

# Il confondimento: un bias della stima dell'effetto fra una esposizione ed un esito di salute

Giornale di Tecniche Nefrologiche e Dialitiche  
2018, Vol. 30(3-4) 216–218

© The Author(s) 2019

Article reuse guidelines:

[sagepub.com/journals-permissions](http://sagepub.com/journals-permissions)

DOI: 10.1177/0394936218823521

[journals.sagepub.com/home/gtn](http://journals.sagepub.com/home/gtn)



Francesco Franco e Anteo Di Napoli

## Abstract

### Confounding: a bias in the estimation of the effects of exposure to health outcome

An epidemiological study aims to investigate the relationship between an exposure and an outcome. The presence of another variable associated with the exposure, and with the disease outcome, may introduce a bias in that relationship. However, if a variable is part of the causal chain between exposure and outcome, it is not a confounding factor (or confounder). The presence of potential confounders must always be considered in the design and analysis of epidemiological studies. Controlling for confounders may be done both at the design stage and during analysis of data. The methods used for controlling for confounders at the design stage are restriction, matching, and randomization. These methods allow making similar distributions of confounders for each study group. To detect the presence of confounders during the analysis stage, the researcher may evaluate differences in the estimates of the association between the exposure and the outcome without and with the confounder (called crude and adjusted estimates, respectively). Stratification, standardization and multivariate analysis are also methods to control for the presence of confounders.

## Keywords

Confounding, Bias, Confounders, Causal diagram, Spurious effect

“A Dictionary of epidemiology”<sup>1</sup> definisce l’epidemiologia come “lo studio della distribuzione dei determinanti degli stati o eventi di salute in popolazioni specifiche e l’applicazione di questo studio per il controllo dei problemi di salute”.<sup>1</sup>

Stabilire l’esistenza di un’associazione statistica tra un’esposizione e un esito di salute è uno degli elementi che caratterizzano l’approccio epidemiologico. Se ipotizziamo, ad esempio, che un certo fattore di rischio sia un determinante di una malattia, ci aspettiamo che la sua frequenza sia diversa tra i soggetti esposti a quel fattore rispetto ai soggetti non esposti. Tuttavia, tale confronto deve tener conto dell’eventuale ruolo svolto da altri fattori di rischio (antecedenti o concomitanti), al fine di valutare se l’associazione trovata tra fattore di esposizione ed esito di salute sia vera (causale, eziologica) oppure falsa (non causale, spuria). Per definire valida un’associazione si deve prima escludere che essa possa essere spiegata dal caso, da una distorsione sistematica (bias) all’origine delle informazioni su esposizione ed esito o da altri fattori, i cosiddetti fattori di confondimento. Un fattore di confondimento è una variabile in grado di produrre un’associazione apparente (confondimento

positivo), oppure il mascheramento di un rapporto realmente esistente (confondimento negativo) tra un’esposizione ed un esito di salute.

Nel precedente articolo ci siamo occupati dei bias di selezione e di informazione,<sup>2</sup> mentre in quello presente approfondiremo una terza tipologia di bias: il confondimento.

Il termine confondimento si riferisce ad una situazione in cui un’associazione non causale fra una data esposizione ed un esito è osservata come conseguenza dell’influenza di una terza variabile (o gruppo di variabili) denominata variabile di confondimento.<sup>3</sup>

Si ha confondimento quando le differenze osservate, confrontando gli esposti e i non esposti a un fattore di rischio rispetto all’insorgenza di una malattia, sono in realtà causate da fattori di rischio estranei che coinvolgono meccanismi diversi da quelli di interesse dello studio. Si determina così una distorsione dell’associazione tra

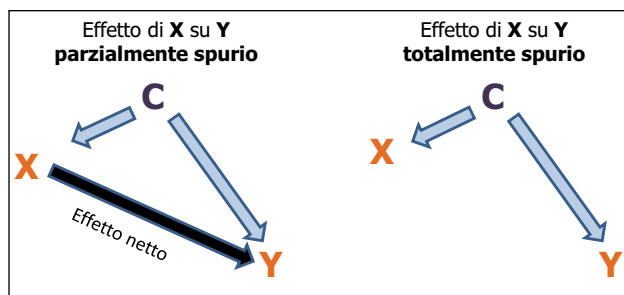
---

Comitato Editoriale GTND, Italy

### Corresponding author:

Francesco Franco, Comitato Editoriale GTND, Italy.

E-mail: [franco\\_francesco@fastwebnet.it](mailto:franco_francesco@fastwebnet.it)



**Figura 1.** Diagramma causale del confondimento. Associazione spuria fra esposizione ed esito.

esposizione ed esito di salute che confonde e disturba lo studio.

In presenza di confondimento i gruppi di esposti e non esposti a un fattore di rischio non sono confrontabili, perché hanno differenti rischi di base per una certa malattia a causa della diversa distribuzione del fattore che confonde l'associazione. Questo determina la presenza di un'associazione statistica che sarebbe errato interpretare in senso causale.

