

UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

DIPARTIMENTO DI STATISTICA E METODI QUANTITATIVI
DOTTORATO DI RICERCA IN EPIDEMIOLOGIA E
BIOSTATISTICA
CICLO XXVII



Metodi per il controllo del confondimento non misurato e dell'errore di misura

Coordinatore:	Prof.ssa	Stefania Galimberti
Tutor accademico:	Prof.ssa	Antonella Zambon
Candidata:	Dott.ssa	Arianna Ghirardi

ANNO ACCADEMICO 2013-2014

Indice

I	Introduzione	3
1	Introduzione	4
1.1	I database amministrativi	4
1.2	Il confondimento non misurato	7
1.3	L'errore di misura	9
1.4	Obiettivo	10
II	Metodi per il controllo del confondimento non misurato	11
2	Quando agire?	12
3	Controllo del confondimento tramite disegno	14
3.1	User-only design	14
3.2	Case-only design	16
3.2.1	Case-crossover	17
3.2.2	Case-time-control	22
3.2.3	Case-case-time-control	28
3.2.4	Self-controlled-case series	31
4	Controllo del confondimento tramite tecniche analitiche	35
4.1	Monte-Carlo Sentivity Analysis	35
4.2	Approccio Rule-out	40
4.3	Propensity Score Calibration	42

III	Metodi per il controllo dell'errore di misura	45
5	Quando agire?	46
6	Metodi per il controllo dell'errore di misura tramite tecniche analitiche	47
6.1	Correzione dei Time-Related Bias	47
6.1.1	Immortal time bias	48
6.1.2	Time-window bias	52
6.1.3	Immeasurable time bias	54
6.2	Regression Calibration	58
IV	Applicazioni	61
7	Fonte dei dati	62
7.1	User-only design	64
7.1.1	Bifosfonati e Fratture, prevenzione secondaria	64
7.2	Case-only design	69
7.2.1	Antiinfiammatori e Sanguinamento gastrointestinale	69
7.2.2	Antidepressivi e Fratture	74
7.3	Monte-Carlo Sensitivity Analysis	79
7.3.1	Bifosfonati e Fratture, prevenzione primaria	79
7.4	Metodi per la correzione dell'ITB	83
7.4.1	Bifosfonati e Complicazioni gastrointestinali	83
7.5	Approccio Rule-Out	87
7.5.1	Bifosfonati e Complicazioni gastrointestinali	87
7.6	Regression Calibration	90
7.6.1	Antiinfiammatori e Sanguinamento gastrointestinali	90
V	Considerazioni finali	94

Parte I

Introduzione

Capitolo 1

Introduzione

1.1 I database amministrativi

Nonostante gli studi clinici randomizzati forniscano importanti informazioni sull'efficacia e sulla sicurezza dei farmaci, i risultati provenienti da tali studi non possono essere generalizzati a tutta la popolazione [1].

Questi studi infatti sono basati su campioni altamente selezionati, costituiti da un numero limitato di individui, che spesso sotto-rappresentano i pazienti più vulnerabili. La maggior parte degli studi clinici, inoltre, viene condotta in strutture ospedaliere, solitamente prestigiose, nelle quali i pazienti sono sottoposti a regimi dietetici e comportamentali strettamente controllati e non sono liberi di gestire le proprie abitudini (alimentari, stili di vita, etc.) secondo una propria scelta, una caratteristica che rende tali studi lontani dalla reale pratica clinica.

Di conseguenza, malgrado i risultati provenienti dagli studi clinici randomizzati siano sufficienti per approvare l'immissione in commercio di un farmaco, non lo sono per rispondere ai quesiti più rilevanti che devono affrontare quotidianamente medico e pazienti [2].

Un altro fattore da tenere in considerazione riguarda il fatto che l'aderenza al trattamento nella pratica clinica è decisamente inferiore a quella ottenuta in un contesto sperimentale [3-7]. Per questi motivi è necessario che l'osservazione degli individui in trattamento continui anche dopo la commer-

cializzazione del farmaco [8].

A tale proposito, nell'ultimo decennio si è assistito all'ampio utilizzo di database (DB) di pertinenza sanitaria, per la conduzione di studi di farmacoepidemiologia [9]. Tali database possono essere classificati in due grandi categorie: quelli che raccolgono informazioni a fini amministrativi (DB amministrativi) e quelli che vengono utilizzati dai medici per registrare informazioni cliniche sulla salute dei propri pazienti (DB dei medici di medicina generale o DB clinici). Attraverso le informazioni reperibili dai DB amministrativi, quali dati anagrafici, ricoveri ospedalieri e prescrizioni farmaceutiche, sono stati condotti diversi studi che hanno fornito importanti risultati circa il profilo di sicurezza ed efficacia dei farmaci.

Una delle principali caratteristiche di tali DB, che li rende molto apprezzati, riguarda la possibilità di riflettere la reale pratica clinica per popolazioni ampie e non selezionate, un aspetto che sicuramente supera uno dei principali limiti degli studi clinici randomizzati [10].

Nonostante i numerosi vantaggi che i DB amministrativi hanno apportato nella conduzione degli studi farmacoepidemiologici, sono emerse alcune criticità legate soprattutto all'incompletezza di importanti informazioni di tipo clinico, come i risultati di test diagnostici, il Body Mass Index o di informazioni relative allo stile di vita dei pazienti, come l'abitudine al fumo, il consumo di alcol, etc . . . [11, 12].

La mancanza di tali informazioni costituisce una delle principali barriere da superare in quanto, da un lato può essere causa di un'incorretta misura delle variabili di interesse (*errore di misura*) che può portare ad un'errata classificazione degli individui in base all'esposizione al farmaco di interesse o allo stato di malattia (misclassificazione), dall'altro può rendere problematica la fase relativa al controllo del confondimento (con particolare riferimento al *confondimento non misurato*). Al contrario, malgrado i DB dei medici di medicina generale siano più ricchi di informazioni circa il profilo clinico e lo stile di vita dei pazienti, spesso sono poveri di informazioni riguardo a visite specialistiche e ricoveri ospedalieri [13, 14].

Alla luce di queste considerazioni, per poter rispondere in modo esaustivo ai diversi quesiti della farmacoepidemiologia, sarebbe opportuno utiliz-

zare un'integrazione delle informazioni provenienti da ciascun DB e correggere le distorsioni derivanti dalle limitazioni sopracitate attraverso strategie adeguate.

1.2 Il confondimento non misurato

Si genera confondimento ogniqualvolta l'effetto dell'esposizione di interesse si combina con l'effetto di altre variabili (C) che sono associate con l'outcome in studio (O) e con l'esposizione stessa (E) [15]. La situazione descritta è rappresentata in Figura 1.1.

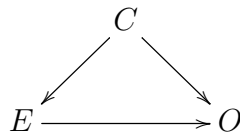


Figura 1.1: *Rappresentazione grafica della presenza di confondimento (C).*

Identificare e tenere sotto controllo il confondimento, nell'ambito della ricerca epidemiologica, è cruciale per verificare se l'associazione osservata è, di fatto, causale. Tale distorsione potrebbe rendere l'associazione osservata superiore o inferiore a quella reale, o potrebbe addirittura invertirne il verso.

Metodi statistici per l'aggiustamento del confondimento sono ormai entrati nella routine della pratica scientifica di epidemiologi e ricercatori. Restrizione, appaiamento, stratificazione o aggiustamento tramite modelli statistici multivariati, rappresentano ottime strategie per tenere sotto controllo fattori che potrebbero confondere o addirittura modificare l'associazione di interesse. Tali metodi tuttavia, sono validi sotto l'assunto che non ci siano confondenti *non misurati*, ovvero fattori, associati sia con l'esposizione che con l'outcome, non inclusi nell'analisi poiché non disponibili (Figura 1.2).

Come già visto nella Sezione 1.1 gli studi condotti tramite DB amministrativi sono caratterizzati dalla carenza di informazioni di tipo clinico. Variabili quali l'abitudine al fumo, il consumo di alcol, la pressione sanguigna, ecc..., possono giocare un ruolo fondamentale quando si parla di confondimento non misurato, nel contesto dei DB amministrativi, in quanto, in molte situazioni risultano associate sia con l'esposizione che con l'outcome. Anche aggiustando o tenendo sotto controllo tutti i fattori misurati dal potenziale

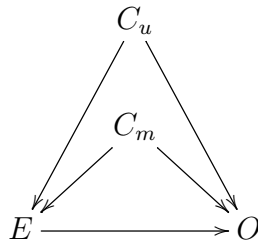


Figura 1.2: Rappresentazione grafica della presenza di confondimento non misurato (C_u). C_m rappresenta un generico confondente misurato.

effetto confondente, quali ad esempio il genere, l'età, l'utilizzo concomitante di farmaci, ecc.. il rischio di possibili distorsioni derivanti dalla presenza di confondimento residuo permane.

L'aggiustamento per il confondimento misurato, infatti, non elimina il confondimento non misurato.

Per valutare la robustezza delle stime ottenute da una prima analisi, sarebbe opportuno condurre un'analisi della sensibilità per quantificare l'impatto del confondimento non misurato (sempre ammesso che sia presente). Malgrado la letteratura sul potenziale effetto del confondimento non misurato si sia notevolmente arricchita nell'ultimo decennio, sono ancora molti gli autori che decidono di affidarsi alle stime dell'analisi principale senza effettuare ulteriori controlli sul potenziale effetto del confondimento non misurato.

I risultati di studi condotti tramite DB amministrativi, se non dotati di una valutazione del confondimento non misurato, pertanto, sono da interpretare con cautela e obiettiva criticità.

1.3 L'errore di misura

Misurare accuratamente l'esposizione, l'insorgenza dell'outcome di interesse così come le covariate è necessario per verificare la relazione causale tra l'esposizione e l'outcome. Tuttavia, in quasi tutti gli studi epidemiologici esiste la possibilità di incorrere in qualche errore di misura. Se la variabile misurata con errore è di natura dicotoma o categorica, si parla di *misclassificazione*.

Gli studi condotti tramite database amministrativi sono vulnerabili, oltre alla presenza di confondimento, anche a questo tipo di distorsione.

L'errore di misura può essere generato dalla presenza di errori di battitura, ad esempio nella data di una prescrizione/ricovero, ecc., oppure dalla mancanza di specifiche variabili cliniche che misurano o definiscono in modo corretto una certa variabile o, come vedremo nelle sezioni successive, dall'impossibilità di identificare una certa caratteristica del paziente in determinati periodi temporali.

Errori da misclassificazione sono, purtroppo, ampiamente diffusi nella ricerca epidemiologica e farmacoepidemiologica.

L'impatto che tali errori possono avere sulle stime degli effetti di interesse, tuttavia, ha ricevuto poca attenzione. Infatti è comunemente pensato che l'effetto dell'errore di misura sposti le stime verso il valore di mancanza di associazione [16]. Purtroppo, spesso questo non accade.

E' dunque preferibile valutare con attenzione se i risultati possano essere verosimilmente affetti da errore di misura, e cercare di applicare valide strategie per minimizzare tale distorsione.

1.4 Obiettivo

L'obiettivo di questa Tesi di Dottorato è quello di (i) approfondire le diverse tecniche per il controllo del confondimento non misurato e dell'errore di misura, nell'ambito degli studi condotti tramite database amministrativi, e (ii) applicare tali metodologie a diversi contesti della farmacoepidemiologia.

Parte II

Metodi per il controllo del
confondimento non misurato

Capitolo 2

Quando agire?

Il controllo del confondimento, sia misurato che non, può avvenire in diverse fasi dello studio:

- (i) in fase di progettazione, prima che lo studio inizi, attraverso un opportuno disegno
- (ii) in fase di analisi, durante lo svolgimento dello studio, attraverso opportune tecniche analitiche.

La Figura 2.1, versione leggermente rivista di quella originale proposta da Schneeweiss [17] rappresenta le tecniche utilizzate in ciascuna fase.

Agire in fase di progettazione, significa restringere l'analisi a pazienti con determinate caratteristiche (gli utilizzatori di un farmaco di interesse o i pazienti che sviluppano l'outcome in studio) e adottare opportuni disegni di studio.

Il controllo tramite tecniche analitiche invece, non comporta l'adozione di un disegno particolare o la restrizione ad un gruppo specifico di pazienti, piuttosto l'utilizzo di tecniche più sofisticate nell'analisi dei dati, da affiancare ai metodi tradizionali.

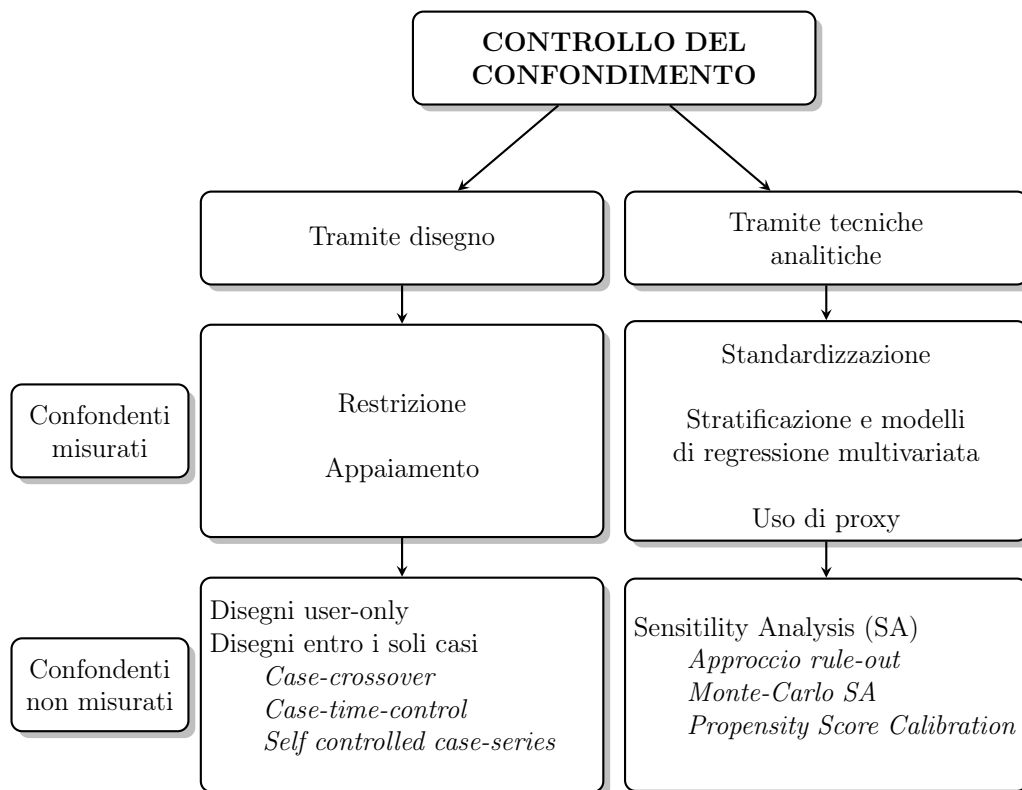


Figura 2.1: Strategie per il controllo del confondimento: è possibile agire a livello del disegno dello studio oppure in fase di analisi.

Capitolo 3

Controllo del confondimento tramite il disegno dello studio

3.1 User-only design

Una delle più ostiche tipologie di confondimento non misurato è sicuramente il *confondimento da indicazione*.

Il confondimento da indicazione si genera quando l'indicazione di un trattamento è, essa stessa, un confondente [18]. Purtroppo, negli studi osservazionali, l'indicazione raramente può essere misurata, a meno che non si disponga di informazioni cliniche dettagliate. Gli studi condotti tramite database amministrativi, come si è già visto in precedenza, sono carenti di questo tipo di informazioni e dunque, particolarmente vulnerabili a questo tipo di distorsione.

Lo scopo di tutti gli studi di farmacoepidemiologia è lo studio degli effetti (benefici o meno) di un farmaco su uno specifico outcome.

Negli studi clinici randomizzati, lo studio degli effetti di un farmaco avviene confrontando il gruppo di pazienti esposti al farmaco di interesse con il gruppo dei non esposti. Tuttavia, negli studi osservazionali, utilizzare tale gruppo di confronto, rende lo studio particolarmente suscettibile al confondimento da indicazione. Infatti, i soggetti a cui viene prescritto il farmaco hanno

un'indicazione medica (ovvero, la malattia per la quale il farmaco è stato prescritto), che manca nella maggior parte dei non utilizzatori. In queste condizioni, se la malattia per cui il farmaco viene prescritto fosse un fattore di rischio per l'outcome in studio, gli esposti avrebbero artificialmente un maggior rischio di base di sviluppare l'outcome.

La scelta del gruppo di pazienti che, per le loro caratteristiche specifiche, può essere utilizzato come gruppo di confronto diventa quindi cruciale per minimizzare i problemi legati al confondimento da indicazione.

Uno dei modi più semplici per minimizzare questa importante fonte di distorsione, è quello di restringere l'analisi ai soli pazienti esposti al farmaco di interesse, ottenendo così un disegno *user-only* [19].

