo e l'aria diventa mefitica. E vengono tutte le epidemie o dal cielo, come le nubi, o se ne sollevano su dalla terra, quando battuta da piogge e soli non tempestivi si trova ad esser fradicia.

da Lucrezio: De Rerum Natura

Là c'è purtroppo la vera cagione ... quella fatale congiunzione di Saturno con Giove.

da Manzoni: Promessi Sposi

Intorno a Oristano, dove il terreno è paludoso, vengono su dei vapori pesanti e l'aria si corrompe ed è veramente malefica per chi non vi è assuefatto.

Sigismondo Arquer 1500

La causa primitiva o esterna di detta malattia furono alcuni uomini e cose che entrarono a Saragozza provenendo da alcuni luoghi della Francia nei quali si moriva di peste.

Giovanni Tommaso Porcell 1564

Descartes Discorso sul metodo

- 1. La prima regola era di non accettare mai nulla per vero, senza conoscerlo evidentemente come tale: cioè di evitare scrupolosamente la precipitazione e la prevenzione; e di non comprendere nei miei giudizi niente più di quanto si fosse presentato alla mia ragione tanto chiaramente e distintamente da non lasciarmi nessuna occasione di dubitarne.*
- 2. La seconda, di dividere ogni problema preso in esame in tante parti quanto fosse possibile e richiesto per risolverlo più agevolmente.*
- 3. La terza, di condurre ordinatamente i miei pensieri cominciando dalle cose più semplici e più facili a conoscersi, per salire a poco a poco, come per gradi, sino alla conoscenza delle più complesse; supponendo altresì un ordine tra quelle che non si precedono naturalmente l'un l'altra.*
- 4. E l'ultima, di fare in tutti i casi enumerazioni tanto perfette e rassegne tanto complete, da essere sicuro di non omettere nulla.*

Londra 1854: Snow

A Londra l'acqua era fornita in concorrenza con una casa servita da una compagnia e quella accanto da un'altra. La scelta non era in relazione con lo stato socio-economico. La mortalità per colera era alta nelle zone servite dalle compagnie Lambeth e Southwark, che attingevano entrambe l'acqua dal Tamigi, ma in punti diversi. Tra il 1849 e il 1854 la Lambeth spostò il luogo di prelievo in una zona meno inquinata. Dall'analisi dei dati (agosto 1853 - gennaio 1854)

Snow dimostrò che il colera era trasmesso con l'acqua.

servite da	Morti	Abitanti	Tasso
Southwark	192 /	167654	114/10000
Entrambe	182 /	301149	60/10000
Lambeth	0 /	14632	0/10000
servite da	Morti	Case	Tasso
Southwark	1263/	40046	315/10000
Lambeth	98 /	26107	37/10000
Altro	1422/	256423	59/10000

Vienna 1846: Semmelweiss da S.B. Nuland I figli d Ippocrate (1992)

I All'Università di Vienna esistevano due divisioni di ostetricia, identiche sotto ogni aspetto, in ognuna delle quali nascevano circa 3500 neonati l'anno. Tra loro c'era un'unica differenza: nella I Divisione tutti i parti erano seguiti da medici e studenti, nella II Divisione da ostetriche. Nella I Divisione morivano ogni anno di febbre puerperale 600/800 donne, mentre nella II Divisione le morti erano 60.

II Mentre la febbre puerperale imperversava con violenza dentro gli ospedali, fuori, nella città di Vienna, non si riscontrava tale epidemia. Le madri che partorivano in casa avevano un rischio di mortalità molto basso, e tra le madri che partorivano senza assistenza non si riscontrava alcuna mortalità.

III La mortalità non era in alcun modo collegata alle variazioni climatiche.

IV Maggiore era la lesione della cervice e dell'utero, maggiori le probabilità che la madre sviluppasse la febbre puerperale.

V La chiusura del reparto arrestava sempre la mortalità per un periodo.

Da tali osservazioni, Semmelweis comprese che la causa non poteva essere attribuita né a epidemie né a miasmi, ma doveva trovarsi in un qualche fattore inspiegato esistente nella I Divisione e non nella II Divisione.

Kolletschka, un docente di medicina legale, subì un'accidentale lacerazione durante un'autopsia e morì rapidamente per infezione.

Rimuginai sul caso, finché, all'improvviso, non mi attraversò la mente un pensiero; d'un tratto mi fu chiaro che la febbre puerperale e la morte di Kolletschka erano la stessa cosa dal punto di vista patologico perché entrambe presentavano gli stessi cambiamenti anatomici. Se, dunque, nel caso di Kolletschka... i cambiamenti nella sepsi... derivavano dall'inoculazione di particelle cadaveriche, allora la febbre puerperale doveva aver origine dalla stessa fonte. ...La fonte di trasmissione delle particelle cadaveriche doveva essere costituita dalle mani degli studenti e dei medici.

Ogni studente e maestro sezionavano parecchi cadaveri al giorno, ci si lavava in modo approssimativo. Semmelweis aveva trovato la spiegazione: la trasmissione di *particelle cadaveriche invisibili* era la causa della febbre puerperale.

La febbre puerperale non era una malattia specifica come il vaiolo. Il vaiolo poteva essere provocato solo da un altro caso di vaiolo ... la febbre puerperale poteva derivare dal contatto con qualunque fonte di pus.

Semmelweis comprese appieno che dovevano esser presenti tre fattori: la fonte di materia putrefatta, un mezzo per trasportare fisicamente tale materia dalla fonte alla vittima che doveva venirvi in stretto contatto, e una superficie ferita, come la parete spoglia di un utero post partum o di un dito ferito.

La febbre puerperale è una malattia trasmissibile ma non contagiosa. Un ascesso dentale o un cancro uterino infetto non possono causare il vaiolo. Ma il pus di entrambi può causare la febbre puerperale.

Semmelweis prescrisse la semplice misura di lavarsi la mani in una soluzione di cloro finché la pelle non fosse scivolosa e l'odore di cadavere sparito.

Nel 1848, il primo anno intero di questa profilassi, la I Divisione ebbe un tasso di mortalità per febbre puerperale dell'1,2% e la II Divisione dell'1,3, pienamente verificabili e senza precedenti. Nell'aprile del 1847, l'ultimo mese in cui non era prescritto il lavaggio delle mani, la percentuale di puerpere uccise per mano dei medici era stata del 18,3%. Erano stata dunque individuate la causa e la prevenzione della febbre puerperale. Il principio era semplice, logico e coerente con tutte le osservazioni fatte da Semmelweis.

La salute e la malattia non sono distribuite a caso tra gli esseri umani

Nel Dizionario di Epidemiologia Last definisce l'epidemiologia nei termini seguenti

L'epidemiologia può essere definita come studio della distribuzione e dei determinanti di condizioni ed eventi legati alla salute in popolazioni specifiche e applicazione di tale studio al controllo dei problemi di salute.

Il termine "studio" include sorveglianza, osservazione, verifica delle ipotesi, ricerca analitica ed esperimento.

Il termine "distribuzione" fa riferimento all'analisi delle persone affette in relazione a tempo, luogo e condizioni.

I "determinanti" sono tutti i fattori fisici, biologici, sociali, culturali e comportamentali che influenzano la salute.

Il termine "condizioni ed eventi legati alla salute" include malattie, cause di morte, comportamenti come uso di tabacco, reazioni ad azioni di prevenzione, disponibilità e uso di servizi sanitari.

"Popolazioni specifiche" sono quelle con caratteristiche identificabili e numeri definiti con precisione.

Il termine "applicazione ... al controllo ..." rende esplicito lo scopo dell'epidemiologia: promuovere, proteggere, ripristinare la salute.

Obiettivi dell'epidemiologia

1. valutare lo stato di salute di una popolazione;
2. valutare se gli esposti a un determinante abbiano una probabilità di malattia differente rispetto ai non esposti;
3. valutare l'efficacia di un intervento preventivo o terapeutico.

Approccio

1. molecolare: biologia, biochimica
2. tissutale / organico: anatomia patologica
3. al singolo: clinica
4. alla popolazione: epidemiologia

L'associazione tra malattia e determinante (terapia) può essere reale o dovuta a:

- 1) caso
- 2) bias (errore sistematico di raccolta / interpretazione dati)
- 3) confondimento
(effetto di fattori interagenti)

- L'associazione può essere modificata da interazione.

Coloro che calcolano molto vinceranno; coloro che calcolano poco non vinceranno. Sun –Tzu L'arte della guerra

Statistica descrittiva Presentazione, organizzazione e sintesi dei dati.

Statistica inferenziale Generalizzazione dei risultati ottenuti dai dati raccolti in un campione e confronto tra gruppi.

Studi epidemiologici

<u>Assegnazione non controllata</u>		<u>Assegnazione</u>
<u>Studi Osservazionali</u>		<u>controllata</u>
<u>Epidemiologia descrittiva</u> Mira a descrivere le caratteristiche generali della distribuzione della malattia in relazione a spazio, tempo e caratteristiche individuali. <u>Studio trasversale:</u> si misurano contemporaneamente la causa e l'effetto <u>Studio ecologico:</u> l'analisi non è effettuata a livello individuale, ma di gruppo.	<u>Epidemiologia analitica</u> Mira a definire la relazione tra malattia e determinanti. <u>Studio di coorte:</u> i gruppi sono formati in base alla causa e si misura l'incidenza di malattia a distanza di tempo. <u>Studio caso-controllo:</u> i gruppi sono formati in base all'effetto e si confronta la frequenza con cui la causa ipotizzata ha agito nel passato.	<u>Studio sperimentale</u> Mira a definire la relazione tra terapia e malattia. <u>Individuale</u> assegnazione di ogni soggetto <u>Di comunità:</u> assegnazione di gruppi interi

DATI STATISTICI

In statistica una popolazione (di uomini, animali, oggetti) è costituita da un insieme di unità statistiche individuali. Per ciascuna unità statistica è determinato il valore assunto da una o più variabili che corrispondono alle caratteristiche dell'unità statistica.

I valori possono essere:

- **QUALITATIVI**

1. qualitativi non ordinabili dicotomici / binari (sano, malato)
2. qualitativi nominali non ordinabili multicotomici (razza, stato civile)
3. qualitativi ordinabili (migliorato, invariato, peggiorato / caldo, tiepido, freddo)

- **QUANTITATIVI**

4. quantitativi discreti (numero di stanze, di conviventi)
5. quantitativi continui (altezza, peso, età, pressione arteriosa)

DATI QUALITATIVI

E' cruciale il criterio di classificazione che deve possedere i requisiti di:

1. coerenza interna (criterio logico di classificazione);
2. completezza (ogni caratteristica deve trovar posto nella classificazione);
3. mutua esclusione (una caratteristica deve trovare posto in una sola classe).

PROPORZIONE (frequenza relativa)

Nella proporzione il numeratore è compreso nel denominatore.

$$\frac{\text{Numero di sieropositivi HIV}}{\text{Popolazione totale}} \quad \frac{\text{Parte}}{\text{Tutto}}$$

La proporzione è un numero adimensionale compreso tra 0 e 1.

Quando la **proporzione** si riferisce ai casi di malattia in una popolazione si parla di **prevalenza** (puntuale o di periodo), se si considerano i casi esistenti di malattia, o di **incidenza cumulativa**, se si considerano i nuovi casi di malattia insorti in una coorte di individui in un periodo di tempo.

Prevalenza * puntuale / di periodo

Casi esistenti di malattia

Popolazione totale

Incidenza cumulativa * in tempo dato

Nuovi casi di malattia

Popolazione sana a inizio studio (coorte)

TASSO

Il numeratore (numero eventi) deriva dal denominatore (popolazione a rischio) e indica il numero di individui che, con l'evento, lasciano la popolazione a rischio. E' riferita a un definito periodo di tempo e a una popolazione di dimensioni 10^n .

Numero di morti * 10^n / Popolazione a rischio

Quando non è possibile definire la popolazione a rischio e questa è prossima alla popolazione totale, la popolazione totale può essere usata come denominatore.

Nuovi casi malattia / Popolazione totale

Quando il tasso è riferito ai nuovi casi in una popolazione a rischio si parla di **densità d'incidenza**. * in tempo dato

Nuovi casi di malattia / Totale tempi-persona di osservazione

Prevalenza = Incidenza * Durata

La prevalenza di una malattia è funzione di incidenza e durata. La diffusione di una terapia che non dia luogo a guarigione ma prolunghi la sopravvivenza può portare ad un aumento della prevalenza. Un azione di prevenzione che riduca l'incidenza riduce sempre anche la prevalenza.

MORTALITA'

$$\text{Tasso di mortalità} = \frac{\text{Morti}}{\text{Popolazione}} * 10^n$$

$$\text{Tasso di mortalità per causa} = \frac{\text{Morti per una singola causa}}{\text{Popolazione}} * 10^n$$

$$\text{Mortalità proporzionale} = \frac{\text{Morti per una singola causa}}{\text{Morti per tutte le cause}}$$

$$\text{Letalità} = \frac{\text{Morti per una malattia}}{\text{Malati per la stessa malattia}}$$

La mortalità è funzione di incidenza e letalità. L'influenza ha alta incidenza e bassa letalità. Il tetano ha bassa incidenza e alta letalità. L'influenza ha mortalità più alta del tetano.

Tutte le misure di morbosità e mortalità possono essere calcolate per gruppi specifici della popolazione, in particolare per sesso e per classe di età.

DETERMINANTI DI SALUTE E MALATTIA

Tempo

Variazioni a lungo termine Le cause di variazione della frequenza di malattie e decessi sono diverse. Una malattia infettiva può diminuire per cause biologiche (carenza di serbatoi), ambientali (condizioni ostacolanti la sopravvivenza dell'agente), o essere il risultato di un intervento sanitario.

Il trend temporale dei decessi dipende sia dalla frequenza con cui ci si ammala per quella malattia sia dalla sopravvivenza.

Variazioni stagionali: le malattie infettive a trasmissione aerea sono più frequenti in inverno, quelle a trasmissione oro-fecale in estate.

Fluttuazioni periodiche: le epidemie di morbillo o varicella insorgono ogni 2-5 anni, quando si forma una sufficiente popolazione suscettibile.

Variazioni a breve (giorno settimana, ora): gli incidenti stradali sono più frequenti sabato e domenica e in determinate ore del giorno.

Spazio

La frequenza di una malattia può essere influenzata dal luogo in cui si verifica, in termini di area delimitata da confini naturali o artificiali. Confronto a livello sovranazionale, nazionale, regionale, cittadino. Una malattia con caratteristica distribuzione geografica in Italia è il cancro dello stomaco, con una fascia di

elevata mortalità sull'Appennino tosco-emiliano e un gradiente di valori sempre più bassi verso il sud. La frequenza della malattia può originare da caratteristiche ambientali o climatiche. In alcuni casi le malattie sono tanto vincolate all'ambiente da verificarsi solo in certe zone (parassitosi delle aree tropicali).

Gruppo etnico / Patrimonio genetico

Le differenze geografiche possono essere legate a fattori genetici, ambientali, culturali, comportamentali:

- l'isolamento della popolazione è perdurato per un numero di generazioni sufficiente a far sviluppare una condizione genetica particolare;
- la situazione ambientale, la presenza di vettori, di sali minerali nell'acqua, di cibo o abitudini particolari sono associati alla malattia.

In realtà l'interazione tra determinanti genetici e ambientali è complessa, legata a molteplici fattori di diversa natura.

Un esempio di un fattore genetico comune ad un gruppo etnico è l'elevata frequenza di talassemia in Sardegna.

Studio dei migranti

Il ruolo di tutte queste variabili può essere saggiato con lo studio dei migranti, con cui lo stesso patrimonio genetico viene posto in situazioni ambientali

differenti. Ad esempio uno studio di questo tipo ha indagato la frequenza di malattie cardiovascolari e ictus cerebrale tra gli uomini Giapponesi residenti in Giappone, alle Hawaii e in California. I Giapponesi alle Hawaii avevano una mortalità intermedia tra quelli residenti in California (mortalità maggiore) e quelli in Giappone (mortalità minore). Poiché i primi e i secondi erano emigranti lo studio indicò che l'ambiente esterno (stile di vita, dieta, etc.) aveva un ruolo eziologico importante nella patologia cardiovascolare.

Età

La frequenza con cui ci si ammala o si muore per qualsiasi causa ha quasi sempre una relazione con l'età. In tutti i paesi sviluppati la mortalità è relativamente elevata nella prima infanzia, diminuisce quindi marcatamente raggiungendo un minimo tra le classi di età 5-14 anni, per poi riaumentare prima gradatamente e , dopo i 40-50 anni, in maniera più rapida.

Anche l'incidenza della malattia è correlata con l'età. All'aumentare dell'età, aumenta il numero dei portatori di patologie croniche.

Alcune malattie infettive hanno una distribuzione per età legata da una parte all'immunità passiva trasmessa con l'allattamento materno e dall'altra all'immunità permanente conseguente all'infezione. In alcuni casi il legame con l'età è legato ad altre variabili come l'occupazione o lo stile di vita.

Sesso

Sia la mortalità che la morbosità si manifestano in modo diversificato nei due sessi. Le donne vivono più a lungo degli uomini, ma si ammalano più spesso. Le differenze fra uomini e donne può essere spiegata come risultato non di uno, bensì di una serie di fattori genetici, ambientali e sociali. Per alcune malattie l'associazione con il sesso è legata a fattori genetici (ad esempio G6PD carenza, emofilia); per altre, a fattori esterni che interessano diversamente uomini e donne (patologie riconducibili al consumo di alcool o di tabacco); per molte altre cause coesistono fattori genetici, ormonali, ambientali e socioeconomici legati al sesso.

Condizioni socio-culturali

Le condizioni socio-culturali influenzano la salute degli individui. Seri problemi insorgono nell'adottare criteri uniformi di definizione delle condizioni sociali: è necessario scegliere un indicatore di condizione sociale, come reddito, grado di istruzione o tipo di occupazione, oppure una combinazione di tali fattori. In Gran Bretagna i dati sanitari sono analizzati da molti anni per classi definite da 5 livelli lavorativi. E' risultato che sia la mortalità infantile che la mortalità generale dipendono dalla classe sociale. Ciò può essere spiegato da una diversa esposizione a determinanti e/o a differenze nell'uso delle strutture sanitarie.

Occupazione

Poiché le persone passano gran parte del loro tempo nei luoghi di lavoro, le condizioni lavorative hanno un ruolo importante sullo stato di salute. E' noto come alcune patologie siano caratteristiche del tipo di occupazione svolta (silicosi per i lavoratori di miniere e di industrie ceramiche, tumore del polmone per i lavoratori dell'asbesto, tumore della vescica per coloro che lavorano con alcune aniline aromatiche). E' quindi importante tenere conto del lavoro svolto dai e delle possibili esposizioni nocive.

D'altro canto le differenze eventualmente osservate tra persone che svolgono attività diverse possono essere state provocate dalla selezione iniziale dei soggetti per un certo *tipo* di lavoro, mostrando un apparente "effetto protettivo" dell'occupazione.

Determinanti ambientali

Sostanze chimiche: aria, acqua, suolo, alimenti

Radiazioni ionizzanti, ultraviolette

Onde elettromagnetiche, rumore.

Stile di vita

Alimentazione, alcol, tabacco, farmaci, droghe illegali, attività fisica, igiene personale, abitudini sessuali.

Determinanti legati ai servizi sociali e sanitari

Tecnologici Nuove tecniche di diagnosi o terapia possono modificare lo stato di salute.

Organizzativi Un sistema capillare di consultori garantisce la diagnosi precoce di alcune patologie. Programmi di promozione della salute possono potenziare la conoscenza dei cittadini e la cooperazione tra i servizi. Un sistema di emergenza garantendo il rapido trattamento può determinare una completa guarigione senza reliquati.

Socioeconomici Il modello di sistema sanitario ha profonde implicazioni sullo stato di salute: legislazione ambientale; copertura parziale o totale della popolazione; presenza di tickets.

DATI QUANTITATIVI

Biglie che cadono attraverso spuntoni a triangolo, sono deviate con probabilità pari a destra o a sinistra, distribuendosi secondo una distribuzione a campana. L'evento più probabile è che la biglia cada al centro; allontanandosi dal centro il numero di biglie diminuisce; un ugual numero cade a destra o a sinistra della linea centrale; la somma delle distanze dal centro delle biglie del lato sinistro è uguale a quella del lato destro. Questa distribuzione (normale o **gaussiana**) caratterizza i fenomeni naturali.

Per descrivere la distribuzione di dati quantitativi si definiscono sia il centro dove i dati tendono a raccogliersi (misure di **tendenza centrale**), che la dispersione tra i dati (misure di **dispersione**).

MISURE DI TENDENZA CENTRALE

La misura di tendenza centrale più usata è la *media aritmetica*, calcolata sommando i singoli valori e dividendo per il numero di misurazioni.

$$\text{MEDIA} = \frac{\text{SOMMA DEI SINGOLI VALORI}}{\text{NUMERO COMPONENTI POPOLAZIONE}} = \frac{\sum X}{N}$$

In realtà tale misura è valida solo quando la distribuzione dei dati è simmetrica.

MODA

La *moda* di una serie di valori è l'osservazione che più frequente.

La migliore misura di posizione centrale per una serie di dati spesso dipende da come sono distribuiti i valori. Se sono simmetrici e *unimodali* (esiste un solo picco), media, mediana e moda sono approssimativamente uguali. Se la distribuzione di valori è simmetrica ma *bimodale*, la media e la mediana sono approssimativamente uguali, ma il valore potrebbe trovarsi tra i due picchi ed essere quindi una misurazione che si verifica difficilmente. Una distribuzione bimodale spesso indica che la popolazione da cui sono selezionati i valori consiste di due distinti gruppi che differiscono per la caratteristica misurata; in questo caso, è preferibile riportare due mode piuttosto che la media o la mediana.

MEDIANA - ASIMMETRIA

Una misura di tendenza centrale meno sensibile al valore di ciascuna misurazione è la *mediana*, che può essere utilizzata per dati ordinali, discreti e continui. La *mediana* è il 50° percentile di una serie di misurazioni.

Se una lista di osservazioni è classificata in ordine crescente, la metà dei valori sarà maggiore o uguale alla mediana, mentre l'altra metà sarà minore o uguale.

Se n è pari, la mediana è la media dei due valori centrali.

Uno stiramento orizzontale di una distribuzione di frequenza (così che un'estremità sia più allungata e presenti più osservazioni dell'altra), viene definito asimmetria. Se una distribuzione è asimmetrica, la media (più influenzata dai valori estremi) si muove verso l'estremità lunga più della mediana. Se media e mediana non coincidono è plausibile un'asimmetria.

MISURE DI DISPERSIONE

Valori equidistanti dalla media dovrebbero contribuire in uguale misura all'indice di variabilità (*dispersione* dei dati intorno alla media) indipendentemente dal fatto di essere superiori o inferiori alla media stessa, anche se nel primo caso lo scostamento è positivo e nel secondo caso è negativo. Elevando un numero al quadrato si ottiene un numero positivo e pertanto descriveremo la variabilità di una popolazione intorno alla media calcolando la *media dei quadrati degli scostamenti dalla media*, definita come **varianza**.

$$\text{VARIANZA} = \frac{\text{SOMMA (SINGOLI VALORI-MEDIA)}^2}{\text{NUMERO COMPONENTI POPOLAZIONE}} = \frac{\sum (X-M)^2}{N}$$

Se la varianza è calcolata per un campione la formula è

$$\text{VARIANZA} = \frac{\text{SOMMA (SINGOLI VALORI-MEDIA)}^2}{\text{NUMERO COMPONENTI POPOLAZIONE} - 1} = \frac{\sum (X-M)^2}{N - 1}$$

Si usa più comunemente la *deviazione standard*, radice quadrata della varianza, (*radice quadrata della media dei quadrati degli scostamenti dalla media*).

$$\text{DEVIAZIONE STANDARD} = \sqrt{\text{VARIANZA}}$$

Nella maggior parte dei fenomeni biologici, o quando i valori sono generati in modo casuale, la curva assume un aspetto simmetrico a campana. La distribuzione è in tal caso definita “**normale**” o “**gaussiana**” e l’area sottesa alla curva rappresenta tutte le osservazioni della distribuzione. In tal caso:

Media, mediana e moda coincidono.

Il 68% delle osservazioni rientra

tra una deviazione standard sopra e una sotto la media ($\alpha=0.32$).

Il 95% delle osservazioni rientra

tra 1.96 deviazioni standard sopra e 1.96 sotto la media ($\alpha=0.05$).

CURTOSI

La curtosi è uno stiramento verticale della distribuzione di frequenza. Una distribuzione curtotica può apparire più appuntita o più appiattita della distribuzione normale a forma di campana.

ERRORE CASUALE IN UN SINGOLO CAMPIONE – ERRORE STANDARD – INTERVALLO DI CONFIDENZA

Media, mediana, tasso, proporzione sono soggetti al caso. E' improbabile che la stima ottenuta da un campione corrisponda esattamente al valore della popolazione. Un secondo campione quasi certamente darà una stima diversa.

Se da una popolazione di 200 individui si estraggono casualmente 200 campioni, si hanno 200 medie campionarie, che si distribuiscono secondo una distribuzione normale con media prossima alla media della popolazione e deviazione standard che rappresenta la precisione della stima ed è definita **errore standard**.

Il valore vero della popolazione è incluso in un intervallo di confidenza $(1-\alpha)\%$ costruito sommando alla stima campionaria il prodotto tra errore standard e valore critico $N_{\alpha/2} = N_{1-\alpha/2}$ (**stima - $N_{\alpha/2}$ ES e stima + $N_{1-\alpha/2}$ ES**) ed escludendo un $\alpha/2$ % ai due estremi della distribuzione. L'ampiezza degli intervalli di confidenza dipende dall'errore standard e dal livello di fiducia prescelto. Per costruire gli intervalli di confidenza al 95% (la scelta più comune); escludendo il 2,5% della distribuzione ad ogni estremo (coda), si ha $N_{\alpha/2} = N_{1-\alpha/2} = 1.96$.

(----intervallo di confidenza 95% --)

.....**I**.....**I**.....**I**.....

stima - 1.96 es

stima

stima + 1.96 es

Se desideriamo un livello di fiducia maggiore o minore, si possono costruire intervalli diversi come 99% o 90%. Con intervalli più ampi si ha una maggiore garanzia che il valore vero della popolazione sia compreso nell'intervallo.

Gli intervalli di confidenza spostano l'attenzione dalla stima di un singolo valore ad un intervallo di valori valido per la popolazione.

Barisione, Mannheim

Quando leggiamo i risultati di un sondaggio in cui viene chiesto alla gente per chi voterebbe alle elezioni ... dovremmo interpretare l'avvertimento che il "margine d'errore" del sondaggio è del 3% nel modo seguente: che un partito cui venga attribuito, il 25% delle preferenze, potrebbe disporre in realtà del consenso tanto del 22% ($25 - 3$) quanto del 28% ($25 + 3$) degli italiani. La "forchetta" all'interno della quale si situa il valore reale è dunque del 6%.

La deviazione standard misura la variabilità fra le singole osservazioni all'interno del fenomeno che si sta studiando, ed è quindi un indice descrittivo. L'errore standard invece fornisce una misura dell'incertezza propria di una statistica campionaria. La deviazione standard è pertinente quando interessa conoscere la variabilità fra singole osservazioni; l'errore standard è pertinente quando dobbiamo trasferire i risultati di un campione alla popolazione.

Per alcune stime, come nel caso della differenza fra medie campionarie, la distribuzione campionaria appropriata è la distribuzione t . Si adotta lo stesso procedimento, con stima $-t_{\alpha/2}ES$ e stima $+t_{1-\alpha/2}ES$. Per campioni grandi $t_{\alpha/2} = N_{\alpha/2}$. L'errore standard della media è calcolato come ds / \sqrt{n} dove ds è la deviazione standard e n le dimensioni del campione.

L'errore standard della proporzione è calcolato come
dove p è la proporzione e n le dimensioni del campione

$$\sqrt{\frac{p(1-p)}{n}}$$

Per altre stime sono pertinenti distribuzioni campionarie diverse, come la distribuzione di Poisson o la distribuzione binomiale.

Barisione, Mannheimer

I risultati di un sondaggio tendono certo a essere tanto più precisi quanto più grande è il campione, cioè quanto più numerose sono le persone intervistate.

La dimensione del campione condiziona la dimensione dell'errore standard e questo a sua volta influisce sull'ampiezza degli intervalli di confidenza. La riduzione della dimensione del campione porta a una minor precisione della stima e ad un aumento dell'ampiezza degli intervalli di confidenza.

Gli intervalli di confidenza esprimono la precisione con cui una statistica campionaria rappresenta l'intera popolazione dei valori di quella stessa statistica. Essi non sono in alcun modo utili per descrivere la variabilità dovuta al disegno o al modo di conduzione dello studio. Essi non possono fornire un controllo per errori che non riguardano il campionamento, ma distorsioni nel disegno, nella conduzione e nell'analisi di uno studio.

ERRORE CASUALE NEL CONFRONTO TRA GRUPPI

IPOTESI NULLA E IPOTESI ALTERNATIVA

Errore I tipo	α	Affermare che è vero	ed è falso
Errore I tipo	α	La probabilità di sbagliare è pari al 5%	$p=0.05$
Errore II tipo	β	Affermare che è falso	ed è vero
Errore II tipo	β	per $b \neq 0.20$, la potenza del test è 0.80	

Obiettivo del confronto è stabilire se la differenza tra i gruppi osservati è **statisticamente significativa**, cioè se è possibile escludere, con ragionevole sicurezza, che sia dovuta al caso.

Il complemento a 1 dell'errore β è definito come potenza del test. Pertanto per $b \neq 0.20$, la potenza del test è 0.80

La potenza è la probabilità di concludere che c'è una differenza quando di fatto la differenza c'è davvero.

L'errore standard della differenza tra due medie è $\sqrt{\frac{ds_1^2 \dots ds_2^2}{n_1 \dots n_2}}$

L'intervallo di confidenza della differenza è compreso, a livello di fiducia 95%, tra

$$\text{Media1-Media2} - 1,96 \sqrt{\frac{ds_1^2 \dots ds_2^2}{n_1 \dots n_2}} \quad \text{Media1-Media2} + 1,96 \sqrt{\frac{ds_1^2 \dots ds_2^2}{n_1 \dots n_2}}$$

Se l'intervallo non comprende **0** la differenza è significativa con un margine di fiducia del 95% ($p < 0.05$).

In alternativa è possibile utilizzare un test di significatività **t di Student**

	Media ₁ -Media ₂
t -	-----
	$\sqrt{\frac{ds_1^2 \dots ds_2^2}{n_1 \dots n_2}}$

L'errore standard della differenza tra due proporzioni è

$$\sqrt{\frac{p_1(1-p_1) + p_2(1-p_2)}{n_1 + n_2}}$$

L'intervallo di confidenza è

$$p_1 - p_2 - 1,96 \sqrt{\frac{p_1(1-p_1) + p_2(1-p_2)}{n_1 + n_2}} \quad \text{e} \quad p_1 - p_2 + 1,96 \sqrt{\frac{p_1(1-p_1) + p_2(1-p_2)}{n_1 + n_2}}$$

Se l'intervallo non comprende il valore **zero** la differenza è significativa con un margine di fiducia del 95% ($p < 0.05$).

In alternativa è possibile utilizzare un test di significatività

	$p_1 - p_2$
$Z =$	$\frac{p_1 - p_2}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$

Valori di significatività associati ai valori di z

Intervalli di confidenza e significatività statistica

Esiste uno stretto legame fra intervalli di confidenza e test di ipotesi. Se si calcolano gli intervalli di confidenza, il risultato del test di ipotesi può essere dedotto rispetto ad uno specifico livello di significatività statistica. La differenza pari a zero fra le medie rappresenta il valore corrispondente all'ipotesi nulla.

Questo sta ad indicare che, se si applica a questi dati il t-test, si otterrà una differenza significativa al 5% fra le medie campionarie ($p < 0.05$). Se il valore zero fosse stato compreso negli intervalli di confidenza al 95%, ciò avrebbe indicato un risultato non significativo al livello del 5%.

Gli intervalli di confidenza forniscono quindi un range di possibilità circa il valore vero della popolazione invece di un'arbitraria dicotomia basata unicamente sulla significatività statistica. Bisogna sempre tenere conto che piccole differenze, di scarso interesse reale, possono risultare significative solo perché il campione è ampio, mentre effetti importanti dal punto di vista clinico possono risultare non significativi solo perché il numero dei soggetti studiati era piccolo.

Dati appaiati

E' essenziale distinguere il caso di osservazioni non appaiate, dove il confronto è tra dati relativi a due gruppi differenti, da osservazioni appaiate, quelle in cui il confronto è tra due misurazioni fatte sugli stessi soggetti in differenti circostanze (come ad esempio prima e dopo il trattamento). Se si sono utilizzati dati non appaiati, dovrebbe essere usato il t test per campioni indipendenti, mentre nel caso di dati appaiati si dovrebbe usare il t test per campioni dipendenti.

Il test U di Mann-Whitney per dati non appaiati è sostituito, per dati appaiati, dal test di Wilcoxon e il test del χ^2 , è sostituito dal test di McNemar.

Analogamente, il metodo per calcolare gli intervalli di confidenza per dati appaiati differisce da quello usato per osservazioni non appaiate

Per dati quantitativi parametrici, invece di confrontare la media delle prime osservazioni con quella delle seconda, si valuta se la media delle variazioni tra prima e seconda osservazione è significativamente diversa da zero.

Test multipli

Se nell'analisi di uno studio eseguiamo molti test, comparando molti sottogruppi o considerando molte variabili, è lecito attendersi un numero di risultati spuri per il semplice aumento della probabilità di ottenere dei risultati falsi positivi.

Chiaramente, più test vengono eseguiti più grande sarà la probabilità di trovare alcuni risultati significativi; non va tuttavia dimenticato però che parallelamente aumenterà anche il numero “atteso” di falsi positivi. Un modo per prevenire il rischio di aumentare eccessivamente la probabilità di risultati falsamente positivi è quello di fissare un valore di p più piccolo quale soglia di significatività statistica (metodo di Bonferroni).

Analisi della varianza - ANOVA

Quando si confrontano le medie di più di due gruppi non è possibile calcolare l'intervallo di confidenza per la differenza tra le medie o il t di Student.

Il test F è un “test t super” che consente di comparare simultaneamente più di due medie. L'ipotesi nulla è che tutti i gruppi derivino dalla stessa popolazione e che la differenza tra le medie sia casuale.

Fisher sviluppò due misure della varianza delle osservazioni, ipotizzando che, se tutti i gruppi provenivano dalla stessa popolazione, le due diverse stime dovessero dare risultati simili.

Una misura è chiamata **varianza tra gruppi** e si basa sulla variazione tra le medie. L'altra è definita **varianza entro i gruppi** e si basa sulla variazione all'interno di ciascun gruppo.

La **varianza tra i gruppi** ne esprime la diversità e quindi, se alta, rende più probabile che la differenza tra i gruppi non sia casuale; **la varianza entro i gruppi** esprime la variabilità della popolazione e quindi, se alta, rende più probabile che la differenza sia casuale, legata alla forte dispersione dei dati.

$$F = \frac{\text{VARIANZA TRA I GRUPPI}}{\text{VARIANZA ENTRO I GRUPPI}} = \frac{N_{\text{gruppi}} * \text{VARIANZA tra MEDIE}}{\text{MEDIA tra VARIANZE}}$$

	N	MEDIA		VARIANZA
GRUPPO 1	26	11,5	$(11,5+10,1+9,1) / 3$	1,3
GRUPPO 2	26	10,1	=	2,1
GRUPPO 3	26	9,1	10,2	2,4
VARIANZA		TRA I GRUPPI		ENTRO I GRUPPI
		N*VARIANZA MEDIE		MEDIA VARIANZE
		$26[(11,5-10,2)^2+(10,1-10,2)^2+(9,1-10,2)^2]$		$(1,3+2,1+2,4)/3 =$
		$= 26 * 1,44 = 37,44$		3,95
F= 37,44 / 3,95 =9,48				

Se F è vicino a **1**, le due stime della varianza sono simili e l'ipotesi nulla che tutte le medie provengono dalla stessa popolazione non viene rigettata. Se il rapporto è molto maggiore di **1**, vi deve essere qualche differenza tra i gruppi, che ha spinto le medie lontano e l'ipotesi nulla non può essere rigettata.

GRADI LIBERTA'	TRA I GRUPPI	ENTRO I GRUPPI
	GRUPPI - 1: 3-1=2	GRUPPI * (N-1):
F = 9,48		3*(26-1)=75
a = 0.05	VALORE CRITICO	4.9

Specificato un livello α , si cerca il suo valore p in una tabella di distribuzione F per determinare se il risultato è significativo.

Se i risultati non sono significativi, l'ipotesi nulla deve essere accettata. Se lo sono, si devono fare passi ulteriori per determinare quali differenze nelle medie siano differenze "reali".

I diversi campioni della variabile dipendente devono essere campioni casuali indipendenti di popolazioni distribuite normalmente con varianze circa uguali.

Scelta del test di significatività

Prima variabile	Seconda variabile	Test
Continua Gaussiana	Dicotomica	t di Student / I. C. Diff. Medie
Continua Gaussiana	Dicotomica appaiata	t appaiato
Continua Gaussiana	Non ordinabile	ANOVA
Continua Gaussiana	Continua Gaussiana	Correlazione Regressione
Continua non Gaussiana / Ordinabile	Dicotomica	Mann Withney / I. C. Diff. Mediane
Continua non Gaussiana / Ordinabile	Dicotomica appaiata	Wilcoxon
Continua non Gaussiana / Ordinabile	Non ordinabile	Kruskal Wallis
Continua non Gaussiana / Ordinabile	Ordinabile	Correlazione ranghi
Continua non Gaussiana / Ordinabile	Continua	Correlazione ranghi
Dicotomica	Dicotomica	c^2 / Test Z / I. C. Diff. Prop.
Dicotomica	Dicotomica appaiata	c^2 McNemar
Dicotomica	Non ordinabile	c^2
Non ordinabile	Non ordinabile	c^2

IMPOSTAZIONE DELLO STUDIO

1. Definizione del fenomeno e della popolazione da analizzare
2. Scelta e costruzione del campione
3. Scelta e predisposizione dello strumento di rilevazione
4. Test di verifica della funzionalità dello strumento di rilevazione
5. Rilevazione dei dati
6. Codifica, registrazione e pulizia dei dati
7. Analisi statistica e interpretazione dei risultati

Definizione del fenomeno da analizzare

Per analizzare lo stato di salute e i suoi determinanti è necessario individuare degli indicatori, che esprimano quantitativamente o qualitativamente l'intensità che il fenomeno stesso assume in una determinata situazione.

La misurazione è un procedimento a seguito del quale si ai fenomeni attributi che permettono di poterli classificare, ordinare, contare.

Il procedimento si attua attraverso:

- 1.identificazione del fenomeno da misurare;
- 2.specificazione di uno o più aspetti osservabili;
- 3.individuazione di una definizione operativa per la misurazione / classificazione in una variabile indicatore;
- 4.valutazione dell'affidabilità e della validità della variabile indicatore.

L'indicatore è una variabile che sintetizza l'informazione relativa a un fenomeno e/o misura il cambiamento in una data situazione.

E' possibile ipotizzare una griglia di criteri per la validità di un indicatore.

- 1.L'indicatore deve essere di semplice raccolta ed elaborazione. Rispondono bene a questo criterio tutti quelli indicatori che derivano da dati raccolti routinariamente (notifiche obbligatorie, dati anagrafici, registri...) e che non richiedono operazioni di calcolo troppo complesse.
- 2.L'indicatore deve essere disponibile in un tempo ragionevole e ad un costo accettabile. L'analisi deve essere tempestiva.
- 3.L'indicatore deve essere disponibile anche a livello di piccola area o di piccolo gruppo e consentire di disporre anche a livello di piccola area di un numero di casi sufficiente per l'analisi.. Un problema può emergere solo in piccoli gruppi.

- 4.L'indicatore deve essere una misura diretta del fenomeno che rappresenta ed essere soggetto il meno possibile all'arbitrarietà dell'osservatore. Diverso è valutare la diffusione dell'epatite B sulla base di elementi diretti come l'indagine sierologica o le notifiche di malattia o valutare la diffusione dell'alcolismo deducendola dalla mortalità per cirrosi epatica, ipotizzando che questa sia dovuta, per il 60-80% al consumo di alcool.
- 5.L'indicatore deve consentire la comparabilità nello spazio e nel tempo. Diverso è il significato dell'essere proprietario di una casa in una grande città o in un piccolo paese. La comparabilità implica anche la necessità di utilizzare classificazioni internazionali o comunque di uso comune. Diverso è il significato di avere la televisione in casa oggi o trenta anni fa.
- 6.L'indicatore deve essere sensibile al cambiamento a breve termine. La mortalità per tumori del polmone non è un indicatore sensibile a breve termine al cambiamento nel consumo di tabacco, in quanto tra consumo e morte passano anni.
- 7.L'indicatore deve essere correlato in modo chiaro alla salute e dotato di una visibilità politica tale da stimolare il cambiamento. In particolare è essenziale che esso sia di semplice comprensione anche al di fuori di un ambito professionale.

Definizione della popolazione da analizzare

Il problema si presenta diversamente a seconda che un'inchiesta miri alla descrizione o all'analisi. Così se si desidera stimare la proporzione di medici che utilizza in un dato momento in Francia un particolare nuovo farmaco la popolazione sarà necessariamente definita come l'insieme dei medici francesi. Supponiamo ora che ci si interessi ai meccanismi di diffusione sociale e che ci si chieda perché certi medici abbiano adottato il nuovo farmaco e altri no. E' evidente che in tal caso sarà del tutto inutile studiare la popolazione dei medici francesi, perché se la relazione che mi aspetto di osservare in virtù della mia ipotesi è verificata in una città X, non c'è nessuna ragione apparente che essa non sia ugualmente ratificata in una città Y.

Tale esempio mostra che la definizione della popolazione obbedisce a vincoli differenti secondo la natura logica dei problemi ai quali si vuole dare risposta. Se si tratta di verificare l'esistenza di certe relazioni tra l'una e l'altra variabile si può scegliere una qualunque popolazione purché si abbia la garanzia che il particolare contesto non rischi di influire sulla relazione in oggetto. Nel caso di problemi di tipo descrittivo la popolazione è al contrario imposta.

da Boudon Les Méthodes en Sociologie

2.Scelta e costruzione del campione

L'idea centrale del campionamento è ottenere informazioni sull'intero (popolazione) studiandone una parte (campione). Il campionamento è basato sul presupposto che la popolazione complessiva possa essere validamente rappresentata da una sua frazione, anche piccola. Il campione è spesso preferibile all'indagine sulla popolazione (censimento): è più economico; richiede meno tempo e consente più cura per raccolta ed elaborazione; spesso la popolazione è troppo ampia o inaccessibile.

Il campione deve essere rappresentativo della popolazione generale, cioè rappresentarne equamente tutte le caratteristiche. Ogni componente della popolazione deve avere uguale probabilità di essere inserito nel campione.

Il campionamento può essere probabilistico o non probabilistico.

Campione probabilistico

Campionamento casuale semplice

Le unità sono selezionate indipendentemente una alla volta fino al raggiungimento della dimensione desiderata del campione. Ciascuna unità di studio della popolazione ha la stessa possibilità di essere inclusa nel campione; la probabilità che una determinata unità sia scelta è uguale a n/N dove n è la dimensione del campione e N è la dimensione della popolazione originaria.

Campionamento casuale sistematico

L'unità di campionamento iniziale è selezionata a caso dalle prime k unità della lista (p. es. 8) dove k è uguale a N/n . Poi, è selezionata ogni k^{esima} unità consecutiva della lista (108, 208, 308...). Il campionamento sistematico richiede la selezione di un unico numero casuale e distribuisce il campione in modo uniforme nell'intera lista della popolazione. Possono verificarsi delle distorsioni (bias) se esiste una sequenza periodica nella lista; ma questa evenienza è rara.

Campionamento stratificato

La popolazione può essere suddivisa (stratificata) in gruppi omogenei (per età, sesso, quartiere, livello di urbanizzazione, area geografica).

Il campionamento casuale semplice non considera elementi che potrebbero influenzare le caratteristiche di interesse e non consente pertanto di escludere la possibilità che un determinato gruppo non sia rappresentato nel campione. Se riteniamo importante includere ogni sottogruppo, possiamo selezionare un *campione casuale stratificato*. Utilizzando questo metodo, dividiamo la popolazione in H distinti gruppi o strati in modo che l' h^{esimo} strato abbia dimensione Nh . Si seleziona, poi, da ciascuno strato un campione casuale semplice di dimensione nh , producendo, così, una frazione di campionamento di nh/Nh per l' h^{esimo} gruppo.

Campionamento a grappolo (cluster)

Il *campionamento a cluster* richiede la selezione di un campione casuale di gruppi e quindi l'inclusione di tutte le unità di studio dei gruppi selezionati. Nel *campionamento bifasico* è prima selezionato un campione casuale di cluster e poi, all'interno di ciascun cluster, un campione casuale di unità di studio. Se i nostri cluster sono le scuole, prima di tutto si seleziona un campione di scuole e poi, da ciascuna delle scuole selezionate, un campione di soggetti.

Barisione, Mannheimer Moltiplicare il numero di campionamenti moltiplica anche la portata degli errori, come nel caso in cui nel primo stadio siano sovrarappresentati comuni ad alto reddito e nel secondo quartieri benestanti. Chi fa parte dello stesso grappolo tende ad assomigliarsi di più dell'insieme, mentre i diversi grappoli tendono a essere particolarmente diversi tra loro. Non è importante disporre di tanti intervistati in ogni grappolo, ma piuttosto contare su una buona varietà di grappoli, in modo da disporre di un panorama sufficientemente vasto della popolazione.

Campione non probabilistico

Il campione è scelto come capita, sulla base del giudizio del ricercatore, costituito da volontari o da chi afferra a un centro (ospedale ambulatorio...).

In generale questo tipo di campionamento è maggiormente soggetto a errori sistematici.

Barisione, Manneheimer I campioni per quote

Il campione per quote assomiglia al campione stratificato nell'esigenza di assicurare un'equa rappresentanza alle diverse categorie sociali, ma la scelta delle persone da intervistare per ogni categoria è affidata alla discrezionalità dell'intervistatore, allo stesso scopo di ottenere un campione rappresentativo della popolazione.

Prepariamo una descrizione dettagliata delle caratteristiche della popolazione totale, ripartendola per tutte quelle che riteniamo importanti in vista del nostro sondaggio. Combinando insieme questi criteri, otteniamo un piano di campionamento per trovare la quantità giusta di persone dotate delle caratteristiche richieste. Se il piano è disegnato ed eseguito correttamente, il risultato sarà quello di un campione che riproduce, per le caratteristiche messe in quota, la composizione della popolazione.

Bias di selezione

Il bias è un errore sistematico che può alterare i risultati dello studio. I criteri di selezione dei diversi gruppi non sono comparabili. I gruppi a confronto sono tratti da popolazioni che differiscono per un determinante legato al determinante oggetto di studio.

Barisione, Mannheim

Come sono state scelte le persone da intervistare?

Questa domanda serve a distinguere i sondaggi scientifici, effettuati su campioni rappresentativi di una popolazione, dagli pseudosondaggi, basati su campioni autoselezionati e perciò incapaci di rappresentare una popolazione più ampia di quella dei rispondenti stessi.

Barisione, Mannheimer

Poniamo di dover effettuare un sondaggio su un campione rappresentativo degli italiani maggiorenni. Se uscissimo per strada e fermassimo un certo numero di persone a casaccio non otterremmo un campione casuale, perché non c'è mai un momento in cui tutti abbiano le stesse possibilità di essere selezionati.