Affinché si verifichi confondimento sono necessarie tre condizioni:

1. il confondente deve essere associato con l'esposizione in studio nella popolazione di origine;
2. il confondente deve essere un fattore di rischio per l'esito (deve, cioè, predire chi svilupperà l'esito) sebbene non sia necessariamente la causa dell'esito;
3. il confondente non deve essere influenzato dall'esposizione o dall'esito.<sup>4</sup>

Inoltre, il confondente non deve essere un fattore intermedio nella catena di causalità tra esposizione ed esito.

In Figura 1 è riportato il diagramma causale del confondimento.<sup>5</sup>

Alcuni fattori sono sempre considerati come potenziali confondenti quali: il sesso, l'età, il periodo di calendario, lo status socio-economico, gli stili di vita (fumo, alcol, dieta) e l'occupazione.

Esistono diverse strategie per il controllo del confondimento che in linea di massima si possono classificare in base al fatto che avvengano prima della raccolta dei dati, cioè in fase di disegno dello studio, oppure dopo la raccolta dei dati, attraverso l'aggiustamento durante l'analisi dei dati. I metodi di controllo del confondimento legati al disegno dello studio sono: la restrizione, l'appaiamento e la randomizzazione; quelli legati all'analisi dei dati sono: la stratificazione, la standardizzazione (diretta e indiretta) e l'utilizzo di modelli di regressione.<sup>6</sup>

La restrizione, per l'appunto, restringe i criteri di eleggibilità ai soggetti appartenenti alla medesima

categoria del confondente. L'idea è che una variabile non produce confondimento se le viene impedito di variare. Se, ad esempio, si sospetta che il sesso possa confondere l'associazione tra un fattore di rischio e una malattia si può restringere lo studio ai soli uomini o alle sole donne. La restrizione ha il vantaggio di essere un metodo estremamente semplice ed efficace, ma ha il limite di ridurre la generalizzabilità dei risultati solo alla tipologia di soggetti inseriti nello studio (solo uomini o solo donne nell'esempio).

L'appaiamento è una metodologia di controllo del confondimento in cui i non esposti sono resi simili agli esposti nella distribuzione di uno o più potenziali confondenti, appaiando i due gruppi per tali caratteristiche. Si rimuove la prima condizione per il confondimento, cioè l'associazione tra esposizione e confondente, visto che il fattore di confondimento non varia più nei due gruppi di esposti e non esposti al fattore di rischio in studio. Il limite dell'appaiamento è il cosiddetto overmatching: se la variabile d'appaiamento è altamente correlata all'esposizione ma non è un fattore di rischio per la malattia, in definitiva si appaia anche per l'esposizione che non può più essere valutata come fattore di rischio. Se la variabile d'appaiamento è un elemento intermedio della catena causale tra esposizione e malattia, rimuovendo la variabilità rispetto a tale variabile, non si potranno più valutare differenze neanche per l'esposizione, come accadrebbe, ad esempio, se in uno studio che valuti l'associazione tra esposizione a alimentazione ricca di grassi e insorgenza di arteriosclerosi si appaiassero i gruppi di esposti e non esposti per livello di colesterolo.

Un altro metodo che consente il controllo del confondimento in fase di disegno dello studio è la randomizzazione. In uno studio clinico randomizzato controllato i pazienti sono assegnati casualmente ai gruppi di trattamento e controllo. L'obiettivo è creare gruppi col medesimo rischio di base verso l'esito in studio prima dell'assegnazione al trattamento (background risk). Il vantaggio della randomizzazione è che consente di bilanciare anche confondenti ignoti, se la numerosità è sufficiente. Al momento dell'analisi statistica è fondamentale che i gruppi a confronto siano valutati tenendo conto dell'assegnazione al trattamento, indipendentemente dal fatto che abbiano o meno effettuato il trattamento medesimo (analisi per intenzione al trattamento o intention to treat).<sup>7</sup>

Il controllo del confondimento in fase di analisi ha lo scopo di ottenere misure di effetto (Rischio Relativo, RR, Odds Ratio, OR, Risk Difference, RD) aggiustate per i potenziali confondenti.

Si possono utilizzare i metodi della stratificazione e della standardizzazione (Standardized Rate Ratio, SRR, Standardized Mortality Ratio, SMR, Mantel-Haenszel), che però sono adatti quando i confondenti sono pochi, due o tre al massimo.

**Tabella 1.** Tavola di contingenza tra esposizione ad un fattore X ed esito Y (dati fittizi).

	Esito (Y)		Totale	
	+	-		
Esposizione (X)	+	340	1.210	1.550
	-	205	1.345	1.550
Totale	545	2.555	3.100	

Percentuale dell'esito Y fra gli esposti a X:  $(340 / 1.550) \times 100 = 21,9\%$ .  
 Percentuale dell'esito Y fra i non esposti a X:  $(205 / 1.550) \times 100 = 13,2\%$ .