Restringere l'analisi ai soli utilizzatori permette di confrontare tra loro soggetti con lo stesso rischio di base di sperimentare l'outcome. Ovviamente in questo particolare caso, anziché confrontare gli esposti con i non esposti, si sceglie di confrontare gli esposti *per molto tempo* con gli esposti *per breve tempo*, oppure gli esposti *subito prima dell'outcome* con gli esposti *molto tempo prima* dell'outcome, ecc ...

Potenzialità e limiti dei disegni user-only

I disegni user-only rappresentano una valida strategia per tenere sotto controllo il confondimento da indicazione, particolare distorsione che si colloca nell'ambito del confondimento non misurato.

Accanto alla semplicità dell'approccio, esistono ovviamente alcuni limiti. In particolare, attraverso tali disegni è possibile confrontare il rischio di insorgenza dell'outcome in studio solamente tra chi è esposto per molto tempo al farmaco, rispetto a chi è stato esposto per un periodo inferiore.

Inoltre, restringendo l'analisi ai soli pazienti che utilizzano il farmaco, la potenza dello studio potrebbe risentirne.

3.2 Case-only design

Nel condurre un studio caso-controllo, la selezione dei controlli è generalmente il compito più impegnativo. Il principio fondamentale è che i controlli selezionati debbano essere rappresentativi della popolazione che ha dato luogo ai casi [20], un principio di difficile attuazione nella pratica. Un metodo semplice per "aggirare" questo problema è quello di utilizzare i casi come controlli di sè stessi. I disegni entro i soli casi hanno suscitato un interesse sempre maggiore in quanto si basano sull'idea che, mantenendo costanti nel tempo i fattori di rischio, questi non possano confondere l'associazione di interesse. Un ulteriore vantaggio, riguarda il fatto che i disegni case-only comportano anche un carico computazionale decisamente inferiore a quello dei classici disegni utilizzati in farmacoepidemiologia.

Tra i più utilizzati in questo campo, il disegno *case-crossover*, il disegno *case-time-control* e il disegno *self-controlled case-series* sono considerati i disegni "emergenti" e anche quelli maggiormente promettenti [21]. Tuttavia, malgrado le potenzialità di questi approcci, è importante sottolineare che la validità dei risultati da essi prodotti dipende da alcune importanti assunzioni, che spesso, non vengono soddisfatte. È dunque necessario valutare con cautela se lo specifico disegno può adattarsi al contesto di interesse e che tutte le assunzioni vengano soddisfatte.

Alcune soluzioni al problema della violazione degli assunti sono state proposte in letteratura, come vedremo nelle sezioni successive.

Ogniqualevolta si conduce uno studio case-only, risulta di fondamentale importanza confrontare i risultati ottenuti, con quelli prodotti da uno dei disegni tradizionali (coorte o caso-controllo), sia per interpretare con maggiore solidità i risultati, sia per aiutare il ricercatore a scovare, se ce ne sono, i difetti metodologici dell'approccio utilizzato.

Conclusioni concordanti, ottenute da approcci statistici differenti, rinforzerebbero i risultati ottenuti, mentre conclusioni discordanti potrebbero suggerire la presenza di una distorsione in uno (o in più) degli approcci seguiti.

3.2.1 Case-crossover

Il disegno case-crossover è stato originariamente proposto da Maclure [22] nel 1991, per studiare l'effetto dell'attività sessuale sul rischio di infarto acuto del miocardio, con l'obiettivo di eliminare la distorsione da selezione e il confondimento dovuto a caratteristiche entro-soggetti.

Il termine *crossover*, solitamente utilizzato in ambito sperimentale per descrivere studi nei quali i soggetti passano dal trattamento in studio al placebo, assume in questo specifico contesto un significato leggermente diverso. Infatti, Maclure utilizza il termine *crossover* per indicare studi retrospettivi (che non prevedono randomizzazione) in cui l'esposizione è di tipo intermittente e con effetti acuti.

Da un altro punto di vista, il disegno case-crossover può essere visto come un disegno caso-controllo con appaiamento, nel quale i controlli sono i casi stessi valutati in un periodo antecedente l'outcome.

L'implementazione di un disegno case-crossover è molto semplice.

Seguendo il razionale degli studi retrospettivi, tale disegno si basa sull'identificazione dell'outcome di interesse e sulla valutazione dell'esposizione a un fattore di rischio in un periodo di tempo immediatamente precedente l'outcome (periodo a rischio). Uno o più periodi di controllo vengono successivamente selezionati (periodi di controllo) e l'esposizione al fattore di rischio nel periodo a rischio viene confrontata con l'esposizione durante i periodi di controllo. L'aspetto innovativo nasce dal fatto che tale confronto viene effettuato all'interno di ciascun individuo (caso), a differenza di ciò che avviene in tutti gli studi convenzionali.

Nel caso di un solo periodo di controllo (Figura 3.1), il disegno può essere visto come un semplice caso-controllo con appaiamento 1:1. Così come negli studi caso-controllo con appaiamento, solamente le coppie "discordanti" contribuiscono alla stima dell'odds ratio (OR), anche in questo caso solo i casi con esposizione "discordante" nelle finestre a confronto contribuiscono alla stima, ovvero i casi che risultano esposti nel periodo a rischio ma non nel periodo di controllo e viceversa. È proprio per questo motivo che un'effi-



Figura 3.1: Rappresentazione del disegno case-crossover. L'esposizione durante un periodo precedente l'outcome (periodo a rischio) viene confrontata con il rischio di esposizione in un periodo di controllo.

ciente applicazione del disegno case-crossover si ottiene quando l'esposizione è di natura transitoria e non costante.

Nel caso invece di più periodi di controllo (Figura 3.2), il disegno può essere visto come un caso-controllo con appaiamento 1:k. Aumentare il numero di

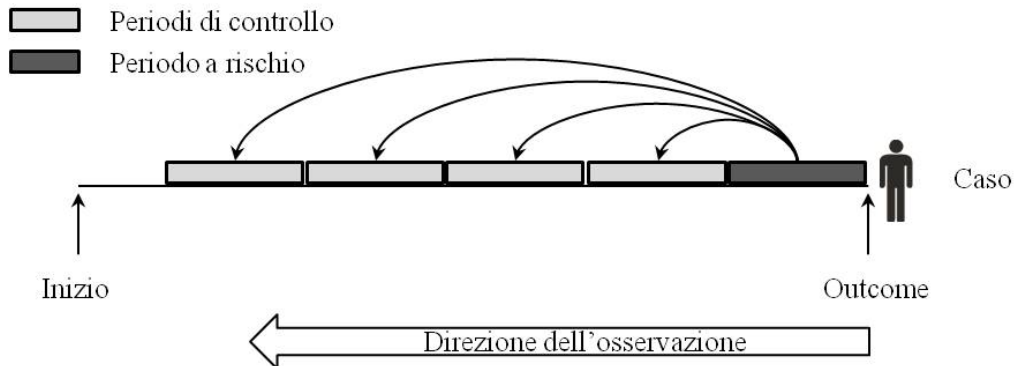


Figura 3.2: Rappresentazione del disegno case-crossover nel caso di più finestre di controllo.

finestre di controllo, può incrementare la potenza dello studio e, di conseguenza, la precisione delle stime. Mittleman, Maclure e Robins hanno infatti mostrato che la precisione delle stime ottenute da un disegno case-crossover dipende fortemente dal numero dei periodi di controllo utilizzati, ma non solo [23]. La formula (3.1) mostra il calcolo dell'efficienza relativa per uno studio

case-crossover:

$$Var(\log(OR)) = \left[\sum_{i=1}^M \frac{T_i i OR (M - i + 1)}{(i OR + M - i + 1)^2} \right]^{-1} \quad (3.1)$$

dove OR è l'OR ottenuto dallo studio, M è il numero di periodi di controllo utilizzati e T_i è il numero di sets con esattamente i soggetti esposti. Per ulteriori dettagli si veda [24].

Variazioni assunzioni sottostanti al disegno

Sebbene inizialmente proposto per studiare esposizioni transitorie che hanno effetti immediati o acuti, il disegno case-crossover è stato utilizzato da alcuni autori per studiare esposizioni prolungate o outcome con periodi di induzione più lunghi. Per adattare tale disegno a queste situazioni, basta semplicemente aumentare l'ampiezza delle finestre dell'esposizione.

In questo modo però, vi è il rischio di incrementare anche la vulnerabilità a confondenti che variano nel tempo, come ad esempio le comorbidità [25].

Confronto tra caso-controllo e case-crossover

E' possibile che vi siano delle discrepanze tra le stime ottenute da uno studio caso-controllo e un disegno case-crossover. Vediamo, in termini di distorsioni, quali possono esserne le cause.

Confondimento

Grazie a quello che viene chiamato il "self-matching" il disegno case-crossover elimina quella parte del confondimento dovuta a caratteristiche costanti nel tempo. Questo non è possibile con un semplice disegno caso-controllo, nel quale il controllo del confondimento avviene tramite un'accurata identificazione dei possibili fattori confondenti, alcuni dei quali, purtroppo non possono essere misurati. Alcuni di questi, come ad esempio la severità della malattia possono generare confondimento da indicazione, distorsione molto difficile da tenere sotto controllo, ma che con il disegno case-crossover è automaticamente tenuta sotto controllo. Tuttavia, è bene considerare che il disegno

case-crossover non corregge per il confondimento dovuto a fattori variabili nel tempo, ma è comunque sempre possibile effettuare un'analisi stratificata o correggere le stime inserendo nel modello tali fattori (valutati in opportune finestre temporali).

Misclassificazione

In uno studio case-crossover, le finestre a rischio e quelle di controllo vengono specificate a priori. Una scelta non adeguata delle finestre potrebbe riflettersi in una misclassificazione dell'esposizione.

Distorsione da selezione

Uno dei principali problemi del disegno caso-controllo riguarda la selezione dei controlli, i quali, se non selezionati accuratamente, possono portare a stime distorte. Il disegno case-crossover, supera questo problema, evitando la scelta di un gruppo di controllo.

Overmatching

L'overmatching si genera quando una variabile di appaiamento è associata con l'esposizione ma non con l'outcome (e dunque non è un vero confondente) oppure quando un fattore causato dall'esposizione è associato con l'outcome. In un disegno case-crossover, se uno dei fattori costanti nel tempo è associato solamente con l'esposizione, potrebbe esserci una riduzione di efficienza, dal momento che il disegno appaia automaticamente per quella e per altre variabili.

La seconda situazione invece, è difficile che si verifichi in uno studio case-crossover.

Potenzialità e limiti del disegno case-crossover

Riassumendo, le principali potenzialità del disegno case-crossover riguardano innanzitutto la possibilità di studiare esposizioni transitorie attraverso un buon controllo del confondimento entro-soggetto dovuto a caratteristiche costanti nel tempo.

Inoltre, presenta il vantaggio di evitare la selezione dei controlli che potrebbe generare selection bias.

Nonostante tali vantaggi, è bene tenere presente che esistono anche degli svantaggi, quali ad esempio il rischio di non controllare il confondimento entro-soggetto, qualora siano presenti dei fattori non costanti nel tempo o il rischio di introdurre distorsioni causate dalla scelta della (delle) finestra (finestre) di controllo.

Infine, se la finestra di controllo viene scelta in un periodo molto distante da quello della finestra a rischio (ad es. 1 anno o più) vi è la possibilità di introdurre una distorsione generata da un *trend temporale* dell'esposizione. In tale periodo infatti, potrebbero esserci stati dei cambiamenti nella pratica clinica, nell'utilizzo e nel commercio del farmaco in studio. Tale cambiamento, potrebbe dipendere dalla diffusione di altri farmaci, da una migliore conoscenza degli effetti benefici/avversi del farmaco, da un'intensa attività di marketing per quel farmaco, ecc...

In queste situazioni, il disegno case-crossover potrebbe non rappresentare la scelta più adatta.

3.2.2 Case-time-control

Il case-time-control, proposto da Suissa [26] nel 1995, è stato sviluppato per eliminare la distorsione dovuta alla presenza di trend nell'esposizione, maggior limite del disegno case-crossover.

Infatti, Suissa ha mostrato che gli ORs ottenuti da un case-crossover possono rappresentare, in parte, il naturale cambiamento nell'utilizzo del farmaco nel tempo e non solo l'incremento associato con l'insorgenza dell'outcome.

Come si è visto nella precedente sezione, questo naturale incremento (o decremento), è normale negli studi di farmacoepidemiologia, poiché spesso riflette cambiamenti nella pratica clinica. Di conseguenza, una porzione dell'effetto di questo trend è inglobata nell'OR ottenuto da un case-crossover e dovrebbe essere rimossa.

Ecco il motivo per il quale, nel disegno case-time-control, vengono reintrodotti i controlli, selezionati con le classiche procedure di un caso-controllo.

Sia i casi che i controlli vengono considerati due volte, una volta per la valutazione dell'esposizione nel periodo a rischio e una volta per la valutazione dell'esposizione nel periodo di controllo (Figura 3.3).

L'OR ottenuto da un case-crossover applicato al gruppo dei controlli (valutati nelle stesse finestre temporali dei casi), riflette solamente l'effetto del trend temporale (ammesso che ci sia) e viene dunque utilizzato per correggere quello ottenuto dai casi che si suppone sia affetto dal trend.

Assunzioni sottostanti al disegno

Il disegno case-time-control si basa su due assunti principali:

- la misura di associazione (OR) ottenuta dal disegno case-crossover è il prodotto di un odds ratio dovuto all'effetto causale dell'esposizione sull'outcome e un odds ratio dovuto al trend temporale nell'esposizione;
- il trend temporale nell'esposizione è lo stesso per casi e controlli.

L'OR di un case-time-control (OR_{ctc}) si calcola semplicemente dividendo l'OR di un case-crossover (OR_{cc}) per l'OR dovuto al trend temporale (otte-

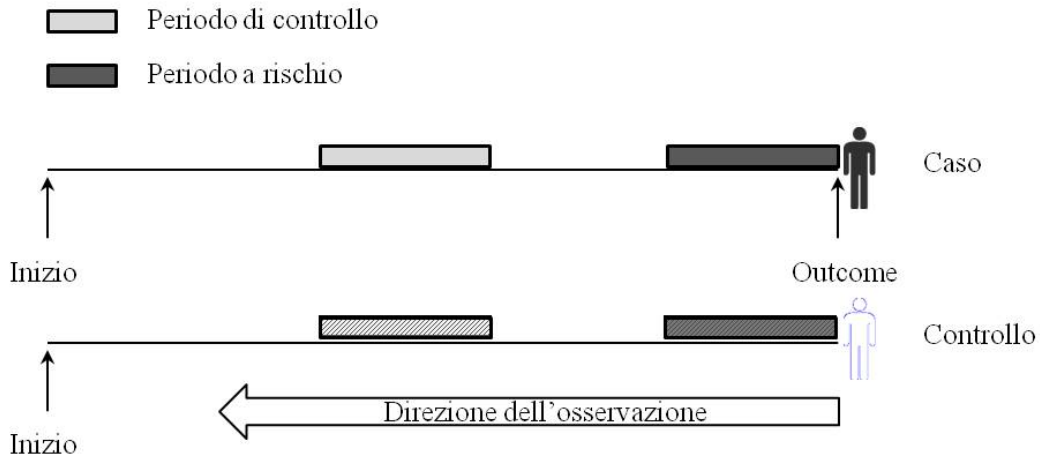


Figura 3.3: Rappresentazione del disegno case-time-control. L'esposizione durante un periodo precedente l'outcome (periodo a rischio) viene confrontato con il rischio di esposizione in uno o più periodi di controllo in casi e controlli.

nuto dai controlli, OR_{contr}):

$$OR_{ctc} = \frac{OR_{cc}}{OR_{contr}}.$$

Il modello case-time-control

Focalizzandosi sul modello basato sul logit dell'esposizione, sia

$$L_{ijkl} = \text{logit}[P(E_{ijkl} = 1)]$$

dove E_{ijkl} rappresenta la variabile dicotoma di esposizione per il gruppo i , il periodo j , l'outcome k e il soggetto l entro il gruppo i . In particolare, $i = 0, 1$ denota il gruppo caso/controllo (1=caso, 0=controllo); $j = 0, 1$ denota il periodo (1=a rischio, 0=di controllo), $k = 0, 1$ denota l'insorgenza dell'evento (1=evento, 0=no evento) ed $l = 0, 1, \dots, n_i$ denota i soggetti all'interno del gruppo i (n_1 caso n -esimo e n_0 controllo n -esimo). Si noti che $k = 1$ rappresenta un caso ($i=1$) e il periodo a rischio ($j=1$).