Nel caso in cui andassimo in giro un mercoledì notte alle tre finiremmo per sovrarappresentare metronotte, panettieri, disoccupati nottambuli e universitari tiratardi, a scapito di tutte le altre categorie.

Ma anche nel caso apparentemente più corretto di un'estrazione effettuata a metà pomeriggio, la regola delle uguali probabilità per tutti sarebbe lontana dall'essere soddisfatta, visto che sarebbero esclusi tutti quelli che al pomeriggio sono al lavoro, o non hanno l'abitudine di andare in centro, o a quell'ora fanno la pennichella, o sono troppo anziani per uscire spesso di casa, e così tanti altri gruppi di persone che fanno parte della popolazione italiana ma che in questo modo non potrebbero figurare nel campione.

Barisione, Mannheimer: Pseudo-sondaggi

Durante le trasmissioni televisive spesso gli spettatori sono invitati a telefonare per partecipare a un sondaggio in diretta. Questi strumenti non possono essere considerati dei veri e propri sondaggi.

Il metodo utilizzato pone almeno tre problemi.

- *A rispondere sono soltanto gli spettatori, non rappresentativi della popolazione italiana.*
- *Coloro che rispondono non sono prescelti in base a un piano di campionamento, ma decidono volontariamente di partecipare al sondaggio, perché sospinti da motivazioni particolarmente forti.*
- *Ogni telespettatore è libero di telefonare più volte.*

Un esempio analogo è quello dei sondaggi su internet o quello dei fax inviati per esprimere la propria opinione. Queste testimonianze possono essere significative da un punto di vista qualitativo, nel senso che leggendo i messaggi ci si può fare un'idea del tipo di argomentazioni utilizzate per difendere una posizione e di quali siano le opinioni. Per la “non rappresentatività” del campione, tuttavia, essi possono dirci ben poco sull'aspetto quantitativo.

Barisione, Mannheimer: Fra chi sono state scelte queste persone?

Oltre alla tecnica di campionamento, è fondamentale conoscere l'oggetto del campionamento stesso, cioè la popolazione statistica campionata. Se la popolazione da cui le persone intervistate sono state estratte è, per esempio, quella degli elettori di un partito o quella degli aderenti a una determinata categoria professionale, a queste stesse precise popolazioni devono essere riferiti i risultati.

Tra gli studenti delle scuole medie inferiori la proporzione di positivi per almeno un marker è del 5.7%. Tra gli studenti delle scuole medie superiori la proporzione è del 5.4%. Il dato è strano: ci aspetteremmo un aumento della proporzione di positivi all'aumentare dell'età.

Ci troviamo di fronte ad un bias di selezione in quanto, mentre gli studenti delle scuole medie (scuola dell'obbligo) sono espressione dell'intera popolazione, quelli delle scuole superiori provengono da un ambiente sociale più ristretto.

Pertanto lo studio è viziato dal fatto di confrontare la sieropositività ad età differenti non nella stessa popolazione, ma in popolazioni differenti.

Barisione, Mannheimer: Il problema delle non-risposte

Non tutte le persone che si tenta di contattare possono essere raggiunte, né tutte quelle raggiunte sono disposte a farsi intervistare. I rifiuti di partecipare al sondaggio possono porre un grave problema di rappresentatività del campione. Se non-rispondenti e rispondenti non sono diversi fra loro, il campione ottenuto con le persone intervistate potrà essere considerato come un campione casuale del più vasto campione originariamente previsto.

Talvolta coloro che rifiutano di rispondere presentano certe caratteristiche in misura maggiore degli altri: essere anziani, appartenere a classi più disagiate, avere un livello d'istruzione inferiore... Coloro che non possono essere contattati sono più spesso giovani e di sesso maschile. In questi casi, il campione finale rischia di non rendere correttamente conto della popolazione.

Poiché non sempre è facile capire se le non-risposte abbiano prodotto delle distorsioni, chi conduce il sondaggio cerca sempre di ridurle al minimo, ripetendo varie volte i tentativi di contatto, offrendo incentivi, con una lista di sostituti.

NUMEROSITA' CAMPIONARIA

Barisione, Mannheim: Quante sono queste persone? Un buon sondaggio può essere eseguito anche con mille o meno intervistati ... Non è quindi la taglia del campione a suggerirci l'affidabilità di un sondaggio. Le dimensioni del campione servono piuttosto per conoscere il margine d'errore

Le dimensioni del campione non sono in relazione con la validità della stima, ma con la precisione dei risultati. Un aumento della numerosità non cambia la stima puntuale, ma riduce l'ampiezza degli intervalli di confidenza.

Barisione, Mannheim: Quante sono queste persone? Purtroppo, il numero degli intervistati è pubblicato solo saltuariamente e il margine d'errore praticamente mai. La diffusione di queste informazioni insieme ai risultati costituirebbe un indice di serietà e di correttezza.

In un confronto tra due campioni, gli intervalli di confidenza sono più piccoli quando entrambi i gruppi sono di pari dimensioni, anche se la loro ampiezza aumenta solo lievemente, salvo che uno dei gruppi sia molto piccolo.

Barisione, Mannheimer Quanto dev'essere grande un campione?

Visti gli equivoci che regnano al riguardo, cominciamo dal dire di che cosa non bisogna tener conto: delle dimensioni della popolazione su cui si svolge il sondaggio. Per una delle stupefacenti ma rigorose leggi della statistica, il fatto che a essere sondati siano gli abitanti di una cittadina di provincia, quelli di una regione o l'intera popolazione nazionale non ha pressoché nessuna influenza sulla quantità di persone necessaria per avere un campione rappresentativo. Contrariamente a quello che l'intuito potrebbe suggerire, insomma, un campione di 1000 persone può rappresentare con la stessa fedeltà e con gli stessi limiti la popolazione di Savona, quella della Liguria, quella dell'Italia e quella dell'Europa. Ciò che conta, in generale, è dunque la "taglia" assoluta del campione, e non quella relativa alla popolazione studiata.

È possibile stimare la dimensione del campione più adeguata per ottenere un determinato valore di intervalli di confidenza. La procedura comunemente impiegata per definire le dimensioni di uno studio è basata sui seguenti passi:

- **Scegliere il livello di significatività α , che rappresenta la probabilità di commettere un “errore del I tipo”, cioè di respingere un’ipotesi statistica quando essa è vera.** Il livello di significatività è generalmente posto uguale a 0.05 o 0.01, e può essere “a due code” se si è interessati sia ad incrementi che a decrementi del rischio fra gli esposti, ovvero a “una coda” se si è interessati solo ad apprezzare un incremento (o solo una diminuzione) del rischio.
- **Scegliere la potenza statistica dello studio, che rappresenta la probabilità di evitare un “errore del II tipo” (β) cioè considerare vera un’ipotesi che di fatto è falsa.** In altre parole, una potenza del 90%, corrispondente a $1-\beta=10\%$, implica la possibilità di non rilevare un’effettiva differenza fra esposti e non esposti.
- **Le probabilità di commettere un errore del I tipo e del II tipo sono antagoniste:** se viene scelto molto piccolo, si è tranquilli di non rilevare effetti inesistenti, ma risulta elevato, e in molti casi si sarà portati a negare l’esistenza di effetti realmente presenti. Solo un aumento delle dimensioni numeriche dello studio può ridurre simultaneamente α e β .

Se

$$ES = \sqrt{\frac{p(1-p)}{n}} \quad ES^2 = \frac{p(1-p)}{n} \quad n = \frac{p(1-p)}{ES^2}$$

$p(1-p)$ è massimo quando $p=0.50$

Ponendo $ES = 0,03$

P	0,50	0,40	0,30	0,20	0,10
1-p	0,50	0,60	0,70	0,80	0,90
$p(1-p)$	0,25	0,24	0,21	0,16	0,09
N	278	267	233	178	100

$z = \frac{p_1 - p_2}{\sqrt{\frac{[p_1(1-p_1)]/n_1 + [p_2(1-p_2)]/n_2}{n}}}$	$z^2 = \frac{(p_1 - p_2)^2}{[2p(1-p)] / n}$	$z^2 = \frac{n(p_1 - p_2)^2}{2p(1-p)}$
--	---	--

$z^2 = \frac{(p_1 - p_2)^2}{2p(1-p)}$	$z^2 = \frac{(p_1 - p_2)^2}{2p(1-p)}$	$n = \frac{Z^2 * 2 p(1-p)}{(p_1 - p_2)^2}$	$n = \frac{(Z_{\alpha} + Z_{\beta})^2 * 2 p(1-p)}{(p_1 - p_2)^2}$
---------------------------------------	---------------------------------------	--	---

$(1.96*0.84)^2*2* 0.60*0.40$
$n=-----$
$(0.20)^2$

Se si considerano le medie

$(Z_{\alpha}+Z_{\beta})^2 *2 ds^2$
$n =-----$
$(media_1-media_2)^2$

La ds dovrebbe essere desunta dalla letteratura o da uno studio pilota.

p_1	0,20	0,55	0,60
p_2	0,10	0,45	0,40
$p=(p_1+ p_2)$	0,15	0,50	0,50
$Z_{\alpha}??5\%$	1,96	1,96	1,96
$Z_b 80\%$	0,84	0,84	0,84
$p_1- p_2$	0,10	0,10	0,20
$n_1=n_2$	200	392	98

Lo strumento di rilevazione

I dati possono consistere in opinioni o informazioni già in possesso dell'individuo esaminato o di altri (in tal caso sono rilevati con questionari o interviste) o in dati da ricavare da esami clinici o di laboratorio.

La Legge di Murphy: Una misurazione accurata vale più delle opinioni di mille esperti.

Intervista: il questionario è compilato da un intervistatore. Le risposte sono filtrate dall'intervistatore che costituisce lo strumento di misurazione. E' pertanto decisiva la preparazione dell'intervistatore.

Questionario autosomministrato il questionario è compilato dall'intervistato.

Il ruolo dell'intervistatore si annulla ma sono più facili gli errori di interpretazione e/o compilazione.

Domande chiuse: è indicata una serie di risposte tra le quali l'intervistato sceglie quella/e preferita/e.

Domande aperte: l'intervistato risponde liberamente.

Barisione, Mannheimer

Nelle interviste semistrutturate, l'intervistatore si limita a guidare la persona interrogata nella scaletta degli argomenti da trattare, ma la lascia libera di esprimersi a lungo e di illustrare dettagliatamente il proprio punto di vista. Nelle interviste in profondità (non strutturate), è invece l'intervistato a gestire il filo del discorso a partire da un soggetto dato, in modo che tutti gli argomenti da lui spontaneamente toccati possano servire a far luce sui processi logici attraverso cui si formano le sue opinioni.

Diverso è il metodo del focus group, che consiste in un dibattito organizzato dall'istituto di ricerca fra un gruppo di 8-12 persone. Una volta definito l'oggetto della discussione un ricercatore fa le veci del "moderatore" e un altro prende nota dei ragionamenti e delle riflessioni proposte dai partecipanti. Alla fine, il materiale raccolto sarà analizzato, nel tentativo di individuare le opinioni, le credenze, le rappresentazioni, le immagini più ricorrenti fra quelle espresse dai partecipanti nel corso del dibattito.

Bias di osservazione

Il bias è un errore sistematico che può alterare i risultati dello studio. L'informazione ottenuta dai diversi gruppi non è comparabile (il caso può ricordare meglio del controllo).

La Legge di Murphy: Un alcolizzato è uno che beve più del suo medico

Non avete mai osservato come la gente non risponda mai a quello che chiedete? Rispondono a quello che volete dire o a quello che credono che vogliate dire. Supponete dunque che una signora chieda ad un'altra in una villa di campagna: "Non c'è nessuno con lei?". La signora non risponde: "Sì c'è il maggiordomo tre servi la cameriera..." quantunque la cameriera sia nella stanza e il maggiordomo dietro la sua sedia; ma risponde: "Nessuno è con me sono sola! intendendo dire che nella casa non c'è alcuno di quelli dei quali suppone che l'interrogante chieda. Ma supponete che il medico in un caso d'epidemia domandi: "Chi c'è con lei?" allora la signora si ricorderà subito del maggiordomo della cameriera e di tutti gli altri in casa. Tutta la lingua è usata così: nessuno risponde letteralmente alle vostre domande neppure quando la risposta è veritiera.

Chesterton: Le storie di padre Brown

Barisione, Mannheim: Come sono state formulate le domande?

Il modo in cui una domanda è formulata ha un impatto fenomenale sulle risposte degli intervistati. Se una domanda presenta parole troppo difficili, se appare sbilanciata in un senso, se è contraddittoria, se sembra suggerire la risposta o se utilizza termini carichi di forti significati positivi o negativi, i risultati finiranno per risentirne.

Barisione, Mannheim: In che ordine sono state poste?

Questa è un'altra delle informazioni che è particolarmente raro avere a disposizione. Tuttavia, sarebbe sempre opportuno poter dare uno sguardo all'intero questionario, perché talvolta le percentuali un po' anomale attribuite a una risposta si spiegano con un effetto contesto dipendente dalle domande precedenti. Altre volte, le risposte possono essere influenzate dall'ordine degli items, vale a dire dalle alternative fra cui l'intervistato è invitato a scegliere. Su quest'ultimo punto, una necessaria precauzione da parte degli istituti è quella di far ruotare a ogni intervista l'ordine degli items stessi.

Bias per “falsa testimonianza”

L'intervistato può scegliere di mentire o di non rivelare qualcosa. Ciò si può verificare in alcuni stati degli Stati Uniti, dove una donna che confessa di aver usato cocaina

in gravidanza rischia di vedersi sottratto il bambino.

In tal caso è decisivo l'anonimato.

A volte l'intervistato mente anche a se stesso riducendo i fattori negativi (consumo di sigarette o grassi) e esaltando quelli positivi (attività fisica, consumo di verdure e frutta).

Quando i comportamenti analizzati riguardano l'osservanza di regole, è utile ricorrere a risposte qualitative ordinabili in modo da non far scegliere tra due sole alternative.

Barisione, Mannheimer Ha senso fare un sondaggio su un tema del genere?

Questa domanda di fondo solleva il problema classico delle non-opinioni, ossia delle risposte fornite a casaccio dagli intervistati a proposito dei temi più lontani dai loro interessi personali e dalle loro sfere di competenza. Il rischio di ottenere risultati privi di senso è elevato, proprio perché gli intervistati tendono a fornire una risposta anche quando non dispongano di opinioni ben precise in merito all'argomento della domanda. Un sondaggio affidabile, dunque, è anche un sondaggio che evita di indagare su temi ignoti alla maggior parte della gente, e che rinuncia a porre tutte quelle domande destinate a raccogliere soltanto un'artificiosa serie di non-opinioni.

La Legge di Murphy: Non si può risolvere un problema che non c'è

Bias di misurazione

Riproducibilità = Concordanza tra diverse osservazioni dello stesso fenomeno
casuale

legata al soggetto

sistematica

dell'osservatore

legata all'osservatore

tra osservatori

La Legge di Murphy: I dati senza generalizzazione sono solo pettegolezzi

STUDIO TRASVERSALE

Sono registrate contemporaneamente la presenza e l'assenza della malattia e dell'ipotetico determinante. Consente il confronto tra sani e malati e quello a livello individuale, ma non di valutare la sequenza temporale.

Il test Chi quadro ? 2 confronta, mediante tabelle a doppia entrata, i casi osservati con quelli attesi nel caso in cui tra i gruppi non ci fosse differenza.

Il test è basato sull'approssimazione normale della distribuzione binomiale.

$$c^2 = (\text{Oss-Att})^2 / \text{Att}$$

	Fumatori Y+	Osservati Non fum. Y-		Attesi Fumatori Y+	Non fum. Y-	$\chi^2=S$	$\frac{(O-A)^2}{A}$
Basso titolo di studio X+	a 389	b 172	X+ 561	Y+ ---- X+ T 371	Y- ---- X+ T 190	0,83+	1,63+
Alto titolo di studio X-	c 89	d 72	X- 161	Y+ ---- X- T 107	Y- ---- X- T 54	2,90+	5,69=
	Y+ 478	Y- 244	T 722			=11.1	p=0,00
	Y+/T 66%	Y-/T 34%					
Basso titolo di studio X+	389 p_{x+} 69%	172 31%	X+ 561	p_{x+} - p_{x-} 14%		I.C. 95%	5% 23%
Alto titolo di studio X-	89 p_{x-} 55%	72 45%	X- 161	e.s. 4%			

Per garantire la validità del test in nessuna casella i valori attesi dovrebbero essere meno di 5. Quando i numeri sono piccoli è opportuno utilizzare la correzione di Yates per la continuità.

$$\text{CHI QUADRO} = \sum (|O-A|-0.5)^2 / A$$

Quando uno dei valori attesi è inferiore a 2 non può essere usato il chi quadro e si deve ricorrere al p di Fisher.

$$p \text{ di Fisher} = \frac{(a+b)! (c+d)! (a+c)! (b+d)!}{N! a! b! c! d!}$$

Per dati raccolti prima e dopo un evento si può utilizzare il test di McNemar.

		Dopo l'evento	
		Positivi	Negativi
Prima dell'evento	Positivi	a	b
	Negativi	c	d

$$\text{CHI QUADRO} = (|b-c|-1)^2 / (b+c)$$

Il termine **gradi di libertà** si riferisce al numero di osservazioni che sono libere di variare. Per qualsiasi test i **gradi di libertà** equivalgono alla numerosità campionaria totale meno 1 grado di libertà per ciascuna media calcolata.

REGRESSIONE E CORRELAZIONE

Esiste una correlazione lineare tra circonferenza cranica ed età gestazionale nella popolazione di neonati con basso peso alla nascita. La relazione è espressa dalla *retta di regressione della popolazione* $m_{y|x} = a + bx$, dove y è la circonferenza cranica media dei neonati di età gestazionale x settimane, a è l'intercetta della retta sull'asse y (circonferenza media all'età 0) e b è la sua pendenza, la variazione in $m_{y|x}$ per la variazione di un'unità (una settimana) di x . Se b è positiva, $m_{y|x}$ aumenta all'aumentare di x ; se b è negativa, diminuisce.

Anche se la relazione tra circonferenza cranica media ed età gestazionale è una retta perfetta, la relazione tra i singoli valori non lo è per la naturale variabilità tra le teste dei neonati. Per tener conto di questa dispersione, adottiamo un modello $y = a + bx + e$, dove e è la distanza di un particolare risultato y dalla retta di regressione e determiniamo la retta $y = a + bx$, che meglio si adatta ai dati reali.

La migliore stima della retta di regressione della popolazione è la retta in cui è minima $S e^2 = S (y_i - \hat{y}_i)^2$, la somma dei quadrati delle differenze tra valori stimati dalla retta \hat{y}_i e dati reali y_i . Posto $S (y_i - \hat{y}_i)^2 = S (y_i - a - bx_i)^2$ si calcola $b = S (x_i - m_x) (y_i - m_y) / S (x_i - m_x)^2$ e $a = m_y - b m_x$. Noti a e b per ogni valore x si calcola il corrispondente valore y . Il calcolo è effettuato dai più comuni software.

Una misura di bontà dell'associazione è il **coefficiente di determinazione R^2** , rapporto tra la devianza spiegata dal modello $S(y-m_y) - S(y'-m_y)$ e la devianza totale tra i valori di y $S(y-m_y)$.

La radice quadrata di R^2 è il **coefficiente di correlazione r** , un numero puro compreso tra -1 e +1. Prossimo a 0 indica assenza di correlazione, prossimo a +1 forte correlazione positiva (al crescere di x cresce y) prossimo a -1 forte correlazione negativa (al crescere di x diminuisce y).

MEDIA

m

DEVIANZA RESIDUA

D res

$\Sigma(y' - \mu_y)$

DEVIANZA SPIEGATA

D sp

$\Sigma(y - \mu_y) - \Sigma(y' - \mu_y)$

DEVIANZA TOTALE

D tot

$\Sigma(y - \mu_y)$

DEV. SPIEGATA / DEV.TOTALE

R^2

Coefficiente di regressione

b

Intercetta

a

Coefficiente di correlazione

r

STUDIO ECOLOGICO

L'unità di studio non è l'individuo, ma il gruppo. Non consente il confronto a livello individuale.

Confronta la frequenza della malattia in differenti popolazioni o in tempi diversi nella stessa popolazione in relazione ad ipotetici determinanti, valutando se dove il determinante è presente con alta frequenza, anche la diffusione della malattia è alta (correlazione positiva) o bassa (correlazione negativa).

I coefficienti di determinazione e di correlazione esprimono in termini quantitativi l'associazione evidenziata da uno studio ecologico.

FASI DELLA SPERIMENTAZIONE CLINICA

- O esplorazione dei meccanismi di attività delle molecole e identificazione-esclusione di meccanismi di tossicità
- 1. prima approssimazione a schemi / dosi di trattamento con effetti farmacologici misurabili e assenza di tossicità macroscopicamente rilevabile
- 2. valuta in un piccolo numero di casi (eventualmente anche non comparativamente) la presenza di effetti terapeutici
- 3. valuta il ruolo terapeutico del trattamento con protocolli comparativi che documentano con probabilità sufficiente che il nuovo trattamento produce, rispetto al miglior trattamento in uso, un beneficio superiore
- 4. consolida (comparativamente o meno) le conoscenze sugli aspetti di sicurezza, apprezzabili solo quando l'uso diviene di routine

Sperimentazioni Controllate Randomizzate (Randomized Clinical Trials)

I **RCT** costituiscono il paradigma della sperimentazione.

Di norma si mettono a confronto due o più trattamenti in gruppi osservati nello stesso arco di tempo.

1. Gli obiettivi usati per la valutazione di efficacia del trattamento sperimentale (end-point) sono importanti per il paziente o, se si tratta di end-point surrogati, sono indicatori affidabili di end-point importanti ?

L'ipotesi sperimentale deve essere chiaramente formulata. In molte malattie, soprattutto se di lunga durata, è difficile osservare se un trattamento è in grado di prevenire un evento sfavorevole.

Per la sperimentazione di un farmaco anti-ipertensivo, gli end-point importanti riguardano la prevenzione degli eventi cerebrovascolari, dell'insufficienza cardiaca, dell'insufficienza renale e la sopravvivenza. Diventa perciò necessario impiegare end-point surrogati che siano facili da valutare e misurare come la normalizzazione della pressione arteriosa.

Il problema è che gli end-point surrogati possono risentire del trattamento sperimentale in modo diverso e talora opposto rispetto agli end-point reali.

2. I criteri di diagnosi e di ammissione sono chiaramente esplicitati ?

Nell'adottare i criteri, si deve da far riferimento a procedure dotate di nota affidabilità e deve essere attento a non indurre selezioni rilevanti a causa, ad esempio, di una bassa accettabilità per il paziente (procedure invasive) o di un'insufficiente agibilità delle procedure.

3. La randomizzazione è eseguita in modo da non poter prevedere a quale trattamento sarà assegnato il paziente ?

La caratteristica dei RCT
è la randomizzazione.

I pazienti sono assegnati a caso al trattamento sperimentale o a quello di controllo. Conseguenza diretta della randomizzazione è che ogni ammesso al trial ha uguali probabilità di vedersi assegnato il trattamento sperimentale e quello di controllo; essa consente così di ottenere gruppi comparabili fra i quali sono equamente distribuite caratteristiche prognostiche note e ignote: la sola variabile che può influenzare un differente esito di malattia tra i gruppi a confronto rimane perciò il trattamento assegnato.

4. La distribuzione dei principali fattori prognostici noti è sufficientemente bilanciata tra pazienti che ricevono il trattamento sperimentale e controlli ? E' utile valutare l'opportunità di un campionamento stratificato per fattori prognostici o, in sede di analisi, la presenza di confondimento.

5. E' riportato l'andamento di tutti i pazienti inclusi nel trial ? L'eventuale numero di pazienti perduti all'osservazione è accettabilmente piccolo ?

Tutti i pazienti esaminati dagli osservatori devono essere inclusi nello studio. Numero e caratteristiche degli esclusi devono essere attentamente considerati e precisati, come pure le vie di accesso dei malati all'osservazione e le possibili distorsioni indotte da meccanismi di filtro. Tutti i pazienti devono essere sotto controllo per un periodo minimo prestabilito. Una perdita superiore al 20% diminuisce decisamente il valore dello studio.

6. Il trial è condotto in doppio cieco o, altrimenti, l'end-point è misurabile obiettivamente senza bias che possano alterare i risultati?

L'effetto *placebo* e il “*bias* ottimistico del medico” possono distorcere i risultati, generalmente in senso favorevole al trattamento sperimentale. L'effetto *placebo* (speranza che il farmaco sia utile) può migliorare sintomi (dolore, impotenza funzionale) e, talora, anche parametri apparentemente obiettivi.

Il *bias* ottimistico del medico deriva dalla sua fiducia nel trattamento che sta sperimentando, che lo induce a interpretare ottimisticamente i risultati.

La randomizzazione consente il disegno in doppio cieco (sia i pazienti che i medici ignorano il trattamento assegnato). che è necessario quando gli obiettivi del trattamento sono influenzabili da valutazioni soggettive (dolore, impotenza funzionale). Inoltre può evitare la disparità dovuta al maggior interesse con cui gli sperimentatori seguono i pazienti assegnati al trattamento sperimentale.

Lo studio può anche essere effettuato in singolo cieco (sa solo lo sperimentatore) o in triplo cieco (non sa nemmeno chi analizza i dati).

7. E' esplicitato il livello di precisione e di significatività per considerare verificata, falsificata o suggestiva l'ipotesi sperimentale ?

Le dimensioni del campione da includere nello studio dipendono dalla frequenza dell'end-point nei gruppi sotto osservazione. Le dimensioni devono essere definite e giustificate prima dell'inizio della sperimentazione.

Se la differenza tra i gruppi è elevata anche un campione piccolo potrà evidenziare differenze significative, se è piccola emergerà solo con campioni numerosi. Un campione troppo piccolo sarebbe destinato a non evidenziare differenze significative e quindi potrebbe far scartare una terapia utile.

8. Il trattamento di controllo è il miglior trattamento disponibile ?

Per poter affermare che il nuovo trattamento apporta un rilevante miglioramento alle possibilità terapeutiche è indispensabile confrontarlo non con un qualunque trattamento alternativo, ma con quello considerato più efficace.

9. Sono state messe in atto procedure di salvaguardia relativa alla libertà degli individui e alla possibilità di insorgenza di danni ?

Ogni sperimentazione deve essere approvata da un Comitato Etico. Nessuno può essere incluso in una sperimentazione senza il suo esplicito consenso informato.

La Legge di Murphy: Una medicina è una sostanza che,

iniettata in un ratto, produce un articolo

Sperimentazioni non controllate

Il nuovo trattamento è applicato a tutti i pazienti eleggibili in base ad alcuni criteri (di solito la diagnosi e la necessità di un trattamento); i pazienti sono seguiti nel tempo ed i risultati valutati in base al confronto con l'andamento della malattia non trattata o trattata con terapia standard, che si ritiene ben noto.

Il limite principale delle sperimentazioni non controllate è la mancanza di un gruppo trattato con terapia standard che serva come termine di confronto. Le malattie hanno infatti una variabilità di decorso e possono migliorare spontaneamente in coincidenza col trattamento. Questo espone all'errore di attribuire ad un trattamento il miglioramento dovuto alle fluttuazioni di malattia.

Le sperimentazioni non controllate possono avere un ruolo nella ricerca di trattamenti per malattie così rare da non consentire un *trial* controllato e randomizzato oppure se sussistono tutte le seguenti condizioni:

- malattia a decorso prevedibile ed esito invariabilmente fatale;
- efficacia terapeutica del nuovo trattamento drammaticamente evidente;
- effetti sfavorevoli accettabili se commisurati al beneficio terapeutico;
- assenza di trattamenti alternativi impiegabili per confronto;
- presupposti fisiopatologici e farmacologici sufficientemente forti, tali da

rendere credibili i risultati favorevoli della sperimentazione.

ANALISI DI SOPRAVVIVENZA

E' analizzato il tempo intercorrente tra un momento iniziale, origine dell'osservazione, e un evento terminale (end-point), fine della stessa.

Il **momento iniziale** è definito dal verificarsi di un evento non ambiguo e ben individuabile nel tempo. Per esempio, con riferimento ad una sperimentazione clinica sul tumore della mammella, potenziali momenti iniziali sono: diagnosi, inizio del trattamento, prima ricaduta, inizio del trattamento per la ricaduta, inizio della chemioterapia. L'introduzione nello studio di pazienti in momenti diversi del decorso della malattia espone a rischi di errore di valutazione della sopravvivenza e dei fattori prognostici.

L'**evento terminale** può essere la morte, la scomparsa del dolore, la ricomparsa dell'ovulazione dopo l'interruzione della contraccezione, l'acquisizione dell'immunità nei confronti di una malattia infettiva, la perdita di un livello di qualità di vita... L'evento terminale deve essere preliminarmente definito in termini obiettivi, affidabili e il più possibile legati alla pratica diagnostica corrente.

L'analisi statistica della sopravvivenza presenta due problemi metodologici. Sono presenti **intervalli di sopravvivenza troncati**, relativi a soggetti in cui l'evento terminale non si verifica nell'arco di tempo intercorrente tra il loro ingresso nello studio e la fine dello stesso. Il reclutamento nello studio avviene in tempi successivi, per esempio nell'arco di un triennio, ed un certo numero di esse è ancora vivo dopo sei, sette, otto, ... anni di follow-up.

La **distribuzione dei tempi di sopravvivenza** è fortemente **asimmetrica** rendendo spesso opportuno il ricorso a tecniche non parametriche per ridurre l'influenza di eventuali tempi di sopravvivenza particolarmente lunghi.

L'analisi di sopravvivenza si basa sulle **tavole di sopravvivenza**, costruite col **metodo attuariale** o col **metodo di Kaplan Meier**. Con entrambi i metodi la probabilità cumulativa di sopravvivenza è il prodotto della probabilità di sopravvivere nei successivi periodi di tempo, ma, mentre nella tavola attuariale si considerano periodi prefissati e regolari (mesi, anni, quinquenni), nella tavola di Kaplan Meier mortalità e sopravvivenza sono calcolate per ogni momento in cui si verifica un evento terminale o termina un follow-up.

Se i tempi di sopravvivenza (in ordine crescente di durata) di un gruppo n di soggetti sono $t_1, t_2, t_3 \dots t_n$, la proporzione di soggetti p_{cum} sopravvissuti oltre ciascun tempo di osservazione (t) è stimata come $p = p (1 - m_x)$ con $m_x = m / n$, dove m_x è il tasso di mortalità in ciascun periodo (metodo attuariale) o al verificarsi di ciascun evento (Kaplan Meier), m il numero di eventi nel periodo (metodo attuariale), n il numero di soggetti vivi fino al tempo t e p indica la *moltiplicativa* di tutti gli intervalli di tempo.