Fisher's exact p-value: 0.000 (differenza della percentuale di esito fra i due gruppi statisticamente significativa).

**Tabella 2.** Tavola di contingenza tra un potenziale fattore di confondimento Z ed esito Y (dati fittizi).

	Esito (Y)		Totale	
	+	-		
Esposizione (Z)	+	262	1.049	1.311
	-	90	1.699	1.789
Totale	352	2.748	3.100	

Percentuale dell'esito Y fra gli esposti a Z:  $(262 / 1.311) \times 100 = 20\%$ .  
 Percentuale dell'esito Y fra i non esposti a Z:  $(90 / 1.789) \times 100 = 5\%$ .

Fisher's exact p-value: 0,000 (differenza della percentuale di esito fra i due gruppi statisticamente significativa).

La stratificazione è la valutazione dell'associazione tra esposizione e malattia entro strati omogenei della variabile confondente. Il vantaggio risiede nel fatto che si tratta di un metodo intuitivo che fornisce immediata comprensione di quanto i dati siano interconnessi. Lo svantaggio è che è difficile da effettuare se ci sono molti confondenti e quindi molti strati.

Come esempio consideriamo uno studio fittizio che illustra l'azione di un fattore di confondimento Z sul presunto rapporto di causa-effetto osservato fra un'esposizione X ed un esito Y. In un campione di 3.100 soggetti, 1.550 sono stati esposti al fattore di rischio X e 1.550 no. Nel gruppo degli esposti a X, 340 soggetti hanno sperimentato l'esito Y in esame mentre nel gruppo dei non esposti sono stati 205 (tabella 1).

Si ipotizza, a questo punto, che la differenza osservata nella percentuale di esito fra esposti e non esposti a X non sia dovuta all'azione del fattore di esposizione considerato bensì all'operato di un terzo fattore Z associato all'esposizione X, e fattore di rischio per l'esito Y (tabella 2).

Per isolare l'effetto dell'esposizione al fattore X da quella al fattore di confondimento Z, eseguiamo un'analisi stratificata dei dati (tabella 3).

All'interno degli strati di Z le differenze della percentuale di esito fra esposti e non esposti al fattore X scompaiono. Il

**Tabella 3.** Tavole di contingenza stratificate per l'analisi del confondimento su esposizione ad un fattore X ed esito Y (dati fittizi).

	Esposizione (X)				Totale	
	+		-			
	Esito (Y)		Esito (Y)			
	+	-	+	-		
Esposizione (Z)	+	238	954	24	95	1.311
	-	18	340	72	1.359	1.789
Totale	256	1.294	96	1.454	3.100	

A parità di Z (analisi di riga della tabella):

Percentuale dell'esito Y fra gli esposti a Z e a X:  $(238 / (238 + 954)) \times 100 = 20\%$ .

Percentuale dell'esito Y fra gli esposti a Z e non a X:  $(24 / (24 + 95)) \times 100 = 20,2\%$ .

Percentuale dell'esito Y fra i non esposti a Z ed esposti a X:  $(18 / (18 + 340)) \times 100 = 5\%$ .

Percentuale dell'esito Y fra i non esposti a Z e a X:  $(72 / (72 + 1.359)) \times 100 = 5\%$ .

risultato dello studio dimostra, quindi, che l'apparente associazione osservata fra il fattore di esposizione X e l'esito Y era dovuta solamente all'azione del fattore di confondimento Z (effetto di X su Y totalmente spurio).

### Dichiarazione di assenza di conflitto di interessi

Gli Autori dichiarano di non avere conflitti di interessi.

### Finanziamenti

Gli Autori dichiarano di non aver ricevuto finanziamenti specifici da qualsiasi ente nei settori pubblico, privato o senza fini di lucro.

### Bibliografia

1. Last JM. A dictionary of epidemiology. Fourth edition. Oxford: Oxford University Press 2001.
2. Franco F and Di Napoli A. Il bias: un'insidia per la validità di uno studio [Bias: a pitfall to the validity of a study]. *Gior di Tecn Nefr e Dial*, published online Nov 2018.
3. Szklo M and Nieto FJ. Epidemiology: beyond the basics. Gaithersburg, MD: Aspen Publishers 2000; 177-180.
4. Rothman KJ, Greenland S and Lash TL. Modern epidemiology. Third ed. Philadelphia: Lippincott Williams & Wilkins 2008; 12:194-196.
5. Corbetta P, Gasperoni G and Pisati M. Statistica per la ricerca sociale. Bologna (Italia): il Mulino. 2001; 7:186-190.
6. Rothman KJ. Epidemiologia. Napoli (Italia): Idelson-Gnocchi 2007; 5:102-105.
7. Wassertheil-Smoller S and Smoller J. Biostatistics and Epidemiology, a Primer for Health and Biomedical Professionals. Fourth Edition New York: Springer 2015; 152-153.