Modellando questa situazione

$$L_{ijkl} = \mu + s_{il} + \pi_j + \theta_k \quad (3.2)$$

dove μ rappresenta il logit dell'esposizione nel complesso, s_{il} l'effetto del soggetto l nel gruppo i , π_j l'effetto del periodo j e θ_k è l'effetto dell'insorgenza dell'outcome k che è il parametro di interesse. Per il gruppo dei casi, i due termini per il Modello (3.2) sono

$$L_{111l} = \mu + s_{1l} + \pi_1 + \theta_1 \quad (3.3)$$

e

$$L_{100l} = \mu + s_{1l} + \pi_0 + \theta_0 \quad (3.4)$$

mentre per il gruppo dei controlli

$$L_{010l} = \mu + s_{0l} + \pi_1 + \theta_0 \quad (3.5)$$

e

$$L_{000l} = \mu + s_{0l} + \pi_0 + \theta_0 \quad (3.6)$$

L'effetto del periodo è dato da $\delta_\pi = \pi_1 - \pi_0$, mentre il netto dell'effetto dell'esposizione sull'insorgenza dell'evento è dato da $\delta_\theta = \theta_1 - \theta_0$. I logit (3.3) e (3.4) indicano che non è possibile stimare separatamente δ_θ e δ_π ma solamente la loro somma $\delta_\theta + \delta_\pi$. D'altro canto, i logit (3.5) e (3.6) indicano che δ_π è stimabile nel gruppo di controllo.

Stima degli effetti

Per stimare gli effetti di interesse, vengono utilizzate le classiche tecniche utilizzate per i disegni caso-controllo con appaiamento. Ogni soggetto infatti, forma una coppia appaiata dai due periodi. Assumendo che l'esposizione in un periodo sia indipendente dall'esposizione nell'altro, si può utilizzare il modello di regressione logistica condizionata per stimare gli effetti di interesse. Tale modello, elimina automaticamente il parametro di disturbo relativo all'effetto del soggetto s_{il} (che comprende anche l'effetto della severità della malattia e di altri confondenti).

In accordo alle tecniche standard, anziché utilizzarla come variabile dipendente come nel modello (3.2), l'esposizione viene considerata come variabile indipendente mentre il periodo (1=a rischio, 0=di controllo) come variabile

dipendente.

I due approcci sono ovviamente identici e producono gli stessi risultati.

Sia E la variabile dicotoma di esposizione, G la variabile che indica il gruppo (1=caso, 0=controllo) e sia T l'indicatrice del periodo di esposizione. Il modello di regressione logistica per stimare gli effetti di interesse è:

$$\text{logit}[P(T = 1|E, G)] = \beta_0 + \beta_1 E + \beta_2 E * G. \quad (3.7)$$

L'OR di interesse, ovvero quello che misura il rischio associato all'esposizione è dato da

$$OR = e^{\beta_2} \quad (3.8)$$

ottenuto dall'interazione tra l'esposizione (E) e lo status del soggetto, ovvero il gruppo di appartenenza (G). Il coefficiente β_2 corrisponde esattamente alla differenza tra il Logit (3.5), (3.6) e (3.3), (3.6).

La porzione del rischio associata con l'effetto del periodo invece è data da

$$OR = e^{\beta_1} \quad (3.9)$$

che corrisponde esattamente al Logit (3.5) e (3.6) dato che è specifico per i controlli per i quali ($G = 0$). L'OR ottenuto dal solo gruppo dei casi ($G=1$),

$$OR = e^{\beta_1 + \beta_2} \quad (3.10)$$

corrispondente ai Logit (3.3) e (3.4), include l'effetto associato con il farmaco.

Pertanto, $OR = e^{\beta_2}$ è la porzione esclusivamente associata con il farmaco di interesse, depurata dall'effetto del trend.

Confronto tra case-crossover e case-time-control

La stima prodotta da un case-time-control è sistematicamente inferiore a quella ottenuta da un disegno case-crossover dal momento che viene rimosso l'effetto del trend temporale dell'esposizione. Inoltre, il disegno case-time-control produce una stima apparentemente meno precisa di quella prodotta

da un case-crossover. Dal momento che, però, la stima ottenuta da un case-crossover è affetta da trend dell'esposizione (ammesso che sia presente), le precisioni dei due metodi non sono confrontabili.

Potenzialità e limiti del disegno case-time-control

Il disegno case-time-control è stato disegnato da Suissa per rimuovere, dall'effetto dell'esposizione, la parte dovuta all'effetto del trend temporale. Tuttavia, reintroducendo i controlli, vi è ancora la possibilità di introdurre un confondimento da selezione qualora non venga selezionato un adeguato gruppo di controllo [27].

Per evitare questo problema, nel 2011 Wang ha proposto di utilizzare i "casi futuri" al posto dei controlli, dando luogo al *case-case time control* (sezione 3.4) [28].

Un ulteriore problema è stato messo in luce da Greenland, il quale ha mostrato che l'aggiustamento per il trend temporale non è universalmente valido [27].

Infatti, ci sono situazioni in cui il trend temporale dell'esposizione in casi e controlli è molto differente. Questo accade soprattutto quando l'outcome in studio è la morte. In tal caso, i casi potrebbero aver avuto un deterioramento più rapido rispetto alla popolazione generale dalla quale sono stati selezionati i controlli.

In altri termini, il trend temporale potrebbe essere a sua volta affetto da confondimento da indicazione. In queste situazioni, l'aggiustamento per il trend temporale introduce una nuova distorsione che può essere addirittura più forte di quella prodotta senza l'aggiustamento per il trend [27].

Malgrado le critiche metodologiche, Suissa ha dimostrato che gli esempi portati da Greenland a sostegno delle proprie ipotesi non potevano essere ritenuti validi, poiché violavano alcuni assunti sottostanti al modello.

In particolare, il modello case-time-control è valido solamente se sono verificati i seguenti tre assunti:

- Il confondente non misurato non deve avere un effetto modificatore sull'associazione esposizione-outcome. Per questo motivo nel modello (3.2) non è presente il termine $s_{il}\theta_k$;

- Il periodo non deve avere un effetto modificatore sull'associazione esposizione-outcome. Per questo motivo nel modello (3.2) non è presente il termine $\pi_j \theta_k$;
- Il confondente non misurato non deve avere un effetto modificatore sull'associazione esposizione-periodo. Per questo motivo nel modello (3.2) non è presente il termine $s_{il} \pi_j$.

3.2.3 Case-case-time-control

Il disegno *case-case-time-control*, proposto da Wang et al. nel 2011, è un'estensione del disegno *case-time-control*. Al posto di utilizzare i controlli, e dunque di introdurre una nuova distorsione, nel caso in cui i controlli selezionati non siano adeguatamente rappresentativi della popolazione generale, Wang ha proposto di utilizzare i "casi futuri" come gruppo di controllo, per correggere il trend dell'esposizione.

Dalla Figura 3.4 (in basso), si può osservare che, durante il periodo a rischio, il soggetto **2b** non ha ancora sperimentato l'outcome, ma lo svilupperà in seguito : è dunque idoneo per essere appaiato a un caso precedente e fungere da suo controllo.

Come si è visto nella sezione precedente, la performance del disegno *case-time-control* nella correzione del trend temporale dell'esposizione, dipende fortemente da quanto il trend nei controlli approssimi il trend dell'esposizione nei casi. Utilizzando un disegno *case-case-time-control*, ci si tutela da questa evenienza, dal momento che solamente i casi che entrano nello studio vengono analizzati.

Considerazioni principali

Come per tutti gli altri disegni *case-only*, ci sono diverse considerazioni che devono essere tenute presente nel caso in cui si decidesse di applicare questo tipo di disegno.

Innanzitutto, è importante sottolineare che il disegno *case-case-time-control*, come il disegno *case-crossover* e il disegno *case-time-control*, è adatto a studiare l'effetto di esposizioni brevi con effetto transitorio sul rischio di eventi acuti. Questo perché l'esposizione nel periodo a rischio non deve contenere alcun "effetto residuo" del rischio associato al periodo di controllo.

La scelta dell'ampiezza delle finestre, così come la distanza tra esse, diventa cruciale (così come per tutti gli altri *case-only*) e deve essere basata sulla conoscenza biologica del meccanismo d'azione del farmaco in studio. Per verificare la robustezza dei risultati ottenuti, sarebbe sempre opportuno condurre delle *sensitivity analysis* variando l'ampiezza delle finestre.

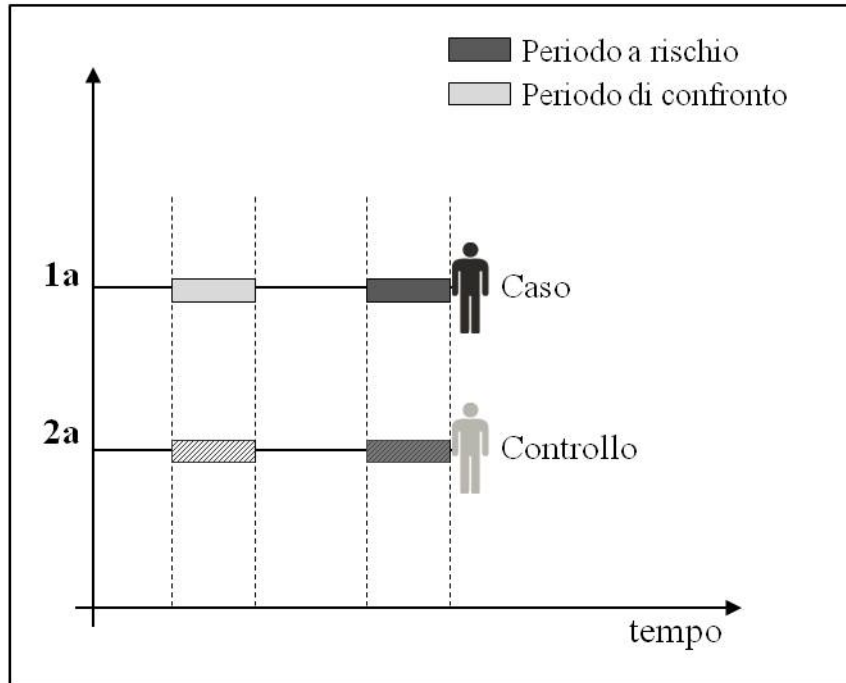
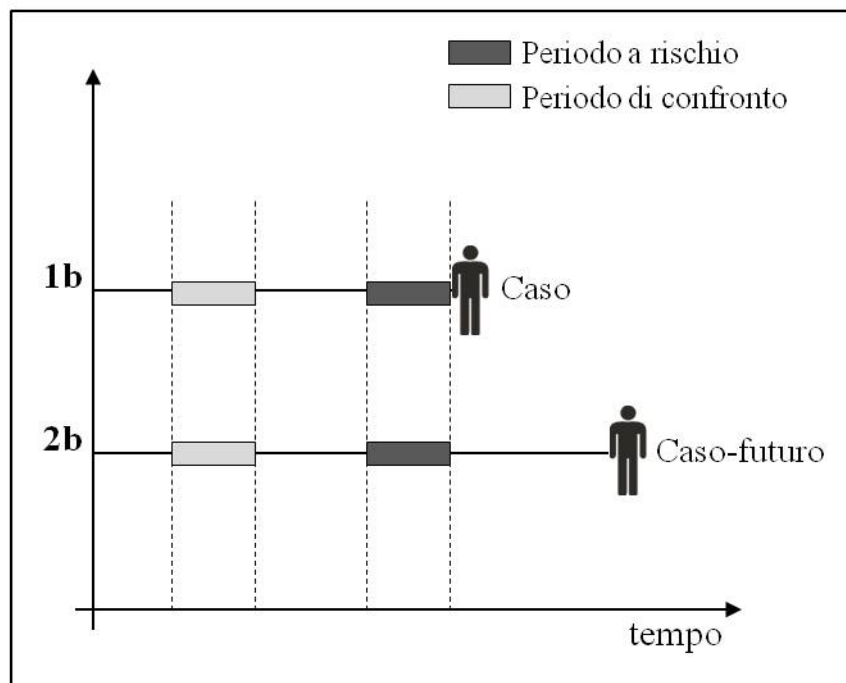
Case-time-control**Case-case-time-control**

Figura 3.4: Confronto disegni case-time-control e case-case-time-control: al posto di utilizzare i controlli per correggere il trend temporale dell'esposizione, nel case-case-time-control si utilizzano i casi futuri.

Un altro aspetto molto importante riguarda la durata del follow-up.

Il ricercatore, infatti, deve tenere in considerazione il fatto che una parte dei casi non potrà essere selezionata come controllo, poiché non più a rischio di sperimentare l'outcome. Questo diventa particolarmente problematico nelle situazioni in cui la maggior parte degli eventi si verifica verso la fine del follow-up. Per questi casi, è chiaro che ci saranno pochissimi individui idonei per il matching.

Tale caratteristica ha importanti implicazioni in termini di potenza dello studio, dato che questi casi non potranno contribuire all'analisi.

Anche la distanza tra l'insorgenza dell'outcome per il caso e l'insorgenza dell'outcome per il caso-futuro (che funge da controllo) assume un ruolo di fondamentale importanza, dal momento che l'esposizione del caso-futuro nel periodo a rischio, deve considerarsi indipendente dall'insorgenza dell'outcome.

Potenzialità e limiti del disegno case-case-time-control

Il disegno case-case-time-control è un'estensione del disegno case-time-control, sviluppato per rimuovere la distorsione da selezione, introdotta qualora i controlli scelti non siano sufficientemente rappresentativi della popolazione che ha dato luogo ai casi.

Nonostante tale potenzialità è bene tenere presente che tale disegno deve soddisfare numerose considerazioni che, se violate, potrebbero compromettere la validità dei risultati.

3.2.4 Self-controlled-case series

Il self-controlled-case series è stato sviluppato da Farrington nel 1995 per indagare l'associazione tra vaccini e potenziali eventi avversi acuti [29, 30]. Successivamente il metodo è stato applicato ad altri contesti epidemiologici. Seguendo il razionale degli studi prospettici, tale disegno di basa sull'identificazione dell'esposizione ad un fattore di rischio e sulla successiva valutazione dell'outcome in periodi temporali specifici.

In particolare, il periodo di osservazione viene partizionato in periodi di esposizione e periodi di non esposizione e successivamente la probabilità dell'outcome di interesse all'interno di un periodo di esposizione viene confrontata con la probabilità dell'outcome in un periodo di non esposizione, dove tale confronto viene effettuato all'interno dello stesso individuo (Figura 3.5) [31].

Uno dei maggiori problemi metodologici riguarda la possibilità di introdurre distorsioni derivanti da confondimento entro-soggetto, qualora vengano considerati fattori variabili nel tempo [32].

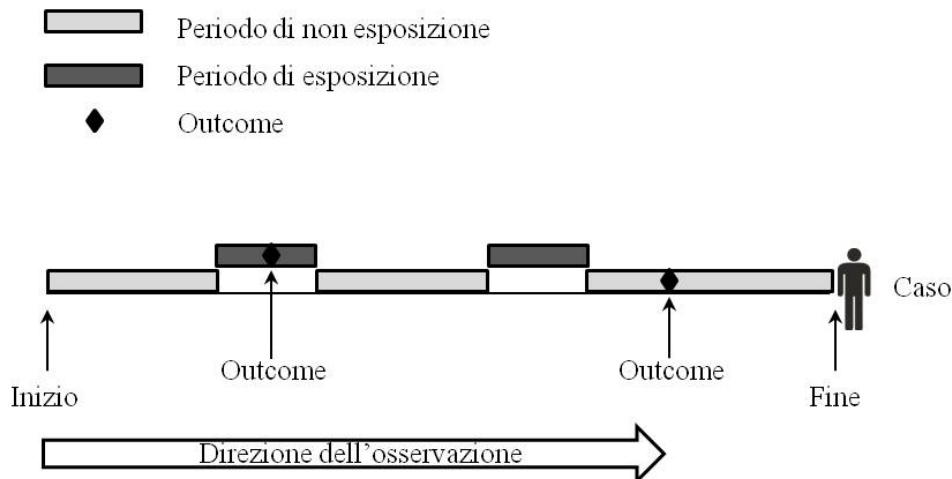


Figura 3.5: Rappresentazione del disegno self-controlled case-series. La probabilità dell'outcome di interesse che si verifica in un periodo di esposizione viene confrontata con la probabilità dell'outcome in un periodo di non esposizione.

Il modello self-controlled case-series

Il disegno self-controlled case-series può essere modellato attraverso un mo-

dello di Poisson.

Per ogni soggetto i viene definito un periodo di osservazione $(a_i, b_i]$, che viene partizionato in periodi di età j , e in k periodi a rischio (di esposizione) con $k=1,2,\dots$. Tutti gli altri periodi temporali, prima, dopo o tra periodi a rischio costituiscono i periodi di controllo, indicizzati da $k = 0$.

Siano n_{ijk} ed e_{ijk} il numero di eventi sperimentati dall'individuo i , durante il periodo di età j e il periodo a rischio k e il tempo trascorso dall'individuo i durante il periodo di età j , e il periodo a rischio k . Il tasso di incidenza, λ_{ijk} , assunto essere costante all'interno di ogni intervallo temporale, è dato dal seguente modello moltiplicativo:

$$\lambda_{ijk} = e^{s_i + \alpha_j + \beta_k} \quad (3.11)$$

dove s_i rappresenta l'effetto del soggetto i , α_j l'effetto dell'età j e β_k l'effetto del periodo a rischio k . Ponendo $\alpha_0 = 0$ e $\beta_0 = 0$, si ottiene l'incidenza al baseline $\lambda_{i00} = e^{s_i}$. Il numero di eventi n_{ijk} avvenuti durante un intervallo di lunghezza e_{ijk} è assunto essere distribuito come una v.c. Poisson con tasso $e_{ijk}\lambda_{ijk}$. Il valore di e^{β_k} rappresenta l'incidenza relativa ed è una misura dell'incidenza nel periodo a rischio k , rispetto al periodo di controllo $k = 0$.

Assumendo che l'esposizione non sia influenzata da eventi precedenti, è possibile condizionare rispetto al numero di eventi $n_i = \sum_{j,k} n_{ijk}$ osservati per il soggetto i durante l'intero periodo di osservazione. Il contributo alla verosimiglianza dell'individuo i risulta:

$$L_i(\alpha, \beta) = \prod_{j,k} \left(\frac{e_{ijk} e^{\alpha_j + \beta_k}}{\sum_{r,s} e_{irs} e^{\alpha_r + \beta_s}} \right)^{n_{ijk}}. \quad (3.12)$$

Come si può notare, l'effetto individuale s_i non contribuisce alla verosimiglianza grazie al self-matching. Pertanto, supponendo ci siano n casi, la verosimiglianza del disegno è:

$$L(\alpha, \beta) = \prod_{i=1}^n \prod_{j,k} \left(\frac{e_{ijk} e^{\alpha_j + \beta_k}}{\sum_{r,s} e_{irs} e^{\alpha_r + \beta_s}} \right)^{n_{ijk}}. \quad (3.13)$$

Assunzioni sottostanti al disegno

Affinché sia applicato correttamente, il disegno self-controlled case-series richiede che siano verificati tre assunti principali.

Assunto 1: Eventi ricorrenti sono tra loro indipendenti

In generale, il metodo self-controlled case-series è adatto a modellare eventi ricorrenti indipendenti. Tuttavia, Farrington and Hoxby nel 2010 hanno presentato un'estensione del modello nel caso in cui gli eventi (di uno stesso individuo) non siano indipendenti tra loro [33].

Assunto 2: L'insorgenza di un evento non deve alterare la probabilità di esposizione successiva

Questo è probabilmente l'assunto più restrittivo di questo metodo. Esistono tuttavia, delle varianti del modello che possono essere utilizzate nelle circostanze in cui tale assunto non possa essere verificato.

Il primo approccio viene utilizzato quando i soggetti ricevono *al massimo* una singola esposizione (come nel caso di una singola vaccinazione). In questo caso, il periodo di osservazione viene ridefinito iniziando con l'esposizione stessa. Solamente gli individui che erano esposti prima dell'evento sono inclusi nell'analisi. Lo svantaggio di tale approccio è che i casi in cui l'outcome di interesse si verifica prima dell'esposizione e i casi che non si espongono mai vengono esclusi dalle analisi, riducendo la potenza dello studio.

Il secondo approccio invece, si utilizza quando i soggetti sperimentano più di un'esposizione. In questo approccio, l'inizio e la fine del periodo di osservazione rimangono invariati, l'esposizione in seguito all'evento viene ignorata e i dati vengono analizzati utilizzando il metodo della pseudo-verosimiglianza. Per maggiori dettagli si vedano [34, 35].

Assunto 3: L'insorgenza dell'evento non deve censurare nè influenzare il periodo di osservazione

Il periodo di osservazione di ciascun soggetto deve essere indipendente dalla data di insorgenza dell'evento.

Questa assunzione può essere violata, qualora l'evento di interesse incrementi

il rischio di morte a breve-termine. Anche in questo caso, sono state messe a punto delle tecniche che consentono di applicare correttamente questo metodo [36].

Potenzialità e limiti del disegno self-controlled case-series

Il disegno self-controlled case-series combina la potenza e la semplicità di un disegno di coorte con l'economia del caso-controllo eliminando, o comunque limitando, il confondimento dovuto a fattori costanti nel tempo.

Sebbene molto intuitivo a livello teorico, a livello computazionale presenta qualche difficoltà, soprattutto se il numero di casi è elevato.

Infatti, l'implementazione di un case-series prevede la generazione di tanti record per soggetto quanti sono i periodi in cui è suddiviso il periodo di osservazione. Se il follow-up è molto ampio e un soggetto alterna periodi di esposizioni e non, è chiaro che sono necessari molti record per descrivere la storia di un solo paziente.

Capitolo 4

Controllo del confondimento tramite tecniche analitiche

4.1 Monte-Carlo Sentivity Analysis

La *Sensitivity Analysis (SA)* è stata descritta come "l'ultima linea di difesa contro le distorsioni dopo aver cercato di eliminarle, ridurle o controllarle attraverso il tipo di disegno, la raccolta e l'analisi dei dati" [37].

Ma che cos'è esattamente una SA?

Dal momento che in uno studio osservazionale, non vi è alcuna randomizzazione, l'associazione osservata tra esposizione e outcome potrebbe essere influenzata da qualche fattore non misurato associato sia con l'esposizione che con l'outcome di interesse.

La SA rappresenta un valido approccio per valutare i possibili effetti del confondimento non misurato. Infatti, consente (i) di determinare quanto forte dovrebbe essere l'effetto di un confondente non misurato per spiegare l'associazione osservata (apparente), e (ii) di aggiustare, talvolta, l'associazione osservata attraverso opportune assunzioni sulle caratteristiche del confondente.

Nelle situazioni in cui è possibile ricalcolare la stima di associazione osservata (cosa che non sempre è possibile fare), è possibile quantificare direttamente l'impatto del confondente non misurato. Infatti, se l'associazione osservata

non si discosta molto dalle stime aggiustate, allora la conclusione causale osservata diventa più "affidabile".

Ad ogni modo, anche se non è possibile aggiustare direttamente la stima di interesse, attraverso la SA è comunque possibile capire se quel risultato è plausibile o potenzialmente affetto da confondimento non misurato.

Questa idea di sensitivity analysis in studi osservazionali risale a Cornfield et al. (1959), il quale concepì una prima forma di SA per rispondere all'ipotesi di Fisher che l'apparente associazione tra il fumo di sigarette e il cancro al polmone potesse essere spiegata dall'effetto confondente di un genotipo particolarmente predisposto sia al fumo che al cancro al polmone [38]. Nello studio condotto da Cornfield et al., si osservò che un tale confondente avrebbe dovuto essere fortemente associato all'insorgenza del cancro al polmone e circa 9 volte più frequente nei fumatori che nei non fumatori, per poter spiegare l'associazione osservata.

Tale analisi non escludeva la possibilità che un confondente con una tale forza potesse esistere, ma piuttosto chiariva quello che un ricercatore dovrebbe essere pronto ad affermare per difendere le proprie conclusioni.

Per illustrare in modo semplice come funziona una SA ordinaria, si assuma un'esposizione dicotomica (E), un outcome dicotomo (O) e un fattore di confondimento non misurato C (Figura 4.5). Attraverso una SA, è possibile

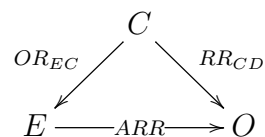


Figura 4.1: *Rappresentazione grafica della presenza di confondimento non misurato (C)*

aggiustare la misura di associazione osservata (o apparente - ARR) dividendo il rapporto tra rischi osservato per un fattore di distorsione, che rappresenta l'effetto confondente di C.

Questo fattore di distorsione (bias) è funzione della prevalenza del confondente negli esposti ad E (P_{C_1}) e nei non esposti (P_{C_0}), nonché del rapporto

tra rischi del confondente di C su D (RR_{CD}).

Nel caso più semplice di un confondente dicotomo, la stima di associazione negli esposti ad E (RR_1) e affetta da C, può essere scritta come:

$$RR_1 = (1 - P_{C_1}) + RR_{CD}P_{C_1} = P_{C_1}(RR_{CD} - 1) + 1 \quad (4.1)$$

mentre nei non esposti ad E, RR_0

$$RR_0 = (1 - P_{C_0}) + RR_{CD}P_{C_0} = P_{C_0}(RR_{CD} - 1) + 1 \quad (4.2)$$

La distorsione derivata da C è il rapporto tra RR_1 e RR_0 ,

$$bias = \frac{RR_1}{RR_0} = \frac{P_{C_1}(RR_{CD} - 1) + 1}{P_{C_0}(RR_{CD} - 1) + 1}. \quad (4.3)$$

Nel caso in cui C non sia dicotoma ma abbia j modalità, con $j=0,1,2,\dots,J$ la formula (4.3) generalizzata è la seguente:

$$bias = \frac{\sum_{j=1}^J P_{C_{j1}}(RR_j - 1) + 1}{\sum_{j=1}^J P_{C_{j0}}(RR_j - 1) + 1} \quad (4.4)$$

dove RR_j rappresenta il rischio relativo dell'outcome di interesse associato alla categoria j-esima del confondente.

La stima del rischio relativo aggiustata per l'effetto del confondente non misurato è data da

$$RR_{adj} = \frac{ARR}{bias}. \quad (4.5)$$

Per tener conto dell'incertezza relativa alle informazioni esterne utilizzate per effettuare una SA, come ad esempio l'incertezza sul vero valore di RR_{CD} , sarebbe necessario effettuare SA per diversi valori dei parametri richiesti. Tuttavia, se l'incertezza riguardo tali parametri è elevata, si dovrebbe eseguire un numero molto elevato di SA.

Un'estensione della SA classica, è stata proposta da Steenland e Greenland [39] ed è nota con il nome di Monte Carlo Sensitivity Analysis (MCSA).

Il procedimento della MCSA è il seguente.

1. La misura di interesse viene ripetutamente stimata sulla base dei dati osservati, in modo tale da tener conto della variabilità casuale:

$$ARR^1, \dots, ARR^k$$

2. Le priori per i parametri di interesse vengono ripetutamente campionate (attraverso un campionamento Monte-Carlo) e i valori estratti vengono poi utilizzati per costruire il fattore di distorsione:

$$(P_{C1}^1, P_{C0}^1, RR_{CD}^1), \dots, (P_{C1}^k, P_{C0}^k, RR_{CD}^k)$$

$$(bias^1, \dots, bias^k)$$

3. I valori campionati al punto (2) vengono dunque utilizzati per correggere l'associazione osservata:

$$RR_{adj}^1 = \frac{ARR^1}{bias^1}, \dots, RR_{adj}^k = \frac{ARR^k}{bias^k}$$

Da questa fase finale, si ottiene un'approssimazione della distribuzione della misura di associazione corretta per la presenza di un confondente non misurato.

Tale distribuzione può essere sintetizzata in modo da ottenere un'unica stima puntuale e intervallare dell'associazione di interesse aggiustata per l'effetto di un singolo confondente non misurato [39].

Potenzialità e limiti della MCSA

La MCSA è un'ottima strategia che consente di correggere le stime di associazione osservate, distorte a causa dell'effetto del confondimento non misurato. Questo metodo, abbastanza semplice da applicare, è adatto nelle situazioni in cui il confondente non misurato, che si suppone abbia distorto i risultati ottenuti, può essere rappresentato da una variabile con 2 o più modalità.

La difficoltà risiede nella capacità del ricercatore di identificare le prevalenze di ogni livello del confondente negli esposti e nei non esposti. Tali prevalenze

possono essere ricavate da studi di validazione condotti in un sotto-campione dello studio principale che contengono l'informazione sul confondente, oppure, nella maggior parte dei casi, possono essere ricavate dalla letteratura scientifica. Quest'ultima, rappresenta sicuramente la scelta più semplice, dal momento che non sempre si dispone delle risorse necessarie per condurre uno studio di validazione e che in letteratura sono disponibili numerosi studi. Accanto all'apparente semplicità, tuttavia, si cela la maggiore difficoltà che deriva dalla necessità di trovare uno studio il più possibile simile a quello condotto dal ricercatore, sia per quanto riguarda la popolazione utilizzata che la definizione dell'esposizione e delle sue categorie.

Si pensi, per esempio, di aver condotto uno studio user-only (sezione 3.1) tramite database amministrativi per valutare l'effetto degli antidepressivi sul rischio di infarto. Si assuma inoltre, di aver categorizzato l'esposizione ad antidepressivi nel modo seguente: durata < 12 mesi e durata ≥ 12 mesi. Supponendo il fumo, un possibile confondente non misurato, tutto quello che serve per poter applicare la MCSA (senza uno studio di validazione) è la stima di RR_{CD} , ovvero la stima del rischio di infarto tra i fumatori rispetto ai non fumatori (quantità facilmente reperibile in letteratura), P_{C_0} , ovvero la prevalenza dei fumatori tra gli utilizzatori di antidepressivi per un periodo inferiore ad 1 anno e P_{C_1} , ossia la prevalenza dei fumatori tra gli utilizzatori di antidepressivi per un periodo superiore ad 1 anno. È chiaro che non è del tutto immediato ricavare dalla letteratura queste due prevalenze.

In queste situazioni, o si riesce a trovare uno studio effettuato su una popolazione sovrapponibile a quella di interesse oppure è possibile identificare dei range di valori plausibili per ciascuna prevalenza e, successivamente, applicare la MCSA a tali valori.

Nel caso si supponga esistano più confondenti non misurati, la MCSA consente di correggere le stime di interesse per un confondente alla volta.

4.2 Approccio Rule-out

Il metodo Rule-Out (RO), permette di valutare la forza di un confondente non misurato necessaria per spiegare completamente l'associazione osservata. Più precisamente, tale approccio consente la quantificazione della forza, espressa tramite le relazioni confondente-esposizione (OR_{EC}) e confondente-outcome (RR_{CD}), che un confondente non misurato dovrebbe avere per spostare la stima puntuale osservata (ARR) verso l'unità, ovvero il valore che rappresenta l'assenza di associazione [40].

Per illustrare il metodo, si assuma un'esposizione dicotomica (E), un outcome dicotomo (D) e un fattore di confondimento non misurato C (Figura 4.5). L'associazione tra C ed E può essere misurata attraverso l'odds ratio OR_{EC} , che è funzione della prevalenza di C tra gli esposti ad E (P_{C_1}) e delle probabilità marginali di esposizione di C ed E nella popolazione in studio (P_C e P_E).

$$OR_{EC} = \frac{P_{C_1}(1 - P_C - P_E + P_{C_1})}{(P_C - P_{C_1})(P_E - P_{C_1})} \quad (4.6)$$

Supponendo che non ci sia associazione tra l'esposizione in studio e l'outcome (ovvero $RR_{ED}=1$), Walker [18] ha mostrato che l'ARR è funzione di P_{C_1} , P_E e P_C , RR_{CD} :

$$ARR = \frac{P_{C_1}(RR_{CD} - 1) + P_E}{(P_C - P_{C_1})(RR_{CD} - 1) - P_E + 1} \cdot \frac{1 - P_E}{P_E} \quad (4.7)$$

Seguendo l'esempio fornito da Psaty et al. [41] è possibile rappresentare tutte le combinazioni dei parametri OR_{EC} e RR_{CD} che dovrebbero generare confondimento da parte di un fattore non misurato abbastanza forte da spostare la stima puntuale di ARR verso l'unità (ARR=1), data una fissata prevalenza P_C .

A livello pratico, per costruire la curva RO, è necessario fissare il valore di P_E , P_C e RR_{CD} (quantità che si possono ricavare dalla letteratura) e ricavare P_{C_1} dalla formula (4.7), ponendo ARR=1. Una volta ricavata P_{C_1} , la si può inserire nella formula (4.6) per ricavare OR_{EC} .

È possibile rappresentare graficamente queste combinazioni di parametri, ottenendo una curva simile a quella in Figura 4.2. I punti sulla curva rappre-

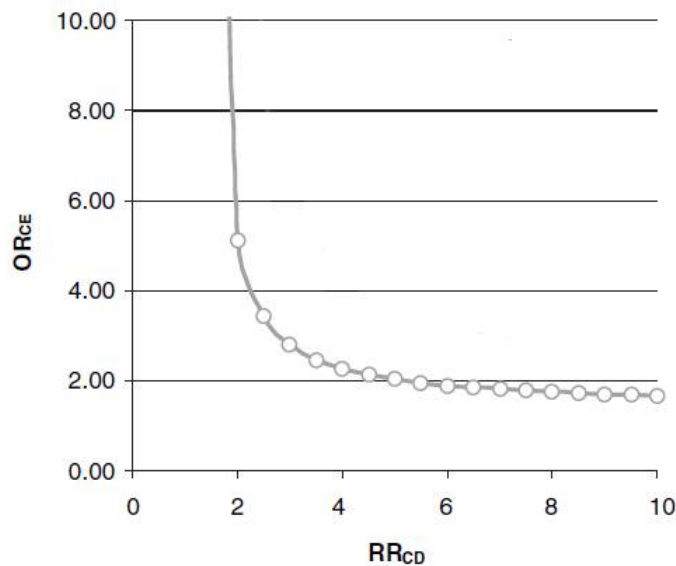


Figura 4.2: Esempio di curva rule-out. La linea divide l'area in due: la zona in alto a destra rappresenta tutte le combinazioni di parametri di OR_{EC} e RR_{CD} che dovrebbero creare confondimento da parte di un fattore non misurato abbastanza forte da spostare la stima puntuale del RR apparente (ARR) verso l'unità ($ARR=1$) o addirittura verso valori inferiori a 1. Al contrario, la zona in basso a sinistra rappresenta tutte le combinazioni di parametri che non sarebbero in grado di spostare l' ARR verso l'unità.

sentano le combinazione dei parametri OR_{EC} e RR_{CD} che annullano l'associazione osservata tra l'esposizione e l'outcome (ARR). La zona in alto a destra rappresenta tutte le combinazioni dei parametri che dovrebbero creare confondimento da parte di un fattore non misurato abbastanza forte da spostare la stima puntuale di ARR verso l'unità ($ARR=1$) o addirittura verso valori inferiori a 1. Al contrario, la zona in basso a sinistra rappresenta tutte le combinazioni di parametri che non sarebbero in grado di spostare l' ARR verso l'unità.

Potenzialità e limiti del rule-out

Al contrario della MCSA, il RO non consente di aggiustare direttamente la stima di interesse, ma solamente di quantificare la forza necessaria che un confondente dovrebbe avere per annullare l'associazione di interesse. Un limite di questo approccio riguarda il fatto che è possibile considerare solamente confondenti binari [37].

4.3 Propensity Score Calibration

Il Propensity Score Calibration (PSC), proposto da Sturmer et al. nel 2005, consente di aggiustare le stime di interesse potenzialmente affette da più di un confondente non misurato [42]. Questo metodo incorpora le informazioni provenienti da uno studio di validazione per calibrare il Propensity Score (PS), utilizzato per la correzione del confondimento.

Il Propensity Score (PS)

Il PS rappresenta la probabilità di esporsi al farmaco di interesse, data una serie di covariate [43]. Ogni soggetto è caratterizzato da un vettore di covariate $\mathbf{X} = (X_1, \dots, X_n)$ e da una variabile dicotoma che rappresenta l'esposizione E . La propensione al trattamento E per un soggetto con covariate (X_1, \dots, X_n) è data da:

$$PS(\mathbf{X}) = P(E = 1 | X_1, \dots, X_n) \quad (4.8)$$

che può essere facilmente stimata tramite un modello di regressione logistica. Ad ogni paziente viene dunque assegnata una probabilità di esposizione che riflette la propensione di ricevere il trattamento di interesse, data una serie di caratteristiche misurate. Gli individui con la stessa stima del PS hanno, in media, la stessa possibilità di ricevere il trattamento, nonostante siano caratterizzati da una diversa storia clinica o prescrittiva (le covariate).

All'interno di ogni strato del PS, vi saranno dunque sia pazienti trattati che pazienti non trattati. Questa caratteristica viene a volte proposta come una sorta di "randomizzazione virtuale" in cui pazienti comparabili (con la stessa chance di esporsi) sono suddivisi in trattati e non.

Una volta che il PS è stato stimato, può essere utilizzato in modi diversi per controllare il confondimento misurato. Infatti, può essere utilizzato come variabile di appaiamento, di aggiustamento multivariato oppure per modellare l'associazione PS-outcome.

Si assuma infatti, di considerare il seguente modello di Cox:

$$\lambda(t|E, PS(\mathbf{X})) = \lambda_0(t) \exp^{\beta_E E + \beta_X PS(\mathbf{X})} \quad (4.9)$$

Questo modello viene utilizzato per l'applicazione del PSC. In teoria, il condizionamento per la stima del PS dovrebbe bilanciare il gruppo degli esposti e quello dei non esposti, e quindi produrre delle stime non affette da confondimento, a patto che non vi siano dei confondenti non misurati.

L'utilizzo della stima del PS in un modello di Cox è descritto nel dettaglio da D'Agostino [44].

Propensity Score Calibration

Nelle situazioni in cui importanti confondenti non sono misurati, la stima del PS utilizzando le sole variabili misurate può essere vista come una variabile misurata con errore. Quello che si vorrebbe fare è riuscire a calcolare il PS come

$$PS(\mathbf{X}^*) = P(E = 1 | X_{m_1}, \dots, X_{m_n}, X_{u_1}, \dots, X_{u_k}). \quad (4.10)$$

dove il vettore di covariate $X_m = (X_{m_1}, \dots, X_{m_n})$ rappresenta le n covariate misurate, il vettore $X_u = (X_{u_1}, \dots, X_{u_k})$ le k covariate non misurate mentre il vettore X^* la totalità delle covariate, misurate e non.

Supponendo di avere a disposizione uno studio di validazione, nel quale sono presenti sia le variabili misurate nello studio principale che quelle non misurate (e dunque la totalità X^*), è possibile utilizzare il metodo della Regression Calibration che vedremo nel dettaglio nella sezione (6.2) a pagina 58.

In particolare, sia

$$PS(\mathbf{X}_m) = P(E = 1 | X_{m_1}, \dots, X_{m_n}) \quad (4.11)$$

la variabile affetta da errore mentre

$$PS(\mathbf{X}^*) = P(E = 1 | X_{m_1}, \dots, X_{m_n}, X_{u_1}, \dots, X_{u_k}) \quad (4.12)$$

la variabile misurata correttamente.

Supponendo che esista una relazione lineare tra le due variabili, è possibile rappresentare tale relazione nel seguente modo:

$$E[PS(\mathbf{X}^*) | PS(\mathbf{X}_m)] = \gamma_0 + \gamma_E E + \gamma_x PS(X_m). \quad (4.13)$$

Il modello di Cox che verrebbe implementato in uno studio di coorte, nel quale tutte le variabili sono disponibili, sarebbe

$$\lambda [t|E, PS(X^*)] = \lambda_0(t)e^{\beta_E E + \beta_X PS(X^*)}. \quad (4.14)$$

La stima dell'effetto dell'esposizione E aggiustata è

$$\hat{\beta}_E^* = \hat{\beta}_E - \frac{\hat{\gamma}_E \hat{\beta}_X}{\hat{\gamma}_X} \quad (4.15)$$

La stima della varianza tiene conto dell'incertezza addizionale causata dalla stima di γ nello studio di validazione [45].

Potenzialità e limiti del PSC

Il PSC è un metodo che trova larga applicazione negli studi condotti tramite database amministrativi. Infatti, per ottenere una buona stima del PS, è importante utilizzare il maggior numero di confondenti disponibili e tramite i database amministrativi, se ne possono identificare moltissimi, ad eccezione delle variabili di natura clinica (confondenti non misurati).

Questo metodo è stato sviluppato per studi di coorte, dal momento che la propensione al trattamento, viene calcolata al baseline per tutti i soggetti dello studio. In uno studio caso-controllo infatti, dal momento che gli individui vengono selezionati sulla base dell'outcome di interesse, sarebbe più complicato definire un periodo (uguale per tutti) per poter calcolare la propensione al trattamento.

Il PSC produce stime approssimativamente consistenti solamente se la varianza dell'errore di misura non è troppo grande ed è valido se l'errore di misura che incide sul valore di $PS(X_m)$ è di tipo non differenziale [42].

Parte III

Metodi per il controllo dell'errore di misura

Capitolo 5

Quando agire?

A differenza di ciò che avviene per il confondimento non misurato, per tenere sotto controllo l'errore di misura non esistono disegni ad hoc o particolari restrizioni che si possano pianificare in fase di progettazione.

Per il controllo dell'errore di misura è possibile agire solamente in fase di analisi, applicando alcune delle tecniche che sono rappresentate in Figura 5.1.

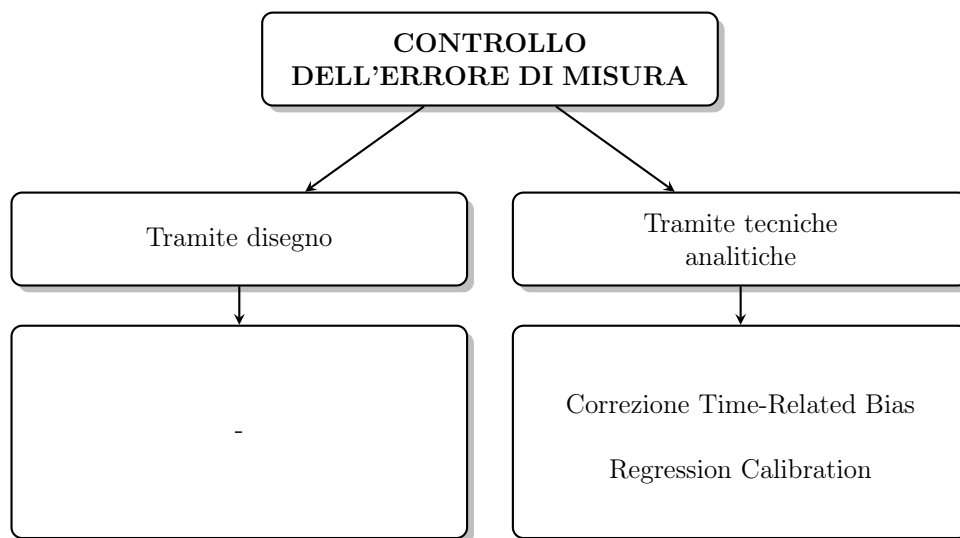


Figura 5.1: Strategie per il controllo dell'errore: al contrario di quanto avviene per il confondimento, è possibile agire solamente in fase di analisi.

Capitolo 6

Metodi per il controllo dell'errore di misura tramite tecniche analitiche

6.1 Correzione dei Time-Related Bias

La crescente disponibilità di database amministrativi, come abbiamo visto, ha portato nell'ultimo decennio ad un'esplosione di studi osservazionali che valutano l'impatto dei farmaci. Tuttavia, questi studi presentano una notevole complessità metodologica spesso sottovalutata. I *time related bias* si attestano tra le distorsioni più sottili e ricorrenti [46].

I time-related bias includono l'*immortal time bias*, una distorsione introdotta da un'analisi statica dell'esposizione negli studi di coorte, il *time-window bias*, una distorsione che affligge gli studi caso-controllo, derivante dalle finestre temporali di lunghezza diversa in cui viene valutata l'esposizione di casi e controlli e l'*immeasurable time bias* che si genera quando l'esposizione non può essere identificata a causa della presenza di periodi di ricovero.

Queste distorsioni sono note per esagerare l'effetto protettivo di un farmaco, anche quando in realtà potrebbe non avere alcun effetto.

6.1.1 Immortal time bias

L'immortal time bias, è stato identificato per la prima volta nel 1970 nel contesto degli studi di coorte sugli effetti benefici del trapianto di cuore. Questa distorsione è recentemente riapparsa in molti studi osservazionali, secondo i quali varie terapie farmacologiche risultavano estremamente efficaci nel ridurre la mortalità e l'incidenza di altri esiti sfavorevoli [47–49].

L'*immortal time* si riferisce ad un periodo di tempo durante il follow-up, durante il quale l'outcome di interesse non può verificarsi.

Questo fenomeno tipicamente nasce negli studi di coorte nei quali la definizione dell'esposizione al farmaco di interesse dipende dal momento in cui il soggetto riceve la prima prescrizione di quel farmaco. Per poter essere classificati come esposti, infatti, gli individui devono necessariamente *sopravvivere* o *essere liberi* dall'outcome in studio fino al momento della prescrizione. Il periodo tra l'ingresso nello studio e la prima prescrizione diventa quindi "immortale" (Figura 6.1), per i soggetti esposti.

A causa di questo fenomeno, i soggetti esposti presentano un vantaggio in

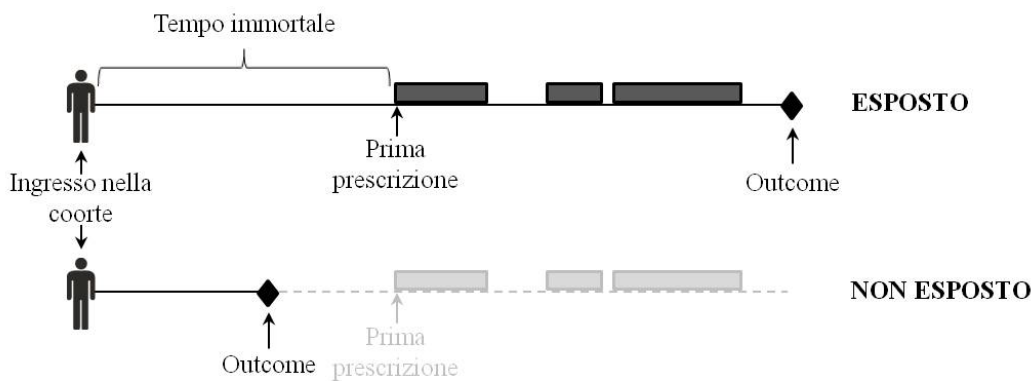


Figura 6.1: Rappresentazione del tempo immortale in uno studio di coorte in cui l'esposizione al farmaco è definita dopo l'ingresso nella coorte. Il tempo tra l'ingresso nella coorte e la prima esposizione al farmaco, per i soggetti che vengono classificati come esposti, è immortale, poiché il soggetto deve sopravvivere per ricevere il farmaco. Tale periodo, viene erroneamente classificato come esposto, quando in realtà, tale periodo non lo è.

termini di sopravvivenza rispetto ai soggetti non esposti, essendo artificialmente protetti durante il periodo immortale.

Questa distorsione genera una misclassificazione del tempo-persona in quanto il tempo immortale dei soggetti esposti viene erroneamente classificato come periodo di esposizione quando invece non lo è.

L'errata attribuzione del tempo-persona determina una sottostima del rischio dell'outcome di interesse negli esposti (il denominatore del rischio risulterà erroneamente incrementato dal tempo-persona immortale, in cui non può, per definizione, manifestarsi l'outcome) e una sovrastima del rischio nei non esposti.

Quantificazione del bias

Si consideri una coorte nella quale i soggetti vengono reclutati e classificati sulla base dello schema in Figura 6.1, ovvero, entrano nella coorte al momento dell'insorgenza di un evento specifico (ad esempio il ricovero per una patologia) e vengono successivamente classificati come esposti se, durante il follow-up, ricevono almeno una prescrizione del farmaco di interesse.

Si identifichi con $i = 0$ il gruppo dei non esposti e con $i = 1$ quello degli esposti. Siano inoltre:

T_i : Persone-tempo per il gruppo i

C_i : Numero dei casi nel gruppo i

k : Rapporto tra le persone tempo T_0/T_1

p : Proporzioni di T_1 misclassificato (periodo immortale e dunque di non esposizione).

Assumendo che il rischio dell'outcome di interesse sia costante nel tempo e che l'esposizione sia classificata correttamente, la stima non distorta del RR, RR_u è stimata da

$$RR_u = \frac{C_1}{C_0} \left[\frac{T_0 + pT_1}{T_1(1-p)} \right] = \frac{C_1}{C_0} \left(\frac{k+p}{1-p} \right) \quad (6.1)$$

mentre la stima distorta RR_b

$$RR_b = \frac{C_1 T_0}{C_0 T_1} = \frac{C_1}{C_0} k. \quad (6.2)$$

Dal rapporto tra la stima distorta e quella non distorta, è possibile quantificare la dimensione del bias, che è dato da

$$Bias = \frac{RR_b}{RR_u} = k \frac{1-p}{k+p}. \quad (6.3)$$

La Figura 6.2 mostra l'andamento della distorsione in funzione di p per tre valori di k , ovvero $k = 0.1, k = 1, k = 10$. Si può notare che la distorsione cresce rapidamente quando $k = 0.10$, ovvero quando il contributo temporale dei non esposti è decisamente inferiore a quello degli esposti. Quando $k = 10$, la distorsione ha un andamento lineare, in funzione di p . In generale, quando k è molto grande, la distorsione tende a $(1 - p)$.

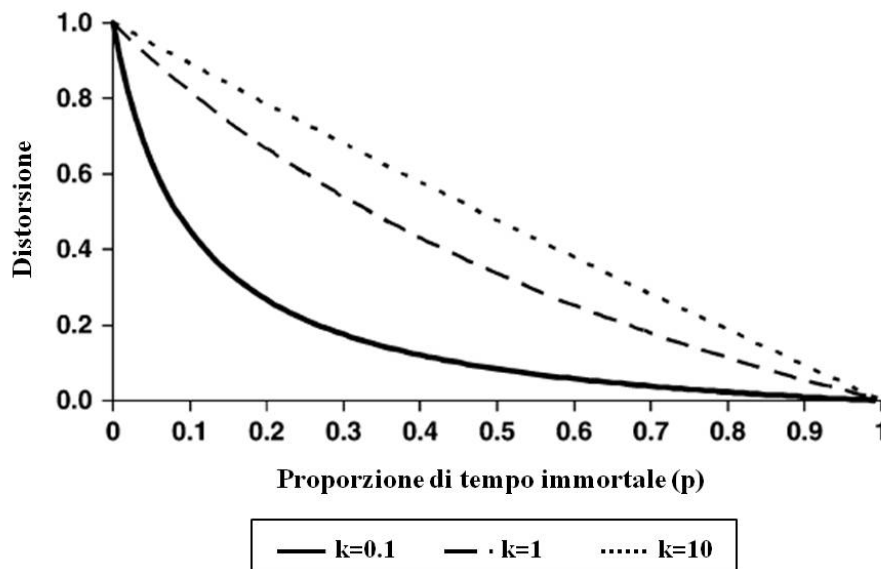


Figura 6.2: Il grafico mostra l'andamento della distorsione, in funzione della proporzione di tempo immortale misclassificato (p), per tre valori del rapporto tra persone-tempo non esposte ed esposte, $k = 0.1, 1, 10$.

Riconoscere e correggere l'immortal time bias

L'immortal time bias è una distorsione che affligge gli studi di coorte nei quali l'esposizione viene misurata in un momento qualsiasi durante il follow-up. Nelle situazioni in cui, in questo tipo di studi, emergono risultati strabilianti circa l'effetto protettivo del farmaco di interesse, è opportuno ricercare criticamente la presenza di tempo-persona immortale nel disegno dello studio e

assicurarsi che questo venga correttamente classificato come non esposto.

Il miglior approccio per tenere sotto controllo la potenziale misclassificazione derivante da questa distorsione, richiede che il periodo immortale venga classificato correttamente in termini di esposizione. Questo è possibile solamente attraverso un approccio tempo-dipendente dell'esposizione che permetta ai pazienti di essere classificati come non esposti nel periodo che va dall'inizio dell'osservazione fino alla data della prima prescrizione e come esposti dopo tale data [50].

6.1.2 Time-window bias

Per alcuni anni, i time-related bias sono stati segnalati esclusivamente negli studi di coorte e sembrava che gli studi caso-controllo fossero immuni da questa distorsione.

Nel 2007, tuttavia, è stato pubblicato uno studio caso-controllo condotto tramite database amministrativi che riportava una riduzione del rischio di tumore al polmone del 55% associata all'utilizzo di statine [51]. Considerando le implicazioni di una così elevata riduzione del rischio, Svizzera ha riveduto l'articolo per ricercare una spiegazione scientifica di tale risultato. La spiegazione di questo risultato così impressionante è dovuta, di nuovo, a un time-related bias, chiamato, *time-window bias* [52].

Il time-window bias, distorsione propria degli studi caso-controllo, viene generato dalla tecnica utilizzata per selezionare i controlli e per misurare la loro esposizione.

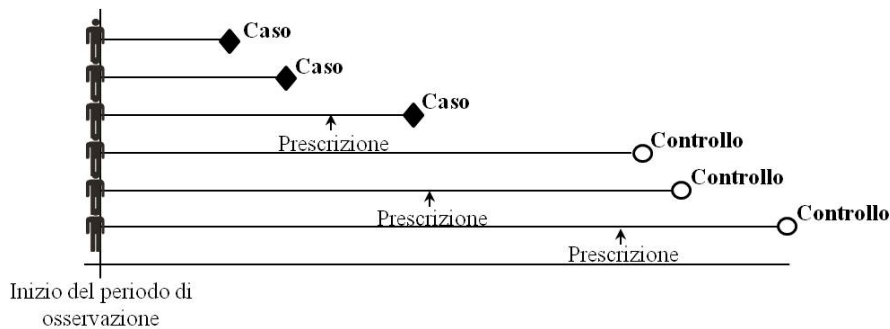


Figura 6.3: Descrizione di casi e controlli, per i quali l'esposizione è definita sulla base delle prescrizioni ricevute prima della fine dell'osservazione

In particolare, nello studio che indagava la relazione tra statine e tumore al polmone, l'esposizione è stata definita come qualsiasi prescrizione durante il periodo di osservazione. Perciò, per i casi di tumore al polmone l'esposizione a statine è stata misurata nel periodo precedente alla data di tumore, mentre per i controlli per tutto il periodo di osservazione. Di conseguenza, i casi avevano avuto meno tempo per potersi esporre al farmaco di interesse. Questo può risultare in una sovra-rappresentazione dei casi non esposti e perciò in una sottostima dell'effetto reale del farmaco (Figura 6.3).

Riconoscere e correggere il time-window time bias

Per correggere l'effetto di tale distorsione, è necessario valutare l'esposizione di casi e controlli in finestre temporali sovrapponibili.

Per fare questo, è necessario appaiare casi e controlli per lunghezza del follow-up. In questo modo la durata del periodo di osservazione risulta identica per casi e per controlli, neutralizzando così il time-window bias (*incidence density sampling*).

6.1.3 Immeasurable time bias

Come si è visto nelle sezioni precedenti, gli studi di coorte sono potenzialmente affetti da immortal time bias, nelle situazioni in cui l'esposizione viene trattata in modo statico, mentre gli studi caso-controllo possono essere influenzati da time-window bias, nei casi in cui le finestre in cui vengono confrontati casi e controlli sono molto diverse tra loro. Purtroppo, esiste un'altra distorsione da cui entrambi gli studi non possono ritenersi del tutto esenti: l'*immeasurable time bias* (ITB) [53].

Gli studi condotti tramite database amministrativi sono particolarmente suscettibili a questa distorsione, dal momento che l'esposizione viene valutata sulla base di database che contengono informazioni sui farmaci (rimborsabili dal SSN) erogati dalle farmacie territoriali.

Può succedere, che un soggetto in terapia cronica con un determinato farmaco venga ricoverato in ospedale (per motivi diversi dall'outcome in studio) e che durante tale periodo, gli venga comunque somministrato il farmaco che è solito assumere.

In questa situazione, non è possibile in alcun modo catturare correttamente l'esposizione, poiché i farmaci somministrati in ospedale non vengono registrati nei database amministrativi.

Se un soggetto ha sperimentato diversi periodi di ricovero durante il follow-up, è molto probabile che la sua esposizione venga misclassificata.

Con *immeasurable time*, infatti, ci si riferisce ad un periodo di tempo durante il follow-up, per uno studio di coorte, o prima della data di insorgenza dell'evento, per un studio caso-controllo, durante il quale l'esposizione ad un farmaco non può essere identificata.

Pertanto, l'effetto della distorsione generata da *immeasurable time* si riflette sul rischio di classificare erroneamente un soggetto come "non esposto" quando in realtà lo è. Per controllare l'effetto dell'ITB esistono diverse tecniche, quali ad esempio: (i) l'eliminazione dei giorni di ricovero dal periodo di follow-up mantenendo inalterata l'ampiezza della finestra temporale antecede-

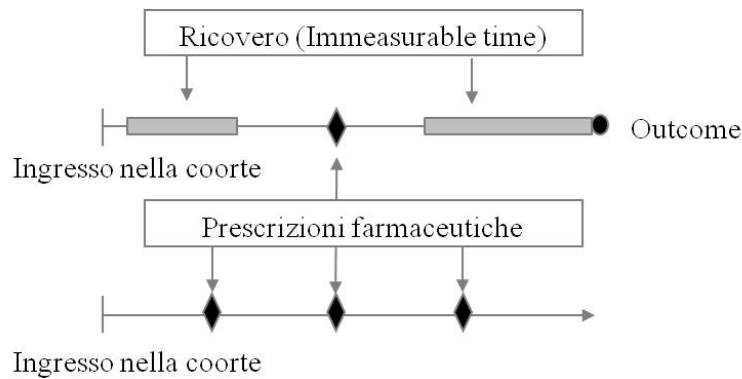


Figura 6.4: Descrizione di due soggetti partecipanti ad uno studio di coorte. I periodi di ricovero (primo soggetto) rendono impossibile la valutazione dell'esposizione ad un farmaco in tali periodi, per la mancanza di informazioni sulle prescrizioni farmaceutiche, generando immeasurable time bias. Questo soggetto sembrerà avere meno prescrizioni rispetto al primo soggetto per effetto di tale distorsione

dente all'evento (nel caso di uno studio caso-controllo); (ii) l'assunzione che i giorni di ospedalizzazione siano interamente coperti dal farmaco di interesse in presenza di una prescrizione nel periodo immediatamente precedente al ricovero; (iii) l'applicazione di alcune delle strategie proposte da Suissa [53].

In particolare, Suissa ha proposto cinque diversi metodi per affrontare il problema dell'immeasurable time bias. Per illustrare i procedimenti elencati di seguito, si supponga di condurre uno studio caso-controllo per valutare l'associazione tra l'esposizione ad un farmaco e l'insorgenza di una malattia. Si assuma inoltre che l'esposizione al farmaco di interesse venga valutata nei 30 giorni precedenti l'evento.

1. Restrizione ai soli soggetti senza ospedalizzazioni nei 30 giorni precedenti l'evento.
2. Inclusione di tutti i soggetti e aggiustamento del modello per la presenza di un ricovero nei 30 giorni precedenti l'evento.
3. Inclusione di tutti i soggetti e valutazione degli stessi in merito all'esposizione nei 30 giorni precedenti l'evento, assegnando a ciascun individuo dei pesi a seconda della proporzione di giorni misurabili (di non ospe-

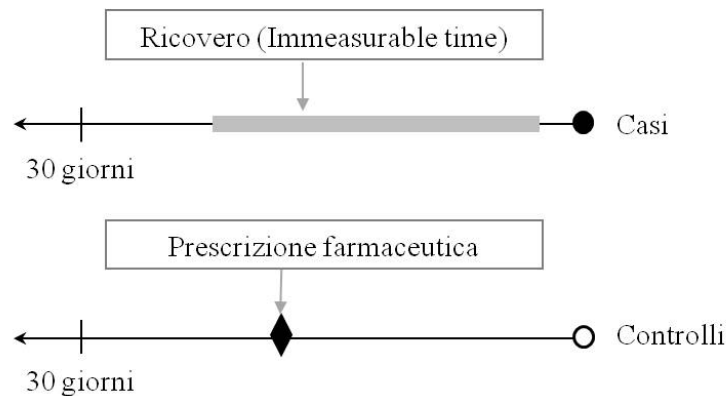


Figura 6.5: Descrizione di due soggetti partecipanti ad uno studio caso-controllo. Il periodo di ricovero (primo soggetto) rende impossibile la valutazione dell'esposizione ad un farmaco nei 30 giorni precedenti la data indice, generando immeasurable time bias (ITB). Tale soggetto può essere erroneamente classificato come "non esposto" qualora abbia ricevuto una prescrizione farmaceutica in tale periodo.

dalizzazione) nello stesso periodo (pesi=1 per i soggetti senza alcun ricovero nei 30 giorni precedenti l'evento).

4. Inclusione di tutti i soggetti e valutazione degli stessi in merito all'esposizione nei 30 giorni precedenti l'evento, considerando esposti tutti coloro con almeno una prescrizione nei 30 giorni precedenti l'evento e assegnandogli un peso pari a 1 (a prescindere dalla lunghezza del periodo misurabile) e pesando i soggetti non esposti proporzionalmente al numero di giorni misurabili in tal periodo.
5. Calcolo dei tassi di esposizione e dell'incidenza cumulativa di esposizione nei 30 giorni precedenti l'evento, determinando il contributo temporale di casi e controlli in accordo ai giorni misurabili in tale periodo.

Riconoscere e correggere l'immeasurable time bias

L'immeasurable time bias può essere particolarmente problematico quando si studiano malattie croniche che richiedono l'utilizzo costante di farmaci e possono portare a ospedalizzazioni multiple.

Pur essendo una distorsione che può influenzare qualsiasi studio condotto tramite database amministrativi, ad oggi, sono pochi gli autori che utilizza-

no metodi per tener sotto controllo la potenziale misclassificazione dovuta al tempo non misurabile. Infatti, in una review del 2010 sulle caratteristiche metodologiche degli studi di coorte, Nikitovic, Solomon e Cadarette hanno trovato che solamente 2 studi su 9 menzionavano l'immeasurable time bias, nei quali si affermava che l'aderenza ai trattamenti oggetto di studio era stata calcolata senza considerare la possibilità di periodi di ricovero [54].

La poca attenzione prestata all'effetto di questa distorsione, può essere dovuta al fatto che non esiste una tecnica precisa da utilizzare per la correzione. Suissa, nel suo lavoro del 2008 ne propone diverse e lascia al ricercatore la scelta di quale potrebbe essere quella che meglio si adatta alle specifiche esigenze.

6.2 Regression Calibration

La Regression Calibration (RC) è stata proposta originariamente da Rosner et al. per modelli di regressione logistica [55] anche se successivamente è stata estesa a differenti contesti da altri autori [56–62].

La RC è un metodo statistico per aggiustare la stima puntuale e intervallare del parametro di interesse, ottenuto da un modello di regressione logistica, che si suppone essere distorto a causa della presenza di errore di misura.

Per illustrare il funzionamento dell'approccio RC, si consideri uno studio (studio principale) per valutare l'associazione tra una variabile di esposizione X continua affetta da errore di misura e un outcome dicotomo Y ($Y = 1$ per gli eventi, 0 per non-eventi). Si supponga che nello studio principale non sia disponibile alcuna informazione sulla vera variabile di esposizione, X^* , ovvero quella misurata correttamente ma che tale variabile sia presente in uno studio di validazione, che è spesso un sottocampione dello studio principale, nel quale è disponibile anche la variabile affetta da errore X (Figura 6.6).

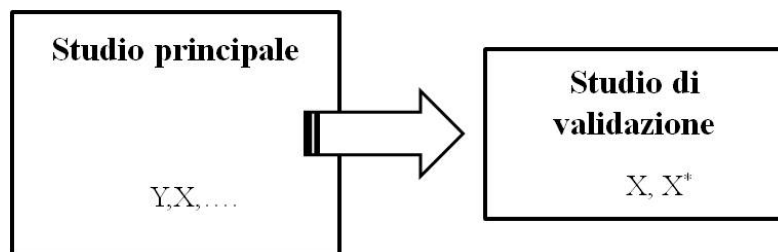


Figura 6.6: *Illustrazione dell'approccio RC: uno studio di validazione condotto in un sottocampione casuale dello studio principale può essere utilizzato per convalidare la variabile di esposizione osservata (X) nello studio principale. La vera variabile di esposizione (X^*) è disponibile solo nello studio di validazione insieme a X .*

L'idea di base della RC è quella di implementare una regressione lineare all'interno dello studio di validazione, che modelli la relazione funzionale tra la vera variabile di esposizione, X^* , e la variabile affetta da errore X . Il parametro di regressione risultante viene successivamente utilizzato per correggere la stima ottenuta dal modello di regressione logistica di partenza.

Ignorando l'errore di misura, si supponga di implementare nello studio principale un modello di regressione logistica che descriva la relazione tra Y e X ,

$$\text{logit}[P(Y = 1|X)] = \beta_0 + \beta X \quad (6.4)$$

e di ottenere una stima del parametro di interesse $\hat{\beta}$. Alla luce di quanto appena detto, tale stima è distorta.

La stima aggiustata mediante l'approccio RC, $\hat{\beta}_{RC}$, è

$$\hat{\beta}_{RC} = \frac{\hat{\beta}}{\hat{\gamma}} \quad (6.5)$$

con varianza

$$\text{Var}(\hat{\beta}_{RC}) = \frac{1}{\hat{\gamma}^2} \text{Var}(\hat{\beta}) + \frac{\hat{\beta}^2}{\hat{\gamma}^4} \text{Var}(\hat{\gamma}), \quad (6.6)$$

dove $\hat{\gamma}$ è la stima del parametro ottenuto fittando il seguente modello di regressione lineare nello studio di validazione:

$$E(X^*|X) = \alpha + \gamma X. \quad (6.7)$$

La RC prevede diverse assunzioni che necessitano di essere verificate.

Infatti, il metodo risulta appropriato quando:

- La variabile di esposizione, assunta essere l'unica variabile affetta da errore, è di natura continua;
- La vera variabile di esposizione è disponibile in uno studio di validazione così come la variabile misurata con errore;
- La relazione tra la vera variabile e quella misurata con errore è lineare.

Per maggiori dettagli si veda [56].

Potenzialità e limiti della Regression Calibration

La Regression Calibration è un metodo molto semplice e intuitivo che consente di correggere le stime di un modello di regressione, distorte a causa dell'errore di misura.

Questo metodo è stato sviluppato nell'ambito dell'epidemiologia nutrizionale. L'applicazione alla farmacoepidemiologia è ancora piuttosto limitata, probabilmente poiché il metodo richiede delle assunzioni piuttosto forti sulla relazione tra la variabile vera e quella affetta da errore.

Per evitare di dover porre assunti troppo forti, è possibile (i) implementare un modello che interpoli adeguatamente la relazione tra la vera variabile e quella affetta da errore, (ii) calcolare il valore predetto di (X^*) e (iii) inserire tale quantità nel modello (6.4) al posto di X .

Parte IV

Applicazioni

Capitolo 7

Fonte dei dati

I dati utilizzati ai fini del Progetto sono stati estratti da database amministrativi e clinici, in relazione a 3 specifici contesti.

A. Database amministrativi, Progetto AIFA-BEST

Il Progetto AIFA-BEST (*Bisphosphonates Efficacy-Safety Tradeoff*) è un progetto nazionale che coinvolge quattro regioni Italiane (Abruzzo, Emilia-Romagna, Marche e Toscana) e nove ASL (Caserta, Como, Gorizia, Latina, Lodi, Milano, Monza, Sondrio e Varese), ed è volto a studiare il profilo di sicurezza ed efficacia dei bifosfonati nella pratica clinica corrente.

A questo scopo, ciascuna unità territoriale ha fornito diversi database amministrativi che comprendevano: (i) il database dell'anagrafe, contenente informazioni demografiche di tutti i pazienti assistiti dal Servizio Sanitario Nazionale (SSN) e residenti in ciascuna unità; (ii) il database delle schede di dimissione ospedaliera (SDO), contenente informazioni relative ai ricoveri ospedalieri dei pazienti, e (iii) il database delle prescrizioni farmaceutiche, costituito da tutte le prescrizioni farmaceutiche erogate dalle farmacie territoriali.

B. Database amministrativi, Regione Lombardia

La Regione Lombardia adotta dal 1997 un sistema automatizzato di archivi elettronici sanitari che raccolgono informazioni demografiche di tutti gli

assistiti (DB dell'anagrafe), informazioni sulle prestazioni sanitarie erogate in regime di rimborso SSN, comprendendo i ricoveri ospedalieri e gli accessi in day-hospital (DB delle schede di dimissione ospedaliera, SDO), la specialistica ambulatoriale e la diagnostica strumentale, le analisi ematochimiche erogate (DB delle prestazioni ambulatoriali), la farmaceutica territoriale, che comprende i farmaci in fascia A (DB delle prescrizioni), e quella ospedaliera che comprende tutti i farmaci erogati dalle farmacie ospedaliere (File F).

Attraverso un procedura di record linkage deterministico, tutte le informazioni relative ad uno stesso paziente e provenienti da DB differenti (nei contesti A e B), sono state collegate tra loro in modo tale da ricostruire l'intera storia prescrittiva e clinica di ogni assistito.

Questa procedura è stata effettuata attraverso un codice personale univoco presente in ciascun database. Per garantire la riservatezza e impedire qualsiasi riconoscimento di dati individuali, questo codice è stato automaticamente convertito in un nuovo codice, anonimo, creato da un generatore di codici casuali.

C. Database clinico, Progetto SOS

Il Progetto SOS (*Safety Of non-Steroidal anti-inflammatory drugs*), è un consorzio di 11 Istituti di Ricerca europei, coordinato dall'Università Erasmus Medical Center (Paesi Bassi), e ha come obiettivo lo studio della sicurezza dei farmaci antiinfiammatori non steroidei (FANS).

In particolare, i dati utilizzati per il Progetto della Tesi sono stati estratti da uno dei DB partecipanti al Progetto SOS, il database olandese PHARMO che, oltre a contenere dati amministrativi, contiene anche informazioni cliniche.

7.1 User-only design

7.1.1 Bifosfonati e Fratture, prevenzione secondaria

L'obiettivo di questo studio era quello di confrontare differenti strategie per il controllo del confondimento da indicazione, distorsione che si colloca nell'ambito del confondimento non misurato.

In particolare, è stata confrontata la performance di alcune tecniche analitiche convenzionali quali l'aggiustamento multivariato e l'appaiamento per il Propensity Score (PS, Sezione 4.3 a pag. 42) con quella dei disegni **user-only** (Sezione 3.1 a pag. 14) nel produrre stime dell'efficacia di un trattamento corretto.

Per verificare se questa strategia riduca la distorsione derivata da confondimento da indicazione rispetto alle tecniche convenzionali, è stato utilizzato uno studio condotto per valutare l'efficacia dei bifosfonati orali (BF) nel ridurre il rischio di sviluppare una seconda frattura osteoporotica, partendo da una coorte di pazienti ricoverati (per la prima volta) per frattura osteoporotica.

Questo studio si adattava al nostro obiettivo per due ragioni: (i) l'efficacia dei bifosfonati nella prevenzione secondaria delle fratture è dimostrata in numerosi studi [63] e (ii) sono disponibili i risultati di diversi RCTs ¹ e meta-analisi di RCTs, che possono essere considerati come gold-standard.

Fonte dei dati

I dati utilizzati per condurre lo studio sono stati estratti dai database amministrativi del Progetto AIFA (Contesto B, pag. 62).

Selezione della coorte

Sono stati estratti tutti i pazienti con almeno 55 anni di età ricoverati per una frattura osteoporotica tra il 1 Luglio 2003 e il 31 Dicembre 2005. La data di dimissione è stata definita come *data di ingresso*. Da questa popolazione, sono stati esclusi tutti coloro che, nei sei mesi precedenti l'ingresso,

¹RCTs: Randomized Clinical Trials, studi clinici randomizzati

(i) avevano ricevuto una o più prescrizioni di BF, (ii) erano stati ricoverati per frattura osteoporotica, complicazioni gastrointestinali, e altre patologie. Sono stati inoltre esclusi i pazienti affetti da tumore maligno (in qualsiasi momento del periodo di osservazione) e i pazienti con meno di 1 anno di follow-up.

Ogni membro della coorte è stato seguito dopo sei mesi dalla data di ingresso (inizio del follow-up) fino al primo tra i seguenti eventi: ricovero per frattura osteoporotica, morte o fine del follow-up (31 Dicembre 2007).

Esposizione e covariate

I membri della coorte sono stati inizialmente suddivisi in due categorie, *utilizzatori* di BF, se avevano ricevuto almeno una prescrizione del farmaco di interesse entro 6 mesi dalla data di ingresso nella coorte e *non utilizzatori* in caso contrario. Successivamente, è stata valutata la durata della terapia con BF, nel periodo di follow-up, e categorizzata in *utilizzatori per più di 1 anno*, *utilizzatori per meno di 1 anno* e *non utilizzatori*.

Per ogni membro della coorte è stato valutato l'utilizzo di altri farmaci così come le condizioni di salute nell'anno precedente l'inizio del follow-up nello studio e durante il follow-up. I farmaci considerati includevano gli agenti antiipertensivi e statine, anticoagulanti, farmaci per il diabete, corticosteroidi e glicocorticoidi, inibitori della pompa protonica, antipsicotici e antidepressivi, sedativi e narcotici, terapia ormonale sostitutiva, raloxifene, integratori di calcio e vitamina D. La condizione di salute è stata valutata considerando la presenza di uno o più ricoveri per ipotensione ortostatica, malattie cerebrovascolari, disturbi dell'equilibrio, malattia di Parkinson, demenza, epilessia e cecità.

Strategie per il disegno dello studio

A partire dalla coorte di partenza, sono stati innestati 3 studi caso-controllo.

(i) Nel primo approccio (disegno con appaiamento convenzionale), i casi erano i pazienti che durante il follow-up erano stati ricoverati per una frattura osteoporotica. La data del ricovero per l'esito di interesse è stata definita come *data indice*. Sono stati poi selezionati fino a 5 controlli, tra i membri

della coorte ancora a rischio di sperimentare l'esito di interesse al momento della data indice del caso, e appaiati ad ogni caso per sesso, età e data di inizio follow-up.

(ii) Nel secondo approccio (disegno con appaiamento per PS), ogni caso è stato appaiato ad un controllo che avesse, oltre che la stessa età e lo stesso genere, lo stesso valore del PS. Il PS è stato calcolato al baseline (inizio del follow-up) attraverso un modello di regressione logistica per modellare la probabilità di esposizione a BF al baseline, sulla base di 24 covariate misurate nell'anno precedente l'inizio del follow-up.

(iii) Nel terzo approccio, (disegno user-only) la coorte è stata ristretta ai pazienti esposti al trattamento di interesse entro sei mesi dalla data di ingresso. Per ogni caso generato da tale coorte, sono stati selezionati fino a 5 controlli e appaiati ad ogni caso per sesso, età e data di inizio follow-up.

Analisi statistica

Per ognuno dei tre approcci, per stimare (a livello puntuale e intervallare) l'Odds Ratio (OR) di una seconda frattura associato all'uso di BF al baseline, nonché durante il follow-up, è stato implementato un modello di regressione logistica condizionata aggiustato per le covariate considerate. Per i primi due approcci il confronto è stato effettuato utilizzando come categoria di riferimento i non utilizzatori (al baseline o durante l'intero follow-up).

Per il terzo approccio, ovviamente, la categoria di riferimento era costituita dagli utilizzatori per meno di 1 anno.

Controllo del confondimento non misurato

Per tenere sotto controllo il confondimento da indicazione è stato implementato un disegno **user-only** .

Risultati principali

Nel primo approccio (appaiamento convenzionale), la stima dell'OR di frattura osteoporotica associata all'uso di BF al baseline (con riferimento al non uso), aggiustata per l'effetto dell'età, dell'utilizzo di farmaci concomitanti e dello stato di salute, è risultata pari a 1.20 (95% IC 1.01-1.44). L'uso pro-

lungato del farmaco (per più di un anno), rispetto al non uso sembra essere associato ad un aumento del rischio del 18%, mentre l'uso per un breve periodo non sembra far emergere alcun effetto sul rischio di frattura (Figura 3, box **a** e **b**).

Appaiamento tramite PS

L'appaiamento con il PS non ha portato a sostanziali differenze. Infatti si può notare come non emerga alcun effetto del farmaco di interesse sulla prevenzione delle fratture osteoporotiche (Figura 3, box **c** e **d**).

Disegno user-only

Restringendo l'analisi ai soli utilizzatori di BF, emerge l'atteso effetto benefico sulla prevenzione di una seconda frattura. Infatti, confrontando il rischio dei pazienti che assumono BF per più di 1 anno con quello dei pazienti che lo assumono per poco tempo, si può osservare come un uso prolungato del farmaco sia associato ad una riduzione del rischio di fratture del 28% (95% IC 1% - 48%). Scegliendo come riferimento, la categoria dei pazienti che utilizzano BF per un periodo di tempo ancor più breve (meno di 6 mesi), la riduzione del rischio di frattura a seguito di un'esposizione superiore ad 1 anno è risultata pari al 36% (95% IC 7% - 56%) (Figura 7.1, box **e** ed **f**).

Conclusioni

L'incremento del rischio di frattura ottenuto sia dall'approccio convenzionale che dall'appaiamento tramite PS sembra essere influenzato dalla presenza di confondimento da indicazione, dal momento che i pazienti che assumono BF potrebbero essere più ad alto rischio di sviluppare l'esito di interesse.

L'approccio basato sull'appaiamento per PS, sembra non essere sufficiente per tenere sotto controllo tale distorsione. E' importante sottolineare che il PS è utile per tenere sotto controllo il confondimento misurato.

La restrizione ai soli utilizzatori, invece, sembra rappresentare lo strumento ideale per studiare l'efficacia di un farmaco in presenza di confondimento da indicazione, come nel nostro contesto.

Questo studio è stato pubblicato su *Pharmacoepidemiology & Drug Safety* nel 2014 [19].

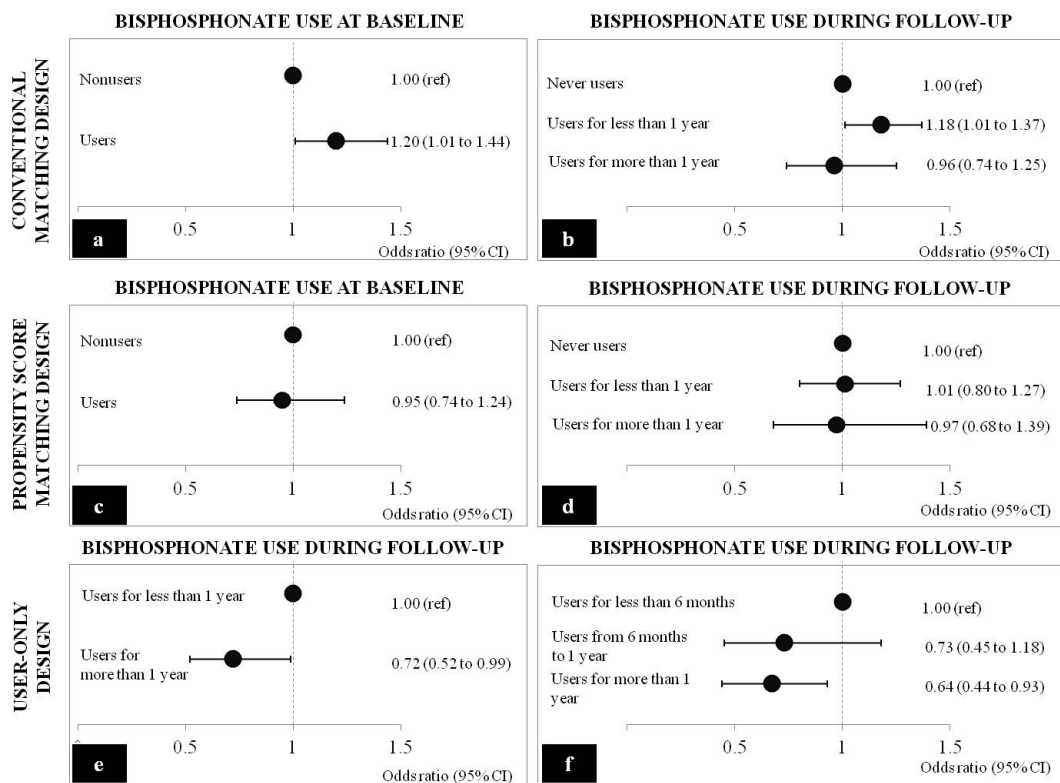


Figura 7.1: Effetto dell'uso di BF sul rischio di una seconda frattura, utilizzando diverse strategie di analisi

7.2 Case-only design

7.2.1 Antiinfiammatori e Sanguinamento gastrointestinale

Questo studio aveva come obiettivo il confronto della performance di alcuni disegni case-only (case-crossover, case-time-control e case-case-time-control) per la correzione del confondimento non misurato, nel contesto di uno studio volto a indagare la relazione tra utilizzo di antiinfiammatori non-steroidi (FANS) e rischio di sanguinamento gastro-intestinale (UGIB).

Fonte dei dati

I dati utilizzati per condurre lo studio sono stati estratti dai database amministrativi della Regione Lombardia (Contesto A, pag. 62).

Selezione della coorte

Per rispondere all'obiettivo dello studio, è stato implementato un disegno caso-controllo innestato in una coorte di utilizzatori di FANS, selezionando tutti i pazienti con almeno una prescrizione di FANS nel periodo 2004-2007. La data della prima prescrizione di FANS è stata definita come *data di ingresso* nello studio. Da questi pazienti, sono stati esclusi coloro che, nell'anno precedente all'ingresso, (i) avevano già ricevuto una prescrizione di FANS, in modo da limitare l'inclusione dei soli utilizzatori incidenti, (ii) erano già stati ricoverati per UGIB o per tumore maligno, al fine di escludere i soggetti a più alto rischio di UGIB.

Ciascun membro della restante coorte è stato seguito dall'ingresso nella coorte fino al primo tra i seguenti episodi: ricovero per UGIB (outcome), morte o fine del follow-up (31 Dicembre 2008).

Selezione di casi e controlli

I casi erano i pazienti che, durante il follow-up, erano stati ricoverati per UGIB. La *data indice* è stata definita come la data del primo ricovero per UGIB. Per ogni caso, sono stati selezionati fino a 10 controlli (tra i membri

della coorte ancora a rischio di sviluppare l'evento al momento dell'insorgenza del caso) e appaiati per età, genere e data di ingresso nello studio.

Esposizione e covariate

Durante il periodo in studio è stata valutata l'esposizione a FANS ed è stata classificata come *corrente*, se il paziente aveva ricevuto almeno una prescrizione di FANS nei 30 giorni precedenti la data indice e *passata*, se aveva ricevuto almeno una prescrizione tra i 90 e i 120 giorni prima della data indice. Per ogni membro della coorte, è stata valutata l'esposizione durante il follow-up ai seguenti farmaci concomitanti: antidepressivi, antiipertensivi, antitrombotici, inibitori della pompa protonica, aspirina, corticosteroidi, bifosfonati e antibiotici.

Analisi statistica

Per stimare (a livello puntuale e intervallare) l'Odds Ratio (OR) di UGIB associato all'uso corrente di FANS (rispetto all'uso passato) è stato implementato un modello di regressione logistica condizionata aggiustato per tutte le covariate.

Controllo del confondimento non misurato

Per tenere sotto controllo il confondimento non misurato, è stato inoltre implementato un disegno **case-crossover** (Sezione 3.2.1 a pag. 17), restringendo l'analisi ai soli casi di UGIB. In questo caso, la finestra a rischio coincideva con la finestra corrente del caso-controllo, mentre la finestra di controllo corrispondeva alla finestra passata.

Dal momento che le stime del disegno case-crossover potrebbero essere affette da un trend temporale dell'esposizione, è stato valutato l'andamento delle prescrizioni di FANS nel periodo considerato.

Successivamente, sono stati applicati anche (i) un disegno **case-time-control** (Sezione 3.2.2 a pag. 22) e (ii) un disegno **case-case-time-control** (Sezione 3.2.3 a pag. 28), utilizzando le stesse finestre di confronto utilizzate nel case-crossover.

Risultati preliminari

Dall'analisi principale è emerso che l'utilizzo corrente di FANS aumenta considerevolmente il rischio di sanguinamento gastrointestinale. L'OR di UGIB associato con l'uso corrente di FANS, infatti, è risultato pari a 2.45, (95% IC 1.99-3.01). Tali risultati, tuttavia, potrebbero essere affetti da confondimento non misurato. La stima dell'OR di UGIB ottenuta dal disegno case-crossover è risultata decisamente inferiore a quella ottenuta dall'approccio classico, e pari a 1.89 (95% IC 1.62-2.21) (Tabella 7.1).

Come si è visto nella Sezione 3.2.1, se fosse presente un trend nelle prescrizioni, da tale stima dovrebbe essere rimosso l'effetto del trend.

Il trend delle prescrizioni è mostrato in (Figura 7.2).

Come si può osservare, le prescrizioni di FANS hanno un andamento decrescente nel primo biennio mentre, successivamente, la direzione si inverte.

Questo andamento, potrebbe aver distorto i risultati ottenuti dal case-crossover.

Applicando un case-time-control, si è ottenuto un OR di 1.87 (95% IC 1.44-2.43), un risultato molto simile a quello ottenuto dal disegno case-crossover, che sembrerebbe smentire la presenza di un trend dell'esposizione.

Utilizzando invece i casi futuri, al posto dei controlli (case-case-time-control) si è ottenuto un OR pari a 2.19 (95% IC 1.73-2.79).

Disegno	OR (IC 95%)
Caso-controllo	2.45 (1.99-3.01)
Case-crossover	1.89 (1.62-2.21)
Case-time-control	1.87 (1.44-2.43)
Case-case-time-control	2.19 (1.73-2.79)

Tabella 7.1: *Confronto delle stime degli ORs di UGIB ottenute da diversi disegni.*

Conclusione

In questo contesto, l'approccio convenzionale sembra essere verosimilmente condizionato dalla presenza di confondimento non misurato. Infatti, i risultati di tutti e 3 i disegni case-only utilizzati, mostrano una riduzione del rischio di UGIB associato all'uso di FANS.

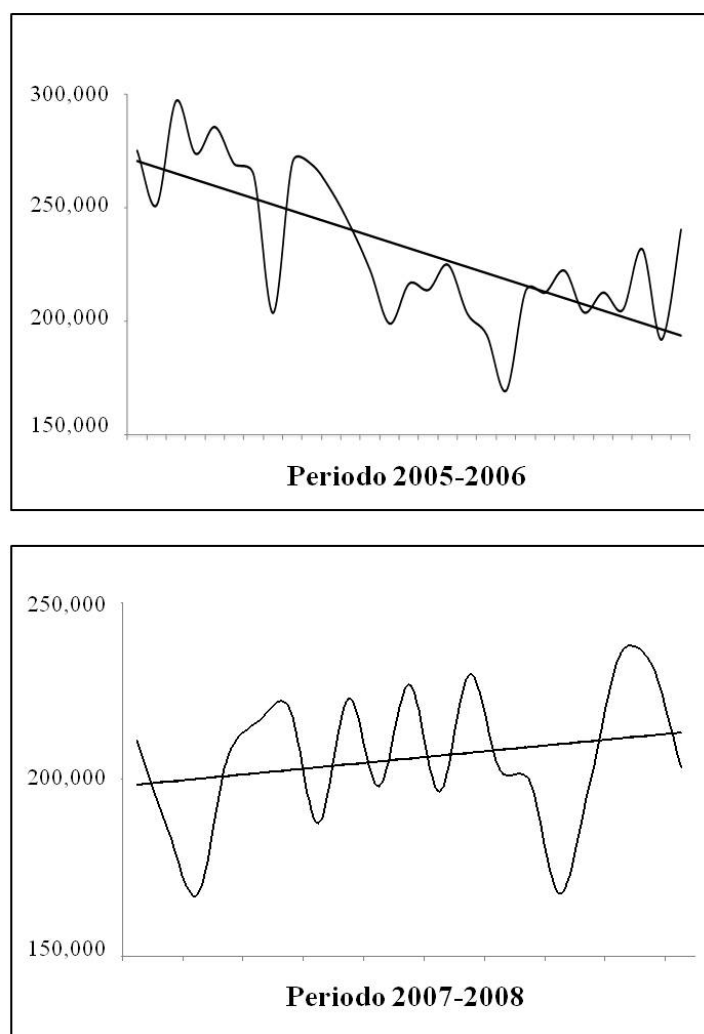


Figura 7.2: *Trend delle prescrizioni di FANS nel periodo di interesse.*

Il case-crossover, a sua volta potrebbe essere affetto dal trend dell'esposizione, dal momento che dalla (Figura 7.2) si vede chiaramente che l'andamento dell'esposizione è decrescente, nei primi anni, e poi crescente.

Il case-time-control, tuttavia, non riesce a "catturare" completamente il diverso effetto nei due periodi. Questo poiché potrebbero essere stati violati alcuni degli assunti sottostanti al disegno. Infatti, affinché questo metodo sia valido, è necessario che il trend delle prescrizioni nei controlli sia sovrapponibile al trend nei casi, e questo potrebbe non essersi verificato nel nostro studio.

Il case-case-time-control, che consente di eliminare il problema del diverso trend delle prescrizioni in casi e controlli, ha prodotto una stima dell'OR di UGIB leggermente incrementata rispetto a quella ottenuta dal case-crossover e dal case-time-control ma pur sempre inferiore a quella ottenuta dal caso-controllo.

Questo studio è ancora in fase di sviluppo.

7.2.2 Antidepressivi e Fratture

L'obiettivo di questo studio era quello di valutare l'associazione tra l'utilizzo di antidepressivi e il rischio di frattura.

In particolare, tra tutte le classi di antidepressivi, gli antidepressivi triciclici (TCA) e gli inibitori della ricaptazione della serotonina (SSRI), sono state considerate le due classi più comunemente utilizzate.

Da recenti studi sono emerse numerose evidenze a favore dell'ipotesi che l'utilizzo di entrambe le classi di antidepressivi aumenti il rischio di frattura, sebbene l'uso a breve termine di TCA sembri essere associato ad un rischio superiore (o tutt'al più simile) a quello associato all'uso a breve termine di SSRI. Per indagare gli effetti a breve termine di TCA e SSRI è stato condotto uno studio caso-controllo innestato in una coorte di utilizzatori di antidepressivi.

Fonte dei dati

I dati utilizzati per condurre lo studio sono stati estratti dai database amministrativi della Regione Lombardia (Contesto B, pag. 62).

Selezione della coorte

La coorte era costituita da tutti i pazienti residenti in Lombardia di età superiore ai 18 anni, con almeno una prescrizione di antidepressivi nel periodo 2005-2007. La data della prima prescrizione è stata definita come *prescrizione indice*.

Sono stati esclusi tutti i pazienti che, nei 5 anni precedenti la prescrizione indice (i) avevano già ricevuto almeno una prescrizione di antidepressivi, in modo tale da selezionare solamente i nuovi utilizzatori di antidepressivi; (ii) erano stati ricoverati con diagnosi di depressione, di osteoporosi o fratture osteoporotiche, osteomalacia, (iii) avevano ricevuto almeno una prescrizione di farmaci per il trattamento delle malattie delle ossa (bisfosfonati, terapia ormonale sostitutiva, raloxifene, calcio, ranelato di stronzio, etc...) in modo tale da escludere i pazienti ad alto rischio di fratture. Infine, (iv) tutti i pazienti affetti da tumore (con ricovero ospedaliero in qualsiasi momento del

follow-up) sono stati esclusi dalle analisi per escludere i pazienti più fragili. Ogni membro della restante coorte è stato seguito dall'ingresso nello studio fino al primo tra i seguenti episodi: ospedalizzazione per frattura, uscita dallo studio, fine del follow-up (31 Dicembre 2010).

Selezione dei casi

I casi erano gli individui che, nel periodo considerato, erano stati ricoverati per frattura. La data del ricovero per frattura è stata definita come *data indice*. Per ogni caso, sono stati selezionati fino a 10 controlli, selezionati tra i membri della coorte ancora a rischio di sperimentare l'evento al momento della data indice del caso e sono stati appaiati per età, genere e data di ingresso nello studio.

Esposizione e covariate

Per ciascun individuo è stata valutata la distanza tra l'ultimo antidepressivo assunto e la data indice. In particolare, sono state considerate le seguenti tre categorie per l'utilizzo di antidepressivi: *corrente*, se la distanza tra l'ultimo giorno di copertura e la data indice era inferiore a 30 giorni, *recente* se compresa tra i 31 e i 60 giorni e *passato* se superava i 61 giorni. Inoltre, l'uso corrente è stato stratificato per la tipologia dell'ultimo antidepressivo assunto (SSRI o TCA).

Per ogni soggetto, è stato valutato l'utilizzo di altri farmaci e la presenza di altre comorbidità nell'anno precedente la prescrizione indice così come durante il follow-up. I farmaci considerati includevano: ansiolitici, antipsicotici, antiipertensivi, bifosfonati, corticosteroidi orali, HRT, inibitori di pompa protonica, sedativi e narcotici. La condizione di salute è stata valutata considerando uno o più ricoveri per artrite, malattie delle ossa, sindrome di Cushing, Parkinson, celiachia, bipolarismo, schizofrenia, disordini mentali, demenza, disturbi cardiaci, ipertensione, ipotensione ortostatica, malattia cerebrovascolare, diabete, epilessia, BPCO e malattie della tiroide.

Analisi statistica

Per stimare (a livello puntuale e intervallare) l'Odds Ratio (OR) di frattura

associato alle categorie di esposizione è stato implementato un modello di regressione logistica condizionata aggiustato per le covariate considerate.

Controllo del confondimento non misurato

Per minimizzare l'effetto del confondimento non misurato è stato implementato uno studio **case-crossover** (si veda la Sezione 3.2.1 a pag. 17), restringendo l'analisi ai soli casi di frattura. Per ogni caso, si è confrontata l'esposizione a SSRI/TCA nel periodo a rischio (1-30 giorni precedenti la data indice) con quello di controllo (61-90 giorni precedenti la data indice). Per stimare (a livello puntuale e intervallare) gli OR di frattura associato all'uso di SSRI/TCA nel periodo a rischio (rispetto all'uso di SSRI/TCA nel periodo di controllo) è stato implementato un modello di regressione logistica condizionata.

Successivamente è stato anche implementato un **self-controlled case-series** (si veda la Sezione 3.2.4 a pag 31). In questo caso, dopo aver selezionato solamente i soli casi di frattura, ciascun caso è stato seguito per tutto il periodo a disposizione (l'intero periodo 2005-2010). Il follow-up di ogni caso di frattura è stato dunque partizionato in diversi periodi di esposizione (a TCA/SSRI) e non esposizione. In particolare, è stato definito come periodo ad alto rischio, il periodo tra gli 0-30 giorni dopo l'inizio di ciascuna prescrizione di TCA/SSRI mentre periodo a rischio, il periodo residuo coperto da ciascuna prescrizione di TCA/SSRI. Il periodo in cui un individuo non risultava esposto è stato definito come periodo baseline ed è stato utilizzato come riferimento.

E' stato implementato un modello di regressione condizionata di Poisson per stimare (a livello puntuale e intervallare) il tasso di incidenza relativo (IRR) di frattura confrontando il tasso di frattura nei periodi ad alto rischio e a rischio con il tasso di frattura nel periodo baseline.

Risultati principali

Dai risultati preliminari ottenuti dal caso-controllo innestato, è emerso che l'uso corrente di SSRI aumenta il rischio di frattura del 70% ($OR = 1.70$, $IC_{95\%} : 1.62 - 1.79$) se confrontato all'uso passato (Figura 7.3). I TCA invece, au-

mentano il rischio del 38% (OR=1.38, IC 95%: 1.15-1.65). Questo risultato inatteso, (ci si aspettava un effetto acuto più marcato, o tutt'al più simile, per i TCA) potrebbe essere spiegato dalla presenza di confondimento non misurato o confondimento da indicazione (ad esempio se gli SSRI venissero prescritti ai pazienti più fragili).

Disegno case-crossover

Dai risultati del case-crossover è emerso che il 20% dell'eccesso di rischio associato agli SSRI potrebbe essere spiegato dalla presenza di confondimento non misurato. Infatti, come si può notare dalla Figura 7.3, si passa da un OR_{SSRI} di 1.70 ottenuto dal caso-controllo ad un OR_{SSRI} pari a 1.36 (IC 95%: 1.22-1.51). L'effetto dei TCA invece è risultato simile a quello ottenuto dal caso-controllo innestato (da $OR_{TCA} = 1.38$ nel caso-controllo a $OR_{TCA} = 1.30$, IC 95% 0.93-1.85 nel case-crossover).

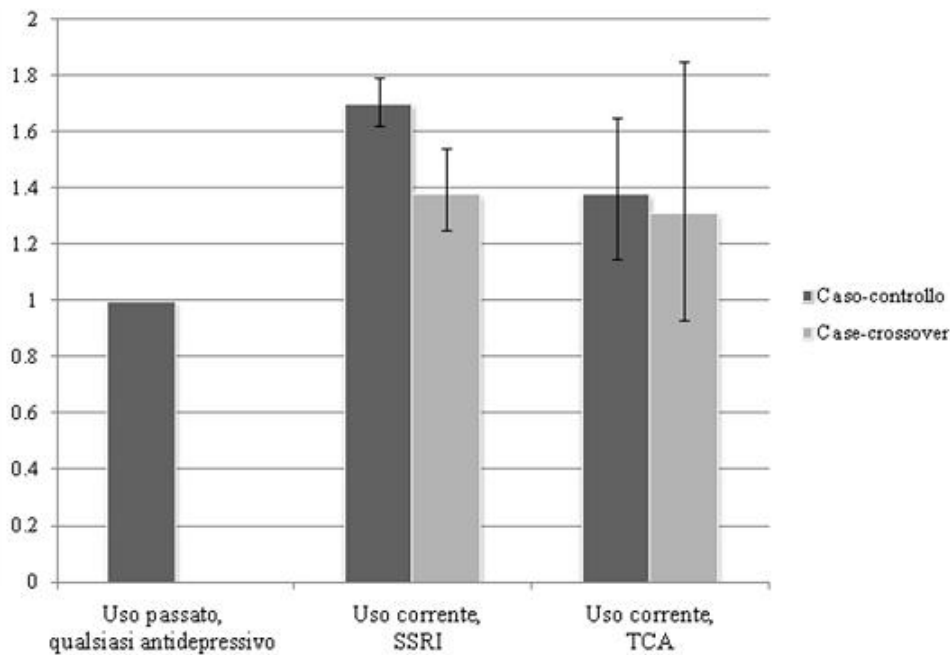


Figura 7.3: Stime degli ORs e dei relativi intervalli di confidenza di frattura associati all'uso corrente di SSRI/TCA ottenute dall'analisi principale (caso-controllo innestato) e dal disegno case-crossover.

Self-controlled case-series

A causa dell'elevata numerosità dei record del dataset costruito per l'analisi

case-series, il modello di Poisson non ha raggiunto la convergenza. Pertanto, si valuterà l'opportunità di spezzare il dataset globale in dataset più piccoli, di implementare il modello di regressione di Poisson in ciascuno di essi e di combinare opportunamente le stime ottenute da ciascun dataset in un'unica stima dell'IRR (e del suo intervallo di confidenza).

Conclusioni

L'effetto degli SSRI sul rischio di frattura, misurato nel caso-controllo innestato, sembra essere influenzato da confondimento non misurato.

L'eccesso di rischio osservato tra gli utilizzatori di questa classe di antidepressivi potrebbe essere in parte dovuto al confondimento da indicazione. In seguito all'applicazione del disegno case-crossover, l'effetto di SSRI e TCA sembra essere sovrapponibile.

Questo lavoro è stato presentato al *VII Congresso Nazionale della Società Italiana di Statistica Medica ed Epidemiologia Clinica (SISMEC)*, tenutosi a Roma nei giorni 25-28 Settembre 2013, attraverso una presