- Copyright - Il Pensiero Scientifico Editore downloaded by IP 18.97.14.88 Mon, 05 Oct 2026, 14:14:27

Note statistiche. Misure di effetto assolute e relative

Ettore Marubini, Giuseppe Reina

Istituto di Statistica Medica e Biometria, Università degli Studi, Milano

Key words:
Effect measures; Randomized controlled clinical trials; Superiority.

The aim of these statistical notes was to give the cardiologist the suitable tools for the understanding of the statistical aspects in the reading of papers presenting the results of randomized controlled clinical trials, in the most simple and intuitive way, without any previous knowledge of basic statistics.

The fundamentals of the most common tools for the comparison of two experimental treatments, are developed by getting cue from classical cardiological examples and by guiding the reader, step by step, toward the understanding and critical discussion of the proposed examples. In particular, this note focuses the attention on the main "indicators" of absolute and relative effect measures by evidencing their most important differences.

(Ital Heart J Suppl 2004; 5 (6): 466-471)

© 2004 CEPI Srl

Questa nota è stata redatta nell'ambito di una ricerca supportata dal Consiglio Nazionale delle Ricerche (contratto n. CU02.00264.ST97).

Ricevuto il 9 febbraio 2004; nuova stesura il 31 marzo 2004; accettato l'1 aprile 2004.

Per la corrispondenza:

Prof. Ettore Marubini

Istituto di Statistica Medica e Biometria Università degli Studi Via Venezian, 1 20133 Milano E-mail: ettore.marubini@unimi.it

Introduzione

La presente è la prima di una serie di note statistiche motivate dalla richiesta di alcuni partecipanti alla discussione che seguì la relazione tenuta da uno degli estensori (EM) al 37° Convegno "Cardiologia 2003" (Milano, 22-26 settembre 2003). Infatti, essi ponevano in evidenza un'estrema difficoltà nella comprensione del paragrafo "Metodi statistici" ed una conseguente incertezza nell'interpretazione dei "Risultati" dei rapporti di sperimentazioni cliniche controllate randomizzate (SCCR) pubblicati nelle riviste scientifiche di lettura usuale per il cardiologo. La richiesta verteva sull'opportunità di disporre di conoscenze metodologiche di base per la programmazione di una SCCR e di strumenti statistici elementari per la corretta valutazione dei risultati presentati, così da poter leggere, in chiave critica, le "Conclusioni" proposte dagli autori.

La sperimentazione clinica controllata randomizzata di superiorità GISSI

Le SCCR di superiorità sono programmate per porre in luce una differenza tra due trattamenti a confronto. La SCCR di superiorità GISSI[1], che aveva l'obiettivo di valutare l'effetto della streptochinasi (SK) contro un gruppo di riferimento, che per semplicità sarà qui indicato come controllo (C), informa che, nel gruppo trattato con SK, il numero di morti entro 21 giorni dal ricovero per infarto miocardico acuto (IMA), è 628 su un totale di 5860 pazienti trattati. Ne consegue che la *frequenza relativa* di morte entro 21 giorni o, per brevità, la mortalità a 21 giorni, m_{SK}, è pari a:

$$m_{SK} = \frac{628}{5860} = 0.1072 \simeq 10.7\%$$

dove il deponente SK sta ad indicare il gruppo del nuovo trattamento sperimentale. L'uso della lettera latina *m* specifica che si tratta di una quantità osservata sull'insieme dei pazienti reclutati in GISSI e ci si aspetta che possa essere diversa, per puro effetto del caso, da quella osservata in un'eventuale altra SCCR avente lo stesso obiettivo. In altri termini, ci si aspetta che le quantità osservate varino da SCCR a SCCR. Infatti, m_{SK} è una *stima* della *vera* frequenza relativa o probabilità di morte nella popolazione *obiettivo* e tale probabilità sarà qui indicata con la lettera greca μ_{SK} [la convenzione di usare lettere greche per indicare quantità *vere* e le corrispondenti lettere latine per indicare le rispettive *stime*, sarà sempre usata in questa serie di note]. Di fatto μ_{SK} è l'ignota quantità che il clinico vorrebbe conoscere ma, per quanto grande possa essere la numerosità dello studio, egli non può fare altro che basare la propria evidenza su una stima empirica m_{SK}.

- Copyright - Il Pensiero Scientifico Editore downloaded by IP 18.97.14.88 Mon, 05 Oct 2026, 14:14:27

Nell'attuale contesto, la popolazione *obiettivo* da trattare con SK è costituita dai pazienti che, fra tutti quelli che si presentano con IMA ad una unità coronarica, soddisfano i criteri di eleggibilità della sperimentazione; nell'esperienza dello studio GISSI, trattasi del 37.1% dei pazienti presentatisi con IMA.

Con riferimento al gruppo controllo, la mortalità m_C è pari a:

$$m_C = \frac{758}{5852} = 0.1295 \simeq 13.0\%$$

ed è stima di μ_C, *vera* frequenza relativa di morte del gruppo di controllo.

Per eseguire il confronto di due quantità, il clinico dispone di due operazioni aritmetiche elementari: la differenza ed il rapporto. In particolare, la differenza:

$$d = m_{SK} - m_C = 10.7\% - 13.0\% = -2.3\% \qquad (1)$$

è misura della riduzione assoluta di mortalità indotta dal trattamento con SK. Si è visto che m_{SK} e m_C sono stime puntuali rispettivamente delle ignote μ_{SK} e μ_C; parimenti, d è stima puntuale dell'ignota *vera* differenza, δ, tra le mortalità dei due gruppi posti a confronto. A seconda dei risultati di una SCCR, d può essere più o meno vicina all'ignota δ e quindi, come stima di δ, può essere più o meno *precisa*. La misura di tale precisione è fornita dall'errore standard. Questo, insieme alla stima d, permette di calcolare l'intervallo di confidenza (IC) di δ ad un *livello di confidenza* prespecificato (usualmente 95%). Tale intervallo rappresenta lo strumento di *inferenza statistica* di maggior importanza in medicina.

I dati di GISSI permettono di ottenere l'IC al 95% di δ che risulta pari a: *[-3.41%, -1.07%]* e di affermare che tale intervallo include l'ignota δ. La probabilità che questa affermazione sia corretta è pari al 95%.

L'insieme dei possibili valori che l'ignota δ può assumere, sono indicati graficamente in figura 1A.

Valori negativi dell'asse orizzontale, specificano una mortalità con SK minore di quella con C (SK "*superiore*" a C); valori positivi dell'asse, specificano una mortalità con SK maggiore di quella con C (SK "*inferiore*" a C). Il segmento riportato sopra l'asse orizzontale, corrisponde all'IC al 95% di δ ed include il sottoinsieme di valori di δ plausibili in base ai risultati GISSI. È importante realizzare che l'IC giace interamente nella parte negativa dell'asse e non include lo zero. Il fatto che tale intervallo non includa lo zero equivale, in termini di test di ipotesi, ad affermare che la differenza tra le mortalità dei due gruppi, d = -2.3%, è statisticamente significativa con $p \leq 0.05$ (o più precisamente, come riferito in GISSI, $p = 0.0002$); ciò può interpretarsi affermando che l'effetto di SK è superiore a quello di C in modo statisticamente significativo.

In alternativa all'uso della differenza assoluta d, il clinico può preferire l'uso della differenza relativa d_R:

$$d_R = \frac{m_{SK} - m_C}{m_C} = \frac{10.7\% - 13.0\%}{13.0\%} \simeq -17.7\% \qquad (2)$$

La relazione (2) può riscriversi:

$$d_R = \frac{m_{SK}}{m_C} - 1 \qquad (3)$$

dove il rapporto $\frac{m_{SK}}{m_C}$ costituisce un'importante misura di effetto della SK relativamente a quello del C ed è noto con il termine di "*rischio relativo*" (RR). Precisamente:

$$RR = \frac{m_{SK}}{m_C} = \frac{10.7\%}{13.0\%} \simeq 82.3\%$$

Come risulta dalla "Tabella IV" di GISSI, l'IC del *vero* rischio relativo, ρ_R, nella popolazione obiettivo è pari a: *[72.0%, 90.0%]*. La figura 1B è l'analogo, in termini di rischio relativo, della figura 1A.