Posti come momento iniziale la diagnosi di SM e come evento terminale il raggiungimento del livello di invalidità, si ha

t	n	m	persi	m_x	p_x	p_{cum}
0	1256	2	63	0,00	1,00	1,00
1	1191	3	69	0,00	1,00	1,00
2	1119	16	82	0,01	0,99	0,98
3	1021	20	61	0,02	0,98	0,96
4	940	22	61	0,02	0,98	0,94
5	857	23	57	0,03	0,97	0,92
6	777	11	57	0,01	0,99	0,90
7	709	17	53	0,02	0,98	0,88

$$m_x = m / n \quad 16 / 1119 = 0.01 \quad p_x = 1 - m_x \quad 1 - 0.02 = 0.98$$

$$p_{cum} = p_x * p_{cum} \quad 0.94 * 0.97 = 0.92$$

Test di Mantel-Cox

Per confrontare due tavole Kaplan Meier e definire la significatività della differenza si usa il test **logrank di Mantel-Cox**, analogo al c^2 Mantel Haenszel.

t	n _{prima}	Eventi	m _x	p _x	p _{cum}	n _{prima}	Eventi	m _x	p _x	p _{cum}
0	100	5	0,050	0,950	0,950	250	0	0,000	1,000	1,000
1	93	4	0,043	0,957	0,909	248	3	0,012	0,988	0,988
2	85	5	0,059	0,941	0,856	240	5	0,021	0,979	0,967
3	79	3	0,038	0,962	0,823	231	1	0,004	0,996	0,963
4	73	7	0,096	0,904	0,744	225	10	0,044	0,956	0,920
5	64	5	0,078	0,922	0,686	211	8	0,038	0,962	0,885
6	55	7	0,127	0,873	0,599	197	9	0,046	0,954	0,845
7	45	10	0,222	0,778	0,466	183	12	0,066	0,934	0,790
8	30	8	0,267	0,733	0,342	162	15	0,093	0,907	0,716
9	20	7	0,350	0,650	0,222	142	13	0,092	0,908	0,651

$$A_1 = (O_1 + O_2) / (n_1 + n_2) * n_1$$

$$(5 + 0) / (100 + 250) * 100$$

$$A_2 = (O_1 + O_2) / (n_1 + n_2) * n_2$$

$$(5 + 0) / (100 + 250) * 250$$

	n₁	O₁	A₁	n₂	O₂	A₂
0	100	5	1,43	250	0	3,57
1	93	4	1,91	248	3	5,09
2	85	5	2,62	240	5	7,38
3	79	3	1,02	231	1	2,98
4	73	7	4,16	225	10	12,84
5	64	5	3,03	211	8	9,97
6	55	7	3,49	197	9	12,51
7	45	10	4,34	183	12	17,66
8	30	8	3,59	162	15	19,41
9	20	7	2,47	142	13	17,53

S
S
S
S

p < 0,01

$$C^2_{MC} = \left(\begin{matrix} \text{61} & \text{-28} \end{matrix} \right)^2 / 28 + \left(\begin{matrix} \text{76} & \text{-109} \end{matrix} \right)^2 / 109 = 48,63$$

logrank $C^2_{MC} = (SO_1 - SA_1)^2 / SA_1 + (SO_2 - SA_2)^2 / SA_2$

EPIDEMIOLOGIA ANALITICA

Mira ad evidenziare la presenza di un'associazione tra un determinante e una malattia. Nello studio prospettico l'evento (insorgere della malattia) non si è ancora determinato quando il ricercatore intraprende lo studio.

Nello studio retrospettivo l'evento (insorgere della malattia) si è già determinato quando il ricercatore intraprende lo studio.

STUDIO DI COORTE

Sono confrontati un gruppo di esposti all'ipotetico fattore causale o protettivo e un gruppo di non esposti. I due gruppi sono seguiti per un tempo definito (spesso anni). In ciascun gruppo è valutata l'incidenza della malattia analizzata.

Un gruppo di fumatori (esposti) è posto a confronto con un gruppo di soggetti simili in tutto ai precedenti tranne che per il fatto di non fumare (non esposti). In entrambi i gruppi è valutata l'incidenza di tumori del polmone.

Selezione degli esposti

Si identifica con criteri definiti con precisione un gruppo di persone che subiscono o hanno subito un'esposizione a un ipotetico determinante.

L'accesso allo studio è limitato a chi può andare incontro alla malattia in esame.

Si escludono i soggetti già malati, gli immuni, o gruppi particolari, come, per il tumore del collo dell'utero le donne isterectomizzate.

Un'impropria definizione dei criteri in base ai quali un soggetto è selezionato come membro di una coorte da studiare (criteri di ammissibilità) potrebbe comportare bias di selezione, ad esempio includendo soggetti per cui è impossibile ammalarsi e quindi diminuendo l'incidenza della malattia.

Selezione dei non esposti

La popolazione di confronto in uno studio di coorte ha come primo requisito l'assenza dell'esposizione in oggetto. Essa deve essere comparabile agli esposti per la maggior parte degli altri fattori.

Il confronto può essere interno (quando è possibile distinguere, all'interno della coorte, esposti e non esposti) o esterno: confronto con un'altra coorte di riferimento, del tutto simile a quella in studio, tranne che per l'esposizione.

Popolazione “aperta” o “chiusa”

Nella popolazione “chiusa” l'appartenenza è definita da un evento, ed i soggetti per i quali tale condizione è verificata appartengono per a questa popolazione, indipendentemente da cambiamenti successivi. Se si è interessati agli effetti a lungo termine dell'esposizione ad un agente chimico, verificatasi nel corso di un incidente, si vuole conoscere la sorte di tutti i lavoratori di una fabbrica che sono stati coinvolti in tale incidente. Il criterio di ammissibilità potrebbe essere definito come “essere stato presente in fabbrica il giorno dell'incidente”.

Una popolazione è definita “dinamica” o “aperta” quando è caratterizzata da ricambio (turnover) dei propri membri. Per esempio alcuni fumatori possono smettere di fumare e alcuni non fumatori iniziare a farlo.

Follow-up

La durata dell'indagine deve tenere conto dei meccanismi biologici del fenomeno studiato, ad es. reazioni collaterali acute di una vaccinazione, oppure cancerogenesi chimica. E' essenziale tenere conto del tempo di induzione-latenza della malattia, cioè il tempo compreso fra l'inizio dell'esposizione e l'induzione della malattia, più quello che trascorre fra l'inizio della malattia e la sua successiva diagnosi (latenza).

Raramente può essere evitata la perdita, durante lo studio, di una certa percentuale di soggetti. C'è sicuramente un bias se le perdite fra gli esposti e i non esposti sono state influenzate dalla successiva insorgenza della malattia.

In alcuni studi in cui la coorte è rappresentata da un'intera popolazione (ad es. il Framingham Cohort Study sullo sviluppo di cardiopatie per diversi fattori di rischio) seguita per molto tempo, è possibile che soggetti classificati come non esposti divengano esposti o cambino la categoria di esposizione a cui appartengono. Si può verificare quindi un'errata classificazione dei soggetti e quindi una distorsione nella stima dell'associazione esposizione—malattia.

Accertamento dell'esposizione e delle condizioni di salute

	Malattia presente	Malattia assente	Incidenza
Esposti	a	b	Esposti $a / a+b$
Non Esposti	c	d	Non Esposti $c / c+d$

L'associazione tra determinante e malattia è misurata col **Rischio Relativo**

RR =incidenza esposti / incidenza non esposti $[a/(a+b)]/ [c/(c+d)]$

Può essere calcolato come **incidenza cumulativa** (proporzione)

Nuovi casi di malattia / Popolazione all'inizio dello studio (coorte)

o, con riferimento ai tempi-persona, come densità d'incidenza (tasso)

Nuovi casi di malattia / Totale tempi-persona di osservazione

La distribuzione campionaria del rischio relativo è asimmetrica a destra, mentre la distribuzione del suo logaritmo naturale è approssimativamente normale.

L'errore standard del rischio relativo è pertanto calcolato come $es(\ln RR)$

$=\sqrt{1/a-1/(a+c)+1/b+1/(b+d)}$ con intervallo di confidenza del Rischio Relativo compreso tra $e^{\ln RR-1.96es}$ e $e^{\ln RR+1.96es}$. Se l'intervallo non comprende il valore

1 la differenza è significativa con un margine del 95% ($p < 0.05$). Il valore **0** è il

valore limite di una differenza in quanto se $a = b$ $a - b = 0$. Il valore **1** è il valore

limite di un rapporto in quanto se $a = b$ $a / b = 1$

Dimensione del campione nello studio di coorte

Si stimano le dimensioni del campione impiegando la procedura di Schlesselman, che assume l'equinumerosità dei campioni confrontati.

$$N = \frac{(z_{\alpha} \sqrt{2pq} + z_{\beta} \sqrt{I(1+R - I*(1+R^2))})^2}{(I * (I - R))^2}$$

N è la dimensione del campione, **z_a** il livello di significatività, **z_b** la potenza, **I** l'incidenza nella popolazione di confronto, **R** il più piccolo valore del rischio relativo che si apprezza come significativo, **p = I (R + 1) / 2**, **q = 1-P**.

Lo studio può prendere in considerazione diversi livelli di esposizione.

Mortalità per coronaropatie (maschi 50-69 anni) (Hammond)

Sigarette/ die	RR
0	1
10-20	1.29
20-30	1.89
30-40	2.15
> 40	2.41

Il **Rischio Attribuibile (RA)** è calcolato come

incidenza fra gli esposti - incidenza fra i non esposti = $[a/(a+b)] - [c/(c+d)]$

Incidenza fra gli esposti	Incidenza fra i non esposti	RR	RA
400 / 1000	100 / 1000	4	300 / 1000
40 / 1000	10 / 1000	4	30 / 1000
4 / 1000	1 / 1000	4	3 / 1000

Il **Rischio Relativo** indica quante volte il rischio è superiore negli esposti rispetto ai non esposti (rapporto). Il **Rischio Attribuibile** indica di quanto aumenta l'incidenza in presenza di esposizione. Se il determinante è raramente presente, il **RR** può essere alto, ma il **RA** resta basso (amianto).

E'anche possibile esprimere il confronto come rapporto tra odd degli esposti e odd dei non esposti. Per odd si intende il rapporto tra la probabilità che un evento si verifichi (in questo caso l'incidenza cumulativa) e la probabilità che lo stesso evento non si verifichi (in questo caso la sopravvivenza).

Odds Ratio OR = $\frac{\text{odd fra gli esposti}}{\text{odd fra i non esposti}} = \frac{a / b}{c / d}$

E' possibile dimostrare che, quando l'incidenza è bassa (< 10-15%), OR = RR.

Infatti se $a \sim 0$ e $c \sim 0 \rightarrow a+b \sim b$ e $c+d \sim d$

$$RR = \frac{[a/(a+b)]}{[c/(c+d)]} \sim OR = \frac{a/b}{c/d}$$

						Odd_	Odd_	Bias
I_esp	I_non	S_esp	S_non	RR	OR	Esp	Non	OR
0,010	0,005	0,990	0,995	2,00	2,01	0,01	0,01	0,99
0,023	0,011	0,978	0,989	2,00	2,02	0,02	0,01	0,99
0,051	0,025	0,949	0,975	2,00	2,05	0,05	0,03	0,97
0,076	0,038	0,924	0,962	2,00	2,08	0,08	0,04	0,96
0,171	0,085	0,829	0,915	2,00	2,21	0,21	0,09	0,91
0,256	0,128	0,744	0,872	2,00	2,34	0,34	0,15	0,85
0,577	0,288	0,423	0,712	2,00	3,36	1,36	0,41	0,59

STUDIO CASO CONTROLLO

Sono posti a confronto un gruppo di casi (affetti dalla malattia di interesse) e un gruppo di controlli (senza la malattia di interesse). In ciascuno dei due gruppi è valutata la proporzione di esposti all'ipotetico fattore causale o protettivo.

Un gruppo di soggetti affetti da tumore del polmone (casi) è posto a confronto con un gruppo di soggetti simili in tutto ai precedenti tranne che per il fatto di non essere affetti da tumore del polmone (controlli). In entrambi i gruppi è valutata la proporzione di fumatori.

Criteria di elezione.

Fissare i criteri di eligibilità dei soggetti da inserire nello studio ha lo scopo di limitare l'accesso ai soggetti (casi e controlli) che hanno avuto la possibilità di essere esposti al determinante studiato. In uno studio sulla possibile associazione tra infarto miocardico e consumo recente di contraccettivi orali non saranno ritenute eligibili le donne affette da infarto del miocardio ma in menopausa, ovariectomizzate o con condizioni che controindicano l'uso di contraccettivi orali. Criteri di eligibilità analoghi dovranno essere fissati anche per i controlli.

L'accertamento dell'esposizione può essere effettuato, sia per i casi che per i controlli, mediante questionari, interviste, registri.

Definizione dei casi

La definizione biologica della malattia considerata deve portare a riconoscere un'entità eziologicamente omogenea. In questo senso sarebbe improprio uno studio caso-controllo sui fattori di rischio del cancro dell'utero. Posto che che adenocarcinoma del corpo uterino e carcinoma della cervice sono differenti per molti aspetti, tra i quali i determinanti. Oltre ad una chiara definizione biologica è necessaria anche una definizione operativa della malattia al fine di riconoscere ed includere nello studio solo casi che siano veramente tali.

Selezione dei casi

Gli **ospedali** sono la principale fonte dei casi negli studi retrospettivi. Altre fonti possono essere i **registri di patologia e di morte**, gli **ambulatori** o altri **presidi sanitari extraospedalieri**. Per le patologie che non necessariamente richiedono un ricovero ospedaliero il reclutamento di casi ospedalieri può avere l'effetto di **selezionare per lo studio solo le forme più gravi** distortendo i risultati.

Abitualmente i casi reclutati per uno studio retrospettivo sono **casi incidenti** (di nuova diagnosi), più omogenei di quanto non lo siano quelli con diversa anzianità di malattia. L'utilizzazione esclusiva di casi incidenti può rendere lo studio molto lungo e pertanto molte indagini sono condotte con casi prevalenti, verificando al momento dell'analisi l'effetto della diversa durata di malattia.

Il ricordo delle esposizioni precedenti la malattia è di solito più accurato nei casi incidenti e non distorto dalla consapevolezza di essere malati. Si è inoltre più sicuri che l'esposizione indagata abbia preceduto e non seguito la comparsa della malattia, come in soggetti con malattie croniche che, proprio a causa della malattia, modificano la loro esposizione a determinanti comportamentali quali l'alcool, il fumo, il caffè, i farmaci.

Selezione dei controlli

Il gruppo dei controlli serve a fornire una stima della frequenza di esposizione al determinante studiato in soggetti non affetti dalla malattia.

Molto spesso i controlli sono selezionati nello stesso ospedale da cui sono stati scelti i casi, tra ricoverati per altra patologia (controlli ospedalieri).

I controlli ospedalieri sono più facili da trovare, più consapevoli della propria anamnesi, più disponibili a collaborare. Il limite maggiore è la possibilità che il ricovero sia dovuto a una patologia in qualche modo legata all'esposizione considerata. Se, ad esempio è analizzata l'associazione tra aspirina e infarto del miocardio non sono elegibili come controlli soggetti affetti da ulcera gastrica, di cui l'aspirina è determinante. Dal momento che abitualmente non si sa a quali malattie possa essere associato l'ipotetico determinante, è opportuno includere tra i controlli più malattie, ritenute non associate al determinante studiato.

Meno spesso i controlli sono scelti tra soggetti non ospedalizzati della stessa popolazione da cui provengono i casi (controlli di popolazione, tratti ad esempio dalle liste anagrafiche) o da gruppi prossimi ai casi e pertanto più motivati a collaborare (parenti, amici, vicini di casa, colleghi).

I controlli di popolazione sono sicuramente i più adatti, dal punto di vista teorico, associando la confrontabilità con i casi alla generalizzabilità dei risultati. I controlli di popolazione infine però di solito meno disposti a collaborare dei controlli ospedalieri e le loro risposte sono più povere.

La bontà della stima dipende più dalla **confrontabilità** dei controlli rispetto ai casi che dalla loro rappresentatività rispetto alla popolazione generale, in quanto i casi selezionati non sono sempre rappresentativi della popolazione dei malati. Il problema non si pone per gli studi su base di popolazione, in cui gruppo dei casi è formato da tutti i casi di malattia in una popolazione definita.

	<u>Casi</u>	<u>Controlli</u>
Determinante presente	a	b
Determinante assente	c	d
Odd	a / c	b / d

Nello studio caso-controllo non è possibile calcolare l'incidenza fra esposti e non esposti (casi e controlli sono selezionati sulla base della malattia) e di conseguenza il rischio relativo. E' invece possibile calcolare

$$\text{Odds Ratio OR} = \frac{\text{odd fra i casi } a / c}{\text{odd fra i controlli } b / d}$$

I rapporti a/c e b/d sono gli odds , il rapporto tra la probabilità che un evento si verifichi (probabilità di essere esposto) e la probabilità che lo stesso evento non si verifichi (probabilità di non essere esposto).

L'errore standard è calcolato come $es(\log_e \text{OR}) = \sqrt{1/a + 1/b + 1/c + 1/d}$

L'intervallo di confidenza di OR è compreso tra $e^{\log \text{OR} - 1.96es}$ e $e^{\log \text{OR} + 1.96es}$.

Se l'intervallo non comprende **1** l'associazione è significativa ($p < 0.05$).

Se la selezione dei controlli è effettuata correttamente, i controlli rappresentano la popolazione sana e la proporzione di esposti e non esposti è identica a quella che sarebbe stata rilevata nella popolazione sana stessa ($F_2 = F_4$). In tal caso l'odds ratio calcolato nello studio caso-controllo è uguale a quello che sarebbe calcolato in uno studio di coorte effettuato nella stessa popolazione.

DETERMINANTI DELL'IVG

	CASI IVG	CONTR. PARTI	CASI IVG	CONTR. PARTI
NON CON.	245	67	43%	8%
CONIUG.	330	813	57%	92%
	575	880		
ODD CASI		245 / 330	=	0.74
ODD CONTROLLI		67 / 813	=	0.08
ODDS RATIO		0.74 / 0.08	=	9.01

CONFONDIMENTO E INTERAZIONE

Uno studio epidemiologico può concentrarsi su uno specifico determinante che è ipotizzato influenzare il rischio di una malattia o esplorare gli effetti singoli e congiunti di più variabili. Anche se lo studio è prevalentemente rivolto a un solo determinante, spesso è necessario controllare e aggiustare gli effetti di altri.

CONFONDIMENTO

E' definito confondente un fattore che distorce l'effetto di un ipotetico determinante sotto studio. Perché un fattore possa essere considerato confondente è necessario che sia contemporaneamente associato all'ipotetico determinante e alla malattia studiata.

In tal caso, esaminando separatamente le relazioni tra determinante e malattia per i diversi livelli del confondente si osserva che la relazione tra determinante ed malattia risulta differente da quanto appariva globalmente.

La presenza non controllata di una variabile confondente può produrre sia una accentuazione che una riduzione dell'effetto del fattore di rischio studiato.

In uno studio caso-controllo sull'associazione fra consumo di caffè e tumore del pancreas, se fumo e alcool, associati alla malattia, lo fossero anche al determinante considerato (caffè), essi potrebbero essere i veri responsabili dell'effetto osservato e da noi attribuito al consumo di caffè. Sia l'alcool che il fumo dovrebbero essere considerati variabili confondenti in quanto entrambi fattori di rischio per la malattia e associati alla variabile d'esposizione (chi beve più caffè ha di solito anche l'abitudine di bere più alcoolici e di fumare di più).

Metodi per il controllo del confondimento

Selezionare casi e controlli con restrizione

La restrizione implica l'uso di criteri di ammissibilità che limitano l'accesso dei soggetti allo studio, in modo da selezionare solo soggetti omogenei rispetto alla variabile di confondimento. Lo studio può ad esempio essere circoscritto ai non fumatori o ai forti fumatori.

Appaiare ad ogni caso uno o più controlli omogenei per il confondente

L'appaiamento implica stabilire una corrispondenza su base individuale fra soggetti, in modo che essi mostrino gli stessi valori del confondente.

Può riguardare una o più variabili che avrebbero potuto mostrare una diversa distribuzione fra casi e controlli e quindi rappresentare fattori di confondimento; ogni caso è appaiato ad uno o più controlli, caratterizzati, ad esempio, da identici età, sesso e professione.

Analisi stratificata di casi e controlli: metodo di Mantel e Haenszel

I dati possono essere rappresentati in più tabelle di contingenza 2×2 , *stratificate* in relazione ai livelli di una determinata variabile che si ritiene possa influenzare il risultato. Ciò significa intraprendere un tipo di analisi stratificata, che implica la stima dell'associazione determinante - malattia in strati definiti in base al livello del confondente.

Il metodo di Mantel e Haenszel agisce a livello di elaborazione dei dati.

E' possibile esaminare separatamente l'associazione per ciascun livello del confondente, ma, di norma, è più utile combinare le informazioni tra le tabelle per giungere ad un'unica conclusione complessiva e calcolare, combinando i rischi per ciascun livello della variabile, il rischio complessivo.

Prima di combinare le tabelle di contingenza, dobbiamo verificare che gli effetti del determinante siano uguali per i diversi livelli del confondente.

In caso contrario non si tratta di confondimento, ma di interazione e non è appropriato calcolare la stima globale della forza dell'associazione, come media ponderata dei valori dei singoli livelli.

INTERAZIONE

L'effetto congiunto di due o più determinanti è differente da quanto ci si potrebbe attendere sullo base del loro effetto disgiunto.

Quando gli effetti dell'esposizione al determinante B variano al variare del determinante A e viceversa, c'è interazione tra B e A. Se c'è interazione significativa sarà necessario riportare gli effetti di B separatamente per ogni livello di A e viceversa.

Sinergismo: una situazione nella quale se B è presente l'effetto di A è superiore e viceversa

Antagonismo: una situazione nella quale se B è presente l'effetto di A è inferiore e viceversa

<p>Rischio per malattia coronarica</p> <p>Femmina, Uricemia bassa 1.0</p> <p>Femmina, Uricemia alta 1.7</p> <p>Maschio, Uricemia bassa 1.8</p> <p>Maschio, Uricemia alta 2.4</p> <p>(RR11) ≠ (RR10) (RR01)</p> <p>(2.4) ≠ (1.8) (1.7) 2.4¹ 3.1</p>	<p>Rischio per malattia coronarica</p> <p>Giovane, attività fisica alta 1.0</p> <p>Giovane, attività fisica bassa 4.1</p> <p>Anziano, attività fisica alta 3.4</p> <p>Anziano, attività fisica bassa 13.7</p> <p>(RR11) ≅ (RR10) (RR01)</p> <p>(13.7) ≅ (4.1) (3.4) 13.7 @ 13.9</p>
---	---

Studio caso-controllo: IVG e numero di figli

$$\text{OR} = (107 / 224) / (330 / 429) = 0.62$$

	Casi	Controlli
1 figlio	107	330
Nessun figlio	224	429

$$\text{OR} = (244 / 224) / (121 / 429) = 3.86$$

	Casi	Controlli
2 figli	244	121
Nessun figlio	224	429

Coniugate

$$\text{OR} = (70/32) / (321 / 374) = 2.55$$

	Casi	Controlli
1 figlio	70	321
Nessun figlio	32	374

Non coniugate

$$\text{OR} = (37 / 192) / (9 / 55) = 1.18$$

	Casi	Controlli
1 figlio	37	9
Nessun figlio	192	55

Coniugate

$$\text{OR} = (228 / 32) / (118 / 374) = 22.58$$

	Casi	Controlli
2 figli	228	118
Nessun figlio	32	374

Non coniugate

$$\text{OR} = (16 / 192) / (3 / 55) = 1.53$$

	Casi	Controlli
2 figli	16	3
Nessun figlio	192	55

Fattore di interazione

FIGLI=0	CAS	CON	
NON CONIUG.	192	55	40,8
CONIUG.	32	374	

CONIUG.	CAS	CON	
FIGLI >=1	70	321	2,5
FIGLI = 0	32	374	

	CAS	CON	
NON CON. + FIGLI	37	9	48,0
CONIUG. 0 FIGLI	32	374	

FATTORE DI INTERAZIONE	48,0	-----	= 0,46
		40,8 * 2,5	

STANDARDIZZAZIONE

La struttura per età della popolazione influenza il tasso grezzo di mortalità. Di conseguenza il tasso grezzo di mortalità può risultare per una malattia la cui frequenza cresce al crescere dell'età più alto nella popolazione più anziana anche in presenza di tassi età-specifici più alti nella popolazione più giovane.

Tasso di mortalità per tumori maligni in due distretti

		GHILARZA			QUARTU	
ETA'	N	POP	TASSO	N	POP	TASSO
	206	37016	55.7	245	78056	31.4
0-44	11	21328	5.2	25	59835	4.2
45-59	40	6822	58.6	59	10066	58.6
60-74	68	5876	115.7	103	5960	172.8
75-W	87	2990	291.0	58	2195	264.2

TASSO STANDARDIZZATO TS_0 (metodo diretto)

E' determinato utilizzando come riferimento una popolazione standard.

$$\sum p_i^* m_{ic} / p^*$$

p_i^* = popolazione nella classe di età i della popolazione standard

p^* = popolazione standard,

m_{ic} = tasso specifico per età i e causa c nella popolazione in studio

CALCOLO DEI CASI ATTESI NELLA POPOLAZIONE STANDARD

POPOLAZIONE STANDARD		GHILARZA TASSO	QUARTU TASSO
0-44	21328 + 59835 = 81163	0.001 * 81163 = 42	0.000 * 81163 = 34
45-59	6822 + 10066 = 16888	0.006 * 16888 = 99	0.006 * 16888 = 99
60-74	5876 + 5960 = 11836	0.012 * 11836 = 137	0.017 * 11836 = 205
75-W	2990 + 2195 = 5185	0.029 * 5185 = 151	0.026 * 5185 = 137
	115072	429	474

TASSI STANDARDIZZATI
=37.3

429/115072*10⁴
=41.2

474/115072*10⁴

STANDARDIZZAZIONE INDIRETTA

Tassi specifici della popolazione standard applicati alle popolazioni in esame.

CASI ATTESI E_c $E = \sum p_i m_{ij}^*$

p_i = popolazione della classe di età i della popolazione in studio

m_{ij}^* = tasso di mortalità specifico per età i nella popolazione standard

L'intervallo di confidenza è calcolato sulla base della distribuzione di Poisson dei casi attesi E

$$LCL = (E_c) \left(1 - \frac{9}{(E_c)} - \frac{1.96}{(3(E_c)^{1/2})} \right)^3$$

$$UCL = (E_c+1) \left(1 - \frac{9}{(E_c+1)} + \frac{1.96}{(3(E_c+1)^{1/2})} \right)^3$$

RAPPORTO STANDARDIZZATO DI MORTALITÀ $SMR_c = D_c / E_c$

L'intervallo di confidenza è calcolato come

$$LCL = \exp \left(\ln SMR - 1.96 (1/D^{1/2}) \right) \quad UCL = \exp \left(\ln SMR + 1.96 (1/D^{1/2}) \right)$$

TASSI ETA'-SPECIFICI NELLA POPOLAZIONE STANDARD

N / POP TASSO

0-44 36/81163 = 0.000

45-59 99/16888 = 0.006

60-74 171/11836 = 0.014

75-W 145/ 5185 = 0.028

CALCOLO DEI CASI ATTESI (EXP) NELLE DUE POPOLAZIONI

GHILARZA

QUARTU

TASSO POP EXP

TASSO POP EXP

0-44 0.000 * 21328 = 9 0.000 * 59835 = 27

45-59 0.006 * 6822 = 40 0.006 * 10066 = 59

60-74 0.014 * 5876 = 85 0.014 * 5960 = 86

75-W 0.028 * 2990 = 84 0.028 * 2195 = 61

218

233

RAPPORTI STANDARDIZZATI DI MORTALITA' (RSM)

CASI OSSERVATI / CASI ATTESI

GHILARZA 206/218=0.95 QUARTU 245/233=1.05

RAPPORTO PROPORZIONALE DI MORTALITÀ STANDARDIZZATO

$$D / \sum d_i (d_{ic}^* / d_i^*)$$

d_i = numero di morti per tutte le cause nell'età i nella popolazione in studio

d_{ic}^* = numero di morti per la causa c e nell'età i nella popolazione standard

d_i^* = numero di morti per tutte le cause nell'età i nella popolazione standard

L'intervallo di confidenza è calcolato come $LCL = \exp (\ln SPMR - 1.96 (1/D^{1/2}))$ $UCL = \exp (\ln SPMR + 1.96 (1/D^{1/2}))$

CAUSALITA'

Se caso, bias e confondimento possono essere esclusi è possibile concludere che esiste un'associazione statistica tra la malattia ed il determinante. Ciò non indica necessariamente che il determinante sia fattore causale o fattore protettivo.

Plausibilità biologica E' possibile ipotizzare un processo biologico che conduce dal fattore di esposizione alla malattia: masticare tabacco è associato a carcinoma orale, fumarlo a carcinoma laringeo o bronco-polmonare.

Sequenza temporale Il fattore di esposizione deve agire prima dell'insorgere della malattia. Tra l'esposizione e l'insorgere della malattia deve trascorrere un tempo adeguato (incubazione, latenza).

Forza dell'associazione E' difficile che una differenza molto alta sia legata al caso o a un fattore di confondimento sconosciuto. Tale fattore dovrebbe avere un effetto straordinariamente forte.

Relazione dose-risposta Il rischio aumenta all'aumentare della dose di esposizione: chi fuma 50 sigarette al giorno rischia più di chi ne fuma 5 chi ha fumato per 20 anni più di chi ha fumato per 2.

Consistenza con altri studi Studi diversi condotti in luoghi diversi da ricercatori diversi con metodi diversi danno risultati analoghi.

Non v'è sforzo di fantasia che possa permettere alla mente umana di associare il tabacco da naso con i brillanti, le candele senza candelieri e le rotelle sconnesse.

Io credo di vedervi un nesso.

- *Era furioso contro la Rivoluzione Francese e cercava di rivivere la vita familiare degli ultimi Borboni. Aveva tabacco da naso perché questo costituiva il lusso del secolo diciottesimo; candele di cera perché erano la luce del secolo diciottesimo; i pezzettini di ferro lavorato, che rappresentavano la passione di fabbro di Luigi XVI, e i brillanti, quale simbolo della collana di Maria Antonietta.*
- *Era un ladro. Non aveva alcun candelieri perché usava le candele nella lanterna che portava con sé. Usava il tabacco per lanciarne una manciata sulla faccia di chi volesse arrestarlo. Ma la prova finale è nella coincidenza dei diamanti e delle ruote d'acciaio. Sono i due soli strumenti con i quali sia possibile tagliare un vetro.*
- *Qualcuno l'aveva ingannato con quei diamanti sciolti, dicendo che erano stati trovati nelle caverne del castello. Le rotelle servono per tagliare i diamanti. Egli era costretto a ricerche in maniera primitiva, con l'aiuto di qualche pastore. Il tabacco da fiuto è il grande lusso di simili pastori; con esso li si può corrompere. Essi non avevano candelieri perché tenevano le candele in mano quando esploravano le caverne.*

Dobbiamo scegliere fra queste tre stupide verità ?

No! Ho suggerito la cosa soltanto perché dicevate che non si poteva trovare un nesso tra tabacco da naso, candele e pietre preziose.

Dieci false filosofie possono spiegare l'universo; dieci false ipotesi possono spiegare il mistero. Ma noi vogliamo la vera spiegazione del mistero.

da Chesterton: Le storie di padre Brown

DISTRIBUZIONE DI POISSON

$$n=1000 \quad p=0.015 \quad l = 15$$

Quando il **numero di casi** è molto piccolo (eventi rari), la distribuzione di Poisson è più appropriata per costruire l'intervallo di confidenza per la proporzione. Infatti calcolando nel modo usuale tale intervallo si otterrebbero proporzioni negative. Va notato come con tale distribuzione l'intervallo di confidenza dipenda dal numero di casi e non dalle dimensioni del campione.

La media è **np**, la varianza è **np(1-p)**. Quando **p** è molto piccolo **(1-p)=1** e pertanto la varianza è circa uguale alla media.

$$P_x = \frac{e^{-\lambda} l^x}{x!} \quad \text{con } l = np$$

L'intervallo costruito con la distribuzione di Poisson è asimmetrico con margine superiore più distante di quello inferiore. L'asimmetria tende a ridursi con l'aumentare del numero di eventi.

DISTRIBUZIONE BINOMIALE

In una sequenza di n eventi indipendenti della variabile casuale Y di Bernoulli, ognuno con una probabilità di successo p il numero totale di eventi favorevoli (successi) X è una variabile casuale binomiale.

Una moneta ideale, a ogni lancio, può mostrare testa o croce con probabilità $1/2$. Il numero x degli eventi «croce» su n lanci non è prevedibile a priori: ciò significa che x è una variabile aleatoria, con associata una legge di probabilità p_n , definita distribuzione binomiale. All'aumentare del numero di prove la distribuzione binomiale converge alla distribuzione normale.

In una popolazione di n soggetti indipendenti per la variabile casuale Y di Bernoulli (fumo), ognuno con probabilità di successo p (prevalenza dei fumatori nella popolazione) il numero totale di fumatori (successi) X è una variabile casuale binomiale.

$$P_x = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \quad (! = \text{fattoriale})$$

Risultato Y Soggetto			Probabilità			Numero di eventi
1	2	3				
0	0	0	(1-p)	(1-p)	(1-p)	0
1	0	0	P	(1-p)	(1-p)	1
0	1	0	(1-p)	p	(1-p)	1
0	0	1	(1-p)	(1-p)	p	1
1	1	0	P	p	(1-p)	2
1	0	1	P	(1-p)	p	2
0	1	1	(1-p)	p	p	2
1	1	1	P	p	p	3

$$p(X=0) = (1-p)^3$$

$$p(X=1) = 2 p (1-p)^2$$

$$p(X=2) = 3 p^2 (1-p)$$

$$p(X=3) = p^3$$

La media di X è **np**, la varianza di X è **np(1-p)**

Statistica non parametrica

La mediana di un campione di osservazioni è utilizzata per stimare la mediana «vera» della popolazione.

I metodi per il calcolo degli intervalli di confidenza per medie e proporzioni sono basati sull'assunzione di una distribuzione normale.

Per variabili di tipo continuo non aventi una distribuzione normale è preferibile utilizzare come misura di tendenza centrale la mediana della popolazione o del campione, parametro appropriato anche quando la distribuzione non è gaussiana.

Intervallo di confidenza per un solo campione

L'intervallo di confidenza al 95% per la mediana è definito dai valori r^{esimo} e s^{esimo} della distribuzione ordinata in ranghi con valore crescente. Si calcolano

$$r = n/2 - 1.96 \sqrt{n} / 2 \text{ e}$$

$$s = n/2 + 1.96 \sqrt{n} / 2$$

e si arrotondano poi r ed s al valore intero più vicino.

Non è sempre possibile calcolare gli intervalli di confidenza ad un esatto valore.

Perciò, per intervalli di confidenza al 95%, la scelta dovrebbe orientarsi sul livello di confidenza più basso superiore al 95% (intervallo conservativo).

Un procedimento analogo può essere utilizzato per altri percentili.

Per campioni piccoli è usato un metodo basato sulla distribuzione binomiale.

Topo	Minuti	n	r	s
4	12	6	1	6
7	15	10	2	9
1	21	20	6	15
3	25	30	10	21
10	34	40	14	27
8	38	50	18	33
6	40	60	22	39
9	53	70	27	44
2	57	80	31	50
5	58	90	36	55
11	364	100	40	61
Mediana	38	>100	$n/2 -$	$n/2 +$
			$1.96\sqrt{n}/2$	$1.96\sqrt{n}/2$
I.C. 15-58		11	2	10
95%				

Intervallo di confidenza per la differenza tra due campioni

La differenza tra le mediane di due gruppi è stimata dalla mediana di tutte le possibili $n_p n_g$ differenze tra ogni valore del gruppo p e ogni valore del gruppo g.

$p \backslash g$	14	44	42	27	53	37	29	30	MEDIANA DELLA DIFFERENZA TRA LE MEDIANE
48	-34	-4	-6	-21	5	-11	-19	-18	
41	-27	3	1	-14	12	-4	-12	-11	
31	-17	13	11	-4	22	6	-2	-1	
66	-52	-22	-24	-39	-13	-29	-37	-36	
60	-46	-16	-18	-33	-7	-23	-31	-30	
84	-70	-40	-42	-57	-31	-47	-55	-54	-18
66	-52	-22	-24	-39	-13	-29	-37	-36	
16	-2	28	26	11	37	21	13	14	
45	-31	-1	-3	-18	8	-8	-16	-15	
58	-44	-14	-16	-31	-5	-21	-29	-28	

Per calcolare l'intervallo di confidenza si calcola

$$k = \frac{n_p n_g}{2} - 1.96 \sqrt{\frac{n_p n_g (n_p + n_g + 1)}{12}}$$

Il K^{esimo} valore più piccolo e il K^{esimo} più grande della tabella delle differenze definiscono l'intervallo di confidenza della differenza tra le due mediane.

$p \backslash g$	14	44	42	27	53	37	29	30	RANGO NELLA
48	63	22	26	45	14	29	44	41	DIFFERENZA
41	52	15	16	34	9	22	31	29	TRA LE MEDIANE
31	40	7	10	22	4	13	19	17	
66	75	47	50	68	32	54	66	64	INT.CONF.95%
60	73	37	41	62	27	49	58	57	
84	80	70	71	79	58	74	78	77	Ranghi 34 46
66	75	47	50	68	32	54	66	64	
16	19	2	3	10	1	5	7	6	Corrispondenti
45	58	17	21	41	12	28	37	36	ai valori
58	72	34	37	58	25	45	54	53	-14 -22

Test di Mann Whitney

Se la distribuzione non è normale, per confrontare due gruppi (variabili quantitative) è anche possibile ricorrere al test di Mann Whitney, basato sulla distribuzione dei ranghi.

Si attribuisce un rango a tutte le osservazioni, indipendentemente dal gruppo di appartenenza, assegnando rango 1 all'osservazione più piccola e si valuta la probabilità che la somma dei ranghi ottenuta dai due gruppi sia espressione del caso rispetto a quella che sia legata ad una differenza tra essi.

Placebo		Farmaco	
Diuresi ml/giorno	Rango	Diuresi ml/giorno	Rango
1000	1	1400	6
1380	5	1600	7
1200	3	1180	2
		1220	
	T=9		

Possibili ranghi e somme di ranghi per 3 individui su 7

1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X																		
2	X	X	X	X						X	X	X	X	X	X	X	X																		
3	X			X	X	X	X			X	X	X	X			X	X	X	X	X															
4	X		X			X	X	X		X		X	X	X		X	X	X		X	X	X													
5		X		X		X		X	X		X		X		X	X		X		X	X	X	X												
6			X		X		X	X	X		X		X	X	X		X	X		X	X	X	X	X											
7				X		X		X	X		X		X	X	X		X		X	X	X		X	X	X										
				1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1										
T	6	7	8	9	0	8	9	0	1	0	1	2	2	3	4	9	0	1	2	1	2	3	3	4	5	2	3	4	4	5	6	5	6	7	8

Ci sono 35 possibili ranghi.. Qual è la probabilità di ottenere un valore estremo ?
 La probabilità di ottenere i valori 6 (minimo) e 18 (massimo) è complessivamente una $2/35=0,057$. Di conseguenza $T=9$ ricade nel 95% centrale della distribuzione e non è pertanto significativo.

REGRESSIONE LOGISTICA

La regressione logistica è simile alla regressione lineare, ma è adatta a situazioni dove la variabile dipendente y è dicotomica, mentre la variabile indipendente x può essere sia qualitativa che quantitativa. Se indichiamo con 1 il successo e con 0 l'insuccesso, p , la proporzione di volte in cui la variabile assume il valore 1 , è la media della variabile dicotomica.

La probabilità p può, per definizione, assumere solo valori compresi tra 0 e 1 . Nel modello lineare $p = \alpha + \beta x$, il termine $\alpha + \beta x$ può assumere anche valori al di fuori di tale range ed è pertanto necessario adattare il modello nella forma

$$p = \frac{e^{\alpha + \beta x}}{1 + e^{\alpha + \beta x}} \quad \text{con} \quad \frac{p}{1-p} = \frac{e^{\alpha + \beta x}}{1 - e^{\alpha + \beta x}} = (e^{\alpha + \beta x}) \cdot \frac{p}{1-p} \quad \text{e} \quad \ln \frac{p}{1-p} = \alpha + \beta x$$

Posto che $p / (1-p)$ è l'**odd di un successo**, la regressione logistica equivale ad una regressione lineare in cui il risultato continuo è stato sostituito dal logaritmo dell'odd di un successo. Invece di assumere che la relazione tra p e x sia lineare, assumiamo che lo sia la relazione tra $\ln[p/(1-p)]$ e x .

La funzione logistica non è calcolabile matematicamente, ma deve essere definita per tentativi (iterazioni) successivi fino a minimizzare le distanze tra valori reali e valori stimati.

$$p_1 = \frac{e^{a+bx}}{1+e^{a+bx}}$$

$$1-p_1 = \frac{1}{1+e^{a+bx}}$$

$$p_0 = \frac{e^a}{1+e^a}$$

$$1-p_0 = \frac{1}{1+e^a}$$

I coefficienti di regressione logistica possono essere usati per valutare bx per più variabili indipendenti secondo il modello

$$p = \frac{\exp(a + b_1X_1 + b_2X_2 + b_3X_3)}{1 + \exp(a + b_1X_1 + b_2X_2 + b_3X_3)}$$

I valori β sono diversi da quelli che si ottengono esaminando le variabili singolarmente, in quanto corretti per l'effetto confondente delle altre variabili

Per includere le interazioni il modello diventa

$$p = \frac{\exp(a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_1X_2 + b_5X_1X_3 + b_6X_2X_3)}{1 + \exp(a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_1X_2 + b_5X_1X_3 + b_6X_2X_3)}$$

i valori β corrispondono alla condizione in cui il determinante è presente solo per la variabile corrispondente

Il modello iniziale comprende tutte le variabili e le interazioni. Con procedura **step by step** sono eliminate una ad una le interazioni meno significative. Ogni volta che un'interazione è eliminata gli altri coefficienti β cambiano. Quando tutte le interazioni non significative sono eliminate, si passa a considerare ed eliminare le singole variabili, fatta eccezione per quelle che hanno mantenuto interazioni significative. Il modello finale che ne risulta include tutte le variabili significative o con interazioni significative.

Per valutare quali determinanti siano implicati nell'insorgenza di DBP in una popolazione di neonati di basso peso alla nascita (< 1500 g) consideriamo: età gestazionale, RDS, ventilazione prolungata, fumo materno in gravidanza, PDA, asma, peso alla nascita con le rispettive interazioni.

Variable	B	S.E.	Sig	Variable	B	S.E.	Sig
ETAGES	2,29	0,97	0,02	RDS*FUMO	-0,44	1,03	0,67
RDS	2,25	0,57	0,00	RDS*PDA	7,26	164,27	0,96
VENT	-0,10	3,26	0,98	RDS*P1000	-1,40	1,37	0,31
FUMO	1,04	0,76	0,17	RDS*ASMA	8,09	70,62	0,91
PDA	-7,57	164	0,96	VENT*FUMO	-0,33	1,49	0,83
P1000	1,60	0,93	0,09	VENT*PDA	0,04	1,96	0,98
ASMA	-8,94	70,63	0,90	VENT*P1000	-0,45	1,84	0,81
ETAGES*RDS	-1,45	1,06	0,17	VENT*ASMA	9,53	57992	1,00
ETAGES*VENT	1,54	1,81	0,40	FUMO*PDA	5,49	70,99	0,94
ETAGES*FUMO	0,01	1,06	0,99	FUMO*P1000	0,05	1,07	0,96
ETAGES*PDA	7,66	34,16	0,82	FUMO*ASMA	0,82	1,79	0,65
ETAGES*P1000	-0,29	1,26	0,82	PDA*P1000	1,03	53,76	0,98
ETAGES*ASMA	1,22	1,66	0,46	PDA*ASMA	-1,81	57992	1,00
RDS*VENT	1,85	3,18	0,56	P1000*ASMA	8,36	47,46	0,86

Nessuna delle interazioni risulta statisticamente significativa.

Variable	B	S.E.	Sig	Variable	B	S.E.	Sig
ETAGES	1,34	0,40	0,00	ETAGES	1,53	0,36	0,00
RDS	1,63	0,40	0,00	RDS	1,64	0,40	0,00
VENT	1,97	0,54	0,00	VENT	2,00	0,53	0,00
FUMO	0,68	0,41	0,10	FUMO	0,69	0,40	0,09
PDA	0,90	0,67	0,18	PDA	0,91	0,66	0,17
P1000	0,48	0,45	0,29	Si eliminano asma, peso nascita, PDA.			
ASMA	0,14	0,55	0,79				

Variable	B	S.E.	Sig	OR	I.C. 95%	
ETAGES	1,55	0,35	0,00	4,7	2,4	9,4
RDS	1,71	0,39	0,00	5,5	2,5	11,9
VENT	2,18	0,52	0,00	8,8	3,2	24,5
FUMO	0,66	0,40	0,10	1,9	0,9	4,3

Il fumo materno in gravidanza, anche se non strettamente significativo (p=0.10) è mantenuta nel modello.

Ipotizziamo di voler valutare in un campione di studenti della scuola dell'obbligo, quali determinanti siano implicati nel consumo di latte nelle ultime 24 ore. Sono considerati 3 possibili determinanti: ordine di scuola (elementare = 1, media = 0), sesso (femmina = 1, maschio = 0), consumo di latte nella madre (consumatrice = 1, non consumatrice = 0), con le rispettive interazioni.

Si procede con l'eliminazione step by step delle interazioni non significative.

Variabile	B	S.E.	Sig	Variabile	B	S.E.	Sig
Elementari vs Medie	-0,06	0,28	0,84	Ele vs Medie	-0,09	0,28	0,74
Femmina vs Mas	-0,20	0,29	0,49	Fem vs Mas	-0,41	0,19	0,03
Madre cons. vs non cons.	0,40	0,27	0,14	Madre cons. vs non cons.	0,24	0,21	0,27
Ele by Fem	0,41	0,28	0,14	<i>Ele by Fem</i>	<i>0,43</i>	<i>0,27</i>	<i>0,12</i>
Ele by Madre cons.	0,53	0,30	0,07	Ele by Madre cons.	0,56	0,30	0,06
<i>Fem by Madre cons.</i>	<i>-0,30</i>	<i>0,30</i>	<i>0,32</i>				
Si elimina l'interazione <i>Fem by Madre cons.</i>				Si elimina l'interazione <i>Ele by Fem</i>			

Variabile	B	S.E.	p	OR	I. C. 95%
Ele vs Medie	0,12	0,25	0,62	1,13	0,70 1,83
<i>Fem vs Mas</i>	-0,21	0,14	0,13	0,81	0,62 1,06
Madre cons. vs non cons.	0,26	0,21	0,23	1,29	0,85 1,96
Ele by Madre cons.	0,55	0,30	0,06	1,73	0,97 3,10

E' eliminato il determinante *Fem vs Mas*

Variabile	B	S.E.	p	OR	I. C. 95%
Ele vs Medie	0,15	0,25	0,54	1,16	0,72 1,88
Madre cons. vs non cons.	0,28	0,21	0,19	1,32	0,87 2,00
Ele by Madre cons.	0,51	0,30	0,08	1,67	0,93 2,98
Constant	0,71	0,18	0,00		

Non appare opportuno eliminare l'interazione *Ele by Madre cons.*. Il valore di significatività (p=0.08) è sufficiente a far ritenere che i due determinanti si influenzino reciprocamente. Il modello finale comprende pertanto due determinanti e la loro interazione

L'analisi multivariata evidenzia un'interazione positiva (1,67) tra ordine di scuola e consumo di latte da parte della madre.

I dati finali possono anche essere presentati come OR per i diversi sottogruppi. Si può osservare come l'odds ratio calcolato con la regressione logistica sia uguale a quello calcolato per il gruppo in cui il cofattore è assente. L'odds calcolato per il gruppo in cui il cofattore è presente è uguale al prodotto tra l'odds ratio precedente e il fattore di interazione.

Si osserva una diminuzione significativa nel consumo di latte tra le scuole ele e le scuole medie solo nei figli di madri consumatrici di latte (OR=1.94, I.C. 95% 1.40-2.67). Il consumo di latte da parte della madre incide quindi significativamente sul comportamento dei figli solo nelle elementari (OR=2,20 I.C. 95% 1.47-3.29).

FATTORE	COFATTORE	OR	I.C.95%
Ele vs Medie	Madre non cons.	1.16	0.72 1.88
Ele vs Medie	Madre cons.	1.94	1.40 2.67 (1.16*1.67)
Madre cons.	Medie	1.32	0.87 1.99
Madre cons.	Ele	2.20	1.47 3.29 (1.32*1.67)

Interazione

1.67