

Il bias: un'insidia per la validità di uno studio

Francesco Franco e Anteo Di Napoli

Giornale di Tecniche Nefrologiche e Dialitiche
2018, Vol. 30(2) 147-149
© The Author(s) 2018
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0394936218807904
journals.sagepub.com/home/gtn



Bias: a pitfall for the validity of a study

A bias is one of the major concerns in almost all types of epidemiological studies. It is defined as any systematic error in the design, conduct, or analysis of a study that causes an incorrect estimate of an exposure's effect on the risk of illness.

There are several types of bias that can be classified in two main groups: selection bias and information bias. In a selection bias the error is introduced when the study population does not represent the target population. An information bias occurs when an error is introduced during the detection of exposures and/or effects, thus producing a misclassification of exposed/diseased subjects.

Keywords

Selection bias, Information bias, Berkson's bias, Healthy worker effect, Salmon effect, Misclassification bias, Recall bias

Nel precedente articolo abbiamo introdotto l'argomento dei presupposti della validità interna di uno studio che si riferisce alla validità dell'inferenza basata sul campione di individui esaminati. Abbiamo poi ricordato che la maggior parte delle violazioni della validità interna possono essere classificate nelle tre categorie del confondimento, del bias di selezione e del bias di informazione.^{1,2}

Nel presente articolo approfondiremo l'argomento del bias, definito come un qualunque errore sistematico presente nel disegno, nella conduzione o nell'analisi di uno studio epidemiologico e che si traduce in una stima errata dell'effetto dell'esposizione sul rischio di malattia.³

Il bias può essere introdotto in uno studio dal modo in cui si selezionano i soggetti, dalla maniera in cui si misurano le variabili e dal non tener conto di potenziali fattori di confondimento.⁴

La presenza di bias in uno studio è particolarmente insidiosa, in quanto si tratta di un tipo di errore che non può essere valutato (ed eventualmente almeno in parte corretto) in fase di analisi statistica dei dati, ma che può essere prevenuto solo attraverso una pianificazione accurata e una corretta conduzione dello studio.

Sono numerose le tipologie di bias descritte⁵, ma la loro trattazione analitica esulerebbe dalle finalità di questa rubrica. In generale, i bias possono essere classificati in due grandi categorie: i bias di selezione, errori sistematici commessi nella fase di selezione della popolazione in studio, sia in fase di arruolamento che in corso di svolgimento (follow-up); i bias di informazione, errori sistematici

commessi durante il processo di acquisizione delle informazioni. Al confondimento, da taluni considerato come una terza tipologia di bias, sarà dedicato un articolo a parte.

Bias di selezione

In relazione al disegno dello studio, è possibile considerare i bias di selezione in due situazioni distinte.

Negli studi descrittivi, il bias si verifica quando la popolazione in studio non è rappresentativa della popolazione target (bersaglio). In particolare, negli studi trasversali, i bias di selezione sono dovuti essenzialmente alla non partecipazione dei soggetti al campione che s'intende arruolare nello studio (azione dei metodi di selezione e dei fattori che influenzano l'adesione dei soggetti allo studio). Se l'associazione fra esposizione ed esito in esame differisce tra i soggetti partecipanti e quelli non partecipanti allo studio si sarebbe in presenza di bias di selezione. Ad esempio, in un'indagine sul consumo di

Comitato Editoriale GTND

Corrispondenza:

Anteo Di Napoli, Comitato Editoriale GTND via Ischia di Castro, 10, 00189 Roma

Email: anteo.dinapoli@libero.it

Francesco Franco, Comitato Editoriale GTND, via Francesco Tovaglieri, 329 scala C

00155 Roma

E-mail: franco_francesco@fastwebnet.it

Tabella 1. Ipotetico studio caso-controllo di pazienti ospedalizzati per lo studio dell'associazione tra consumo di alcol e cirrosi epatica (Berkson's bias).

Fattore di rischio Consumo di alcol	Casi di cirrosi epatica	Controlli (ospedalizzati per trauma)	Controlli (popolazione generale)
Sì	70	50	20
No	30	50	80
Odds di esposizione	2.3	1.0	0.25
Odds ratio	2.3		9.3

alcol, i forti bevitori potrebbero essere indotti a non partecipare allo studio e questo si tradurrebbe in una sottostima sia della prevalenza dei bevitori che del livello di consumo di alcol nella popolazione.

Negli studi analitici, invece, il bias si verifica quando i gruppi a confronto non sono confrontabili. Negli studi caso-controllo, nei quali un gruppo di malati (casi) viene selezionato e confrontato con un appropriato gruppo di non casi (controlli), il bias interviene se i gruppi selezionati non sono rappresentativi dei casi e dei controlli nella popolazione. In particolare, rispetto alla selezione dei casi, il bias si verifica se i casi esposti hanno diversa probabilità di essere inclusi rispetto ai casi non esposti; ad esempio, in uno studio sui fattori di rischio dell'ipertensione, le donne che assumono contraccettivi orali hanno maggiore probabilità di essere incluse tra i casi, perché sono in genere più seguite e questo si traduce in una sovrastima dell'associazione tra contraccettivi orali e ipertensione. Rispetto alla selezione dei controlli, il bias si verifica se i controlli non sono rappresentativi della popolazione che ha generato i casi; ad esempio, in uno studio sui fattori di rischio per cirrosi epatica, i controlli selezionati tra i pazienti ospedalizzati per trauma potrebbero consumare più alcol rispetto alla popolazione generale e questo si traduce in una sottostima dell'associazione tra consumo di alcol e cirrosi epatica. Questo artefatto va sotto il nome di Berksonian bias in quanto descritto per la prima volta dal biostatistico Joseph Berkson (tabella 1). Un altro noto bias di selezione, che porta spesso a trarre conclusioni errate, è quello determinato dall'"effetto del lavoratore sano" (healthy worker effect) che si origina quando si mettono a confronto i tassi di alcuni esiti sfavorevoli di salute in particolari gruppi di lavoratori rispetto a quelli registrati nella popolazione generale. Ad esempio, il tasso complessivo di mortalità tra i lavoratori è spesso più basso rispetto a quello della popolazione generale e questo è dovuto al fatto che nella popolazione generale è presente un numero consistente di persone che non lavorano a causa di patologie in essere.

Negli studi longitudinali, trial clinici randomizzati controllati e studi di coorte, la possibilità di bias si verifica in particolar modo durante il follow-up, nel caso in cui le probabilità dei soggetti di essere persi si associa tanto all'esposizione quanto alla malattia. Un esempio è rappresentato dal cosiddetto "salmon bias", che viene sovente indicato come una delle possibili ipotesi per

spiegare la più elevata sopravvivenza registrata tra gli immigrati in un paese rispetto ai nativi: vi potrebbe essere una quota più elevata di soggetti persi al follow-up tra gli stranieri con un conseguente bias sul decesso tra i due gruppi in studio, sottostimandone la mortalità. Tale distorsione conosciuta appunto come "salmon bias" è stata descritta come effetto della tendenza, osservata talora tra le persone straniere in precarie condizioni di salute, di lasciare il Paese nel quale sono migrate per tornare in quello di origine, dove trascorrere le ultime fasi della propria vita.⁶

Quali sono le strategie da adottare per evitare o ridurre la possibilità di bias di selezione? Negli studi descrittivi è fondamentale minimizzare la percentuale di rifiuti, e comunque provare a raccogliere informazioni minime sui non rispondenti in modo da poter anche valutare se chi ha rifiutato di partecipare differisce o meno da chi ha accettato.

Negli studi caso-controllo, ove possibile, sarebbe utile ricorrere a controlli multipli, provenienti da più fonti, in modo da essere più sicuri che i controlli siano rappresentativi della popolazione che ha generato i casi, o comunque ad essa riferibile anche se la selezione avviene ad esempio in ambito ospedaliero.

Negli studi longitudinali si può scegliere di restringere lo studio a gruppi più facili da seguire (ad esempio residenti), di prevedere frequenti richiami dei partecipanti e anche di utilizzare fonti informative routinarie per accertare gli esiti (ad esempio registro cause di morte, schede di dimissione ospedaliera).

Bias di informazione

I bias di informazione nel disegno di studio derivano da una diversa qualità delle informazioni nei diversi gruppi a confronto (misclassificazione); in particolare, negli studi retrospettivi si determina un bias quando c'è una diversa probabilità di misclassificare l'esposizione tra malati e non malati, mentre negli studi prospettici quando c'è una diversa probabilità di misclassificare la malattia tra esposti e non esposti.

Relativamente al bias di informazione negli studi caso-controllo, data la loro natura fondamentalmente retrospettiva, la conoscenza dell'esito può influenzare, anche solo inconsciamente, l'accertamento dell'esposizione da parte del ricercatore; d'altro canto anche i malati (casi in studio) potrebbero ricordare non correttamente (recall

Tabella 2. Ipotetico studio caso-controllo per lo studio dell'associazione fra assunzione materna di farmaci in gravidanza e difetti congeniti del neonato (recall bias - misclassificazione differenziale).

Fattore di rischio Assunzione di farmaci accertata	Casi	Controlli
Sì	90	180
No	210	420
Odds di esposizione	0.4	0.4
Odds ratio	1.0	
Fattore di rischio Assunzione riferita dalla madre	Casi	Controlli*
Sì	90	126
No	210	474
Odds di esposizione	0.4	0.3
Odds ratio	1.6	

*percentuale di controlli esposti misclassificati come non esposti: 30%.

bias); in particolare, potrebbero essere indotti a ricordare meglio un'esperienza di esposizione e a riferirla attribuendole un ruolo maggiore o minore, a seconda che si verificano meccanismi di colpevolizzazione o auto-assolutori; ad esempio, una donna il cui bambino sia nato con difetti congeniti per i quali si ipotizza il ruolo dell'assunzione materna di farmaci in gravidanza, potrebbe colpevolizzarsi sovrastimando le dosi assunte o viceversa rifiutare il problema, ridimensionando o negando l'assunzione. Si determina quindi una misclassificazione differenziale dell'esposizione in quanto correlata alla presenza di malattia (tabella 2).⁴

Il bias di informazione negli studi prospettici si determina quando la conoscenza dello status di esposizione può influenzare l'accertamento dell'esito (detection bias); ad esempio, in uno studio di coorte sul cancro del polmone, i casi dubbi potrebbero essere più facilmente classificati come positivi tra i fumatori che tra i non fumatori.

Quali sono le strategie da adottare per minimizzare o evitare i bias di informazione? Sicuramente l'adozione di rigidi criteri diagnostici può limitarli, così pure porre diagnosi attraverso l'utilizzo di fonti oggettive (ad esempio esami di laboratorio o indagini strumentali) più che l'esperienza soggettiva. Un'altra possibilità potrebbe essere la cecità (blinding) del ricercatore rispetto all'informazione sull'esposizione al fattore di rischio del soggetto arruolato.

I bias di informazione, essendo causati da una misclassificazione differenziale dell'informazione su esposizione e malattia tra i gruppi degli esposti e non esposti e tra i gruppi dei malati e dei non malati, determinano una distorsione, in eccesso o in difetto, nelle stime dell'associazione tra esposizione a un fattore di rischio ed esito di malattia.

Accanto a questa tipologia di misclassificazione, esistono altre forme, dette non differenziali, che in linea di

Tabella 3. Ipotetico studio caso-controllo per lo studio dell'associazione fra consumo di alcol e K laringeo (misclassificazione non differenziale).

Fattore di rischio Consumo di alcol effettivo	Casi	Controlli
Sì	96	110
No	148	292
Odds di esposizione	0.6	0.4
Odds ratio	1.7	
Fattore di rischio Consumo di alcol riferito	Casi*	Controlli**
Sì	58	66
No	186	336
Odds di esposizione	0.3	0.2
Odds ratio	1.6	

*percentuale di casi esposti misclassificati come non esposti: 40%.

**percentuale di controlli esposti misclassificati come non esposti: 40%.

massima gli epidemiologi tendono a tenere distinte dai bias di informazione. Si verificano, appunto, quando la probabilità di errore nell'informazione non differisce nei gruppi a confronto (malati/non malati, esposti/non esposti). Va sottolineato che negli studi analitici, la presenza di una misclassificazione non differenziale produce l'effetto di attenuare le associazioni in studio tra fattore di esposizione ed esito (tabella 3).

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. Franco F e Di Napoli A. Bias, confondimento e interazione: tre temi fondamentali nell'ambito dell'inferenza causale [Bias, confounding, interaction: three fundamental topics in the field of causal inference]. *G Tec Nefrol Dial.* (in press)
2. Rothman KJ, Greenland S, Lash TL. *Modern epidemiology. Third ed.* Philadelphia: Lippincott Williams & Wilkins. 2008; 9:128–129.
3. Gordis L. *Epidemiology. Fourth ed.* Philadelphia: Elsevier Saunders. 2014; 11:201–214.
4. Rothman KJ. *Epidemiologia. Napoli (Italia): Idelson-Gnocchi.* 2007; 5:89–95.
5. Delgado-Rodriguez M e Llorca J. Bias. *J Epidemiol Community Health* 2004; 58:635–641. doi:10.1136/jech.2003.008466
6. Abraido-Lanza AF, Dohrenwend BP, Ng-Mak DS, et al. The Latino mortality paradox: a test of the "salmon bias" and healthy migrant hypotheses. *Am J Public Health* 1999; 89(10):1543–1548.