Dal confronto delle figure 1A ed 1B, appare evidente che il valore $(100\ \rho_R)\% = 100\%$ è l'equivalente del valore $\delta = 0$ ed ambedue specificano l'equiattività dei due trattamenti. La figura 1B mostra che l'IC al 95% di ρ_R non include il valore 100% e poiché giace interamente alla sinistra di tale valore prova che il trattamento con SK è "superiore" a quello con C in modo statisticamente significativo, analogamente a quanto già visto in termini di differenza assoluta.

Giusta la relazione (3), l'IC di ρ_R permette di ricavare facilmente l'IC di $\delta_R = \rho_R - 1$, *vera* differenza relativa nella popolazione obiettivo:

$$IC\ di\ \delta_R\ al\ 95\% = [72.0\%\text{-}100\%,\ 90.0\%\text{-}100\%] = [-28.0\%,\ -10\%]$$

che è rappresentato graficamente in figura 1C.

A questo punto, è opportuno considerare quanto riportato nel paragrafo "Patients and methods": *"The sample size of about 12 000 patients to be randomized was estimated according to the following criteria: initial estimate of baseline myocardial infarction mortality 12%; expected reduction in overall mortality as a consequence of SK treatment 20%; significance level 1% and power 95%"*.

La *potenza* (power) di una SCCR si può pensare come un indicatore della capacità dello studio di porre in evidenza una differenza tra i due trattamenti, pari al valore clinicamente rilevante (δ_R) previsto nella programmazione, quando tale differenza *effettivamente esista* nella popolazione obiettivo che si pensa sottoposta per metà al trattamento di controllo e per l'altra metà al

- Copyright - Il Pensiero Scientifico Editore downloaded by IP 18.97.14.88 Mon, 05 Oct 2026, 14:14:27

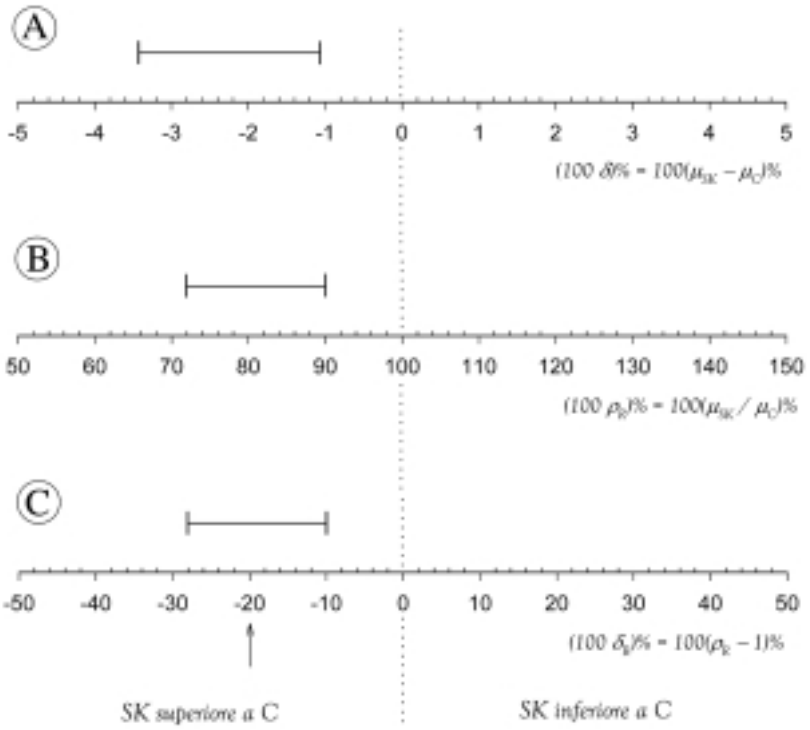

Figura 1. *Rappresentazione grafica dell'intervallo di confidenza al 95% di* δ (A), ρ_R (B) *e* δ_R (C) *calcolato sui risultati di GISSI. La* freccia *indica il valore di interesse clinico previsto nella programmazione della sperimentazione clinica controllata randomizzata. C = controllo; SK = streptochinasi.*

nuovo trattamento. Questo indicatore può assumere valori compresi tra 0 e 100%.

Con riferimento alla figura 1C, è degno di nota il fatto che l'IC al 95% di δ_R includa il valore di riduzione previsto in fase di programmazione dello studio, pari al 20%. Ciò significa che le aspettative dei clinici, riguardo all'entità dell'effetto della SK, sono state completamente soddisfatte.

La sperimentazione clinica controllata randomizzata di superiorità GUSTO III

Nel paragrafo "Statistical analysis" della SCCR di superiorità GUSTO III[2] in cui si confronta l'effetto di reteplase con alteplase, si afferma: *"The study design required the enrollment of 15 000 patients in order to have at least 85% power to detect a 20% relative reduction in mortality with reteplase as compared with alteplase"*. Nel paragrafo "Results" si legge: *"... the mortality rate at 30 days was 7.47% in the reteplase group and 7.24% in the alteplase group"*.

Con questi dati si possono calcolare: la differenza relativa d_R di reteplase (farmaco in esperimento) rispetto ad alteplase (controllo attivo)

$$d_R = \frac{7.47\% - 7.24\%}{7.24\%} = 3.18\%$$

e l'IC al 95% di δ_R pari a: *[-9.1%, 15.3%]*.

La rappresentazione grafica di questi risultati è riportata in figura 2.

Appare evidente come l'IC al 95% di δ_R includa lo zero e si estenda per circa un terzo nella parte negativa e due terzi nella parte positiva dell'asse. Pertanto questa SCCR è "non conclusiva" in quanto non permette di formulare alcun giudizio né di *superiorità*, né di *equivalenza*, né di *inferiorità* di reteplase rispetto ad alteplase. Al contrario, nel paragrafo "Discussion", gli autori di GUSTO III formulano un ben preciso giudizio affermando: *"... the chief finding of this trial is that reteplase is not superior to alteplase"*. Infine, è utile notare che, a differenza di quanto sottolineato a proposito dei risultati di GISSI, in GUSTO III il valore di interesse clinico (-20%) è ben lungi dall'osservato limite inferiore (-9.1%) dell'IC di δ_R. Dato che la potenza dello studio era ≥ 85%, ciò porta a pensare ad un eccessivo ottimismo, nella programmazione di GUSTO III, nel prevedere i vantaggi attesi di reteplase nei confronti di alteplase. Ancorché una riduzione relativa del 20% in mortalità sia *"... considered by many experienced clinicians to be a reasonable threshold for replacing a standard with an innovative therapy"*[3], tale soglia deve essere vagliata criticamente di volta in volta alla luce delle conoscenze sulle proprietà delle molecole che si pongono a confronto e dei risultati delle SCCR di fase II.

Con un valore di $(100 \cdot \delta_R)$% pari a -20%, la SCCR GUSTO III reclutò 15 000 pazienti. Il reclutamento sarebbe stato ben maggiore con un atteggiamento meno ottimistico. Infatti:

- Copyright - Il Pensiero Scientifico Editore downloaded by IP 18.97.14.88 Mon, 05 Oct 2026, 14:14:27

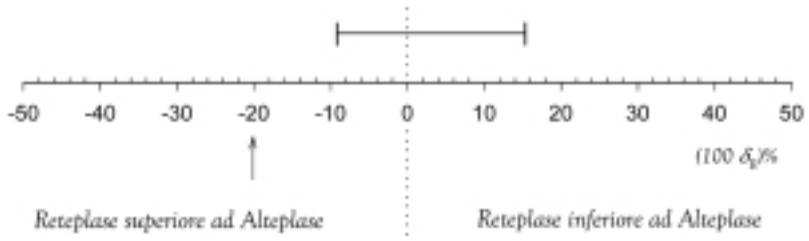

Figura 2. *Rappresentazione grafica dell'intervallo di confidenza al 95% di* δ_R *calcolato sui risultati di GUSTO III. La* freccia *indica il valore di interesse clinico previsto nella programmazione della sperimentazione clinica controllata randomizzata.*

(100 δ_R*)%*	*Dimensione della SCCR*
-20%	15 000
-15%	27 400
-10%	63 250
-5%	259 350

Si sarebbe davvero eseguito uno studio di queste dimensioni?

La sperimentazione clinica controllata randomizzata di superiorità ISIS-2

La SCCR di superiorità ISIS-2[4], che aveva come obiettivi la valutazione degli effetti separati e combinati di SK e di aspirina sulla mortalità in pazienti con "sospetto" IMA, nel paragrafo "Statistical methods" informa: *"Most results are presented in terms of changes in the odds of death ..."*.

Nell'accezione di uso corrente, il termine "*odds*" può essere tradotto in italiano solo mediante una perifrasi. Pertanto si ricorre al neologismo inglese (anche i francesi lo fanno!). L'Oxford Dictionary suggerisce la seguente definizione: "Nelle scommesse, [odds rappresenta] il vantaggio concesso da uno dei due scommettitori in proporzione alle assunte possibilità di vincita ... e consiste nel rapporto tra la somma in danaro che deve essere erogata rispetto a quella che deve essere ricevuta". Il lettore che ha familiarità con le corse dei cavalli, penserà immediatamente alla lavagnetta del bookmaker il quale, con riferimento ad esempio al cavallo Dedalus, con la specificazione 1:15 intende che a fronte di 1 euro ricevuto ne pagherà 15 in caso di vittoria di Dedalus (Dedalus è considerato decisamente un brocco!).

Nel contesto della presente nota, la posta in gioco è la sopravvivenza e quindi è ragionevole parlare di "*odds di morte*" (*o*) come rapporto:

$$o = \frac{numero\ morti}{numero\ vivi}$$

Per quanto attiene all'effetto di SK rispetto a C, è conveniente riportare il numero di morti vascolari entro 5 settimane, osservati in ISIS-2, sotto forma di "*tabella di contingenza*" 2×2:

	C	SK	Totale
Morti	1029	791	1820
Vivi	7566	7801	15 367
Totale	8595	8592	17 187

È immediato osservare che:

$$m_C = \frac{1029}{8595} = 0.1197 \simeq 12.0\%$$

$$m_{SK} = \frac{791}{8592} = 0.0921 \simeq 9.2\%$$

e che gli odds di morte vascolare entro 5 settimane risultano:

$$O_C = \frac{1029}{7566} = 0.1360 \simeq 13.6\%$$

$$O_{SK} = \frac{791}{7801} = 0.1014 \simeq 10.1\%$$

Una riduzione assoluta di mortalità dal 12.0 al 9.2%, ovvero -2.8%, comporta un rischio relativo pari a:

$$RR = \frac{9.2\%}{12.0\%} = 0.7667 \simeq 76.7\%$$

e, dalla relazione (3), una differenza relativa di mortalità:

$$d_R = 76.7\% - 100\% = -23.3\%$$

Analogamente, una riduzione assoluta di odds di mortalità dal 13.6 al 10.1%, ovvero -3.5%, comporta un "*odds ratio*" (*OR*) pari a:

$$OR = \frac{O_{SK}}{O_C} = \frac{10.1\%}{13.6\%} = 0.7426 \simeq 74.3\%$$

ed una differenza relativa degli odds di morte ($d_R o$):

- Copyright - Il Pensiero Scientifico Editore downloaded by IP 18.97.14.88 Mon, 05 Oct 2026, 14:14:27

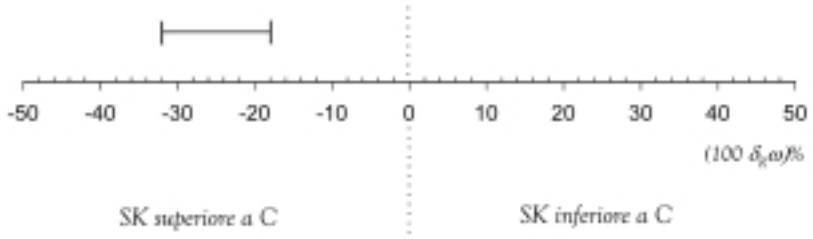

Figura 3. *Rappresentazione grafica dell'intervallo di confidenza al 95% di $\delta_R\omega$ calcolato sui risultati di ISIS-2. C = controllo; SK = streptochinasi.*

$$d_R o = \frac{O_{SK}}{O_C} - 1 = OR - 1 \simeq -25.7\%$$

Per poter passare direttamente da *RR* ad *OR*, il cardiologo può avvalersi della relazione che segue:

$$OR = RR \frac{1 - m_C}{1 - m_{SK}} \qquad (4)$$

da cui, come già visto in precedenza:

$$OR = 76.7\% \frac{100\% - 12.0\%}{100\% - 9.2\%} = 74.3\%$$

La relazione (4) mostra che se, in percentuale, i livelli di mortalità in ambedue i gruppi a confronto sono trascurabili rispetto al valore 100% oppure se sono molto "vicini" tra loro, le quantità *RR* ed *OR* tendono ad assumere valori molto simili. Per una trattazione esauriente riguardo alla definizione di "trascurabilità" si rimanda a Marubini e Valsecchi[5].

Nel paragrafo "Results" di ISIS-2 si legge che l'IC al 95% di $\delta_R\omega$, *vera* differenza relativa degli odds di morte nella popolazione obiettivo, è: *[-32.0%, -18.0%]*.

La rappresentazione grafica di tale intervallo è riportata in figura 3.

Nel paragrafo "Introduction" si legge ancora: "*Reductions in mortality that are realistically moderate (e.g. 20%÷25%) are important ...*". La figura 3 mostra come l'IC al 95% di $\delta_R\omega$ non includa lo zero e quindi si possa affermare che l'effetto di SK è significativamente superiore a quello di C ($p \leq 0.05$). Inoltre, è degno di nota il fatto che questo intervallo include l'insieme dei valori di riduzione relativa di odds ritenuti importanti [20%÷25%] dal punto di vista clinico.

L'interpretazione delle quantità *RR* e d_R è intuitivamente ovvia; meno intuitiva è l'interpretazione di *OR* e $d_R o$. Dato che le due quantità *RR* ed *OR* (e simmetricamente d_R e $d_R o$) veicolano la medesima informazione, è ragionevole che il cardiologo si chieda il motivo dell'introduzione di *OR* e di $d_R o$. La giustificazione è di natura strettamente tecnico-statistica[6] ed esula dal livello di trattazione della presente nota. Tale giustificazione porta a preferire l'uso di *OR* rispetto ad *RR*.

Con riferimento agli IC di ρ_R ed ω_R (*vero* odds ratio nella popolazione obiettivo), dai dati di ISIS-2 si ottengono:

$$IC\ al\ 95\%\ di\ \rho_R = [71.4\%, 82.8\%]$$

$$IC\ al\ 95\%\ di\ \omega_R = [68.7\%, 81.0\%]$$

Si è quindi in grado di asserire che, nell'esempio in questione, non solo le due stime puntuali di ρ_R ed ω_R sono praticamente coincidenti (76.7%, 74.3% rispettivamente), ma anche gli IC stessi: le due quantità sono quindi, in questo caso, interscambiabili. Va infine peraltro aggiunto che *OR* e $d_R o$ sono le misure di effetto di più frequente uso nella valutazione dei trombolitici.

Considerazioni finali

Negli esempi sopra riportati, gli "indicatori" di effetto presentati (d, d_R, $d_R o$, *RR*, *OR*) conducono a conclusioni praticamente sovrapponibili. Per una disamina delle loro proprietà e dei vantaggi e svantaggi di ciascuno si rimanda a Schechtman[7]. Qui sembra sufficiente riportare un suggerimento pratico di questo autore: *"... it is generally more meaningful to use relative effect measures for summarizing the evidence and absolute measures for application to a concrete clinical or public health situation"*.

In ogni caso, le formule sopra riportate permettono al lettore di passare dall'uno all'altro dei tre indicatori in modo immediato, così da poter interpretare i risultati della SCCR nel modo più congeniale a ciascun lettore.

Riassunto

Scopo di queste brevi note statistiche è di fornire al cardiologo gli strumenti necessari per la comprensione degli aspetti statistici relativi agli articoli che riferiscono i risultati di sperimentazioni cliniche controllate randomizzate, nella maniera più semplice ed intuitiva possibile, senza che alcuna nozione di statistica di base sia necessaria per la lettura del presente manoscritto.

L'introduzione dei concetti fondamentali relativi agli strumenti più comunemente usati per il confronto

- Copyright - Il Pensiero Scientifico Editore downloaded by IP 18.97.14.88 Mon, 05 Oct 2026, 14:14:27

di due trattamenti sperimentali, è svolta prendendo spunto, di volta in volta, da esempi classici della letteratura cardiologica e guidando il lettore passo passo alla lettura, comprensione e discussione critica degli esempi proposti. In particolare, questa nota pone l'attenzione sui principali "indicatori" di misure di effetto assolute e relative cui i vari esempi riportati fanno riferimento, ponendone in evidenza le differenze fondamentali.

Parole chiave: Misure di effetto; Sperimentazioni cliniche controllate randomizzate; Superiorità.

Ringraziamenti

Un ringraziamento particolare è dovuto agli anonimi revisori i cui puntuali commenti hanno consentito di rendere questa nota di più immediata lettura.

Glossario

- Differenza assoluta di mortalità:
  d = mortalità nel gruppo "nuovo trattamento" - mortalità nel gruppo "controllo"
- Differenza relativa di mortalità:

$$d_R = \frac{\textit{differenza assoluta di mortalità}}{\textit{mortalità nel gruppo "controllo"}}$$

- Differenza relativa degli odds di morte:

$$d_R o = \frac{\textit{odds di morte nel gruppo "nuovo trattamento"}}{\textit{odds di morte nel gruppo "controllo"}} - 1$$

- Errore standard:
  misura di imprecisione di una quantità osservata utilizzata come stima di una ignota quantità vera
- Intervallo di confidenza al 95% di una ignota quantità vera:
  intervallo che si afferma contenere una ignota quantità vera. La probabilità che questa affermazione sia corretta è pari al 95%
- Mortalità o frequenza relativa di morte:

$$m = \frac{\textit{numero di morti}}{\textit{numero di pazienti reclutati}}$$

- Odds di morte:

$$\omega = \frac{\textit{probabilità di morte}}{\textit{probabilità di sopravvivenza}}$$

è stimato da

$$o = \frac{\textit{numero di morti}}{\textit{numero di vivi}}$$

- Odds ratio:

$$OR = \frac{\textit{odds di morte nel gruppo "nuovo trattamento"}}{\textit{odds di morte nel gruppo "controllo"}}$$

- Rischio relativo:

$$RR = \frac{\textit{mortalità nel gruppo "nuovo trattamento"}}{\textit{mortalità nel gruppo "controllo"}}$$

- Sperimentazioni cliniche controllate randomizzate di superiorità:
  mirano a porre in evidenza il beneficio del "nuovo trattamento" rispetto al "controllo"

Bibliografia

1. Gruppo Italiano per lo Studio della Streptochinasi nell'Infarto Miocardico (GISSI). Effectiveness of intravenous thrombolytic treatment in acute myocardial infarction. Lancet 1986; i: 397-402.
2. The Global Use of Strategies to Open Occluded Coronary Arteries (GUSTO III) Investigators. A comparison of reteplase with alteplase for acute myocardial infarction. N Engl J Med 1997; 337: 1118-23.
3. Ware JH, Antman EM. Equivalence trials. N Engl J Med 1997; 337: 1159-61.
4. ISIS-2 (Second International Study of Infarct Survival) Collaborative Group. Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17 187 cases of suspected acute myocardial infarction: ISIS-2. Lancet 1988; 2: 349-60.
5. Marubini E, Valsecchi MG. Analysing Survival data from clinical trials and observational studies. Chichester: John Wiley & Sons, 1995.
6. Cox DR. Analysis of binary data. London: Chapman and Hall, 1970.
7. Schechtman E. Odds ratio, relative risk, absolute risk reduction, and the number needed to treat - which of these should we use? Value Health 2002; 5: 431-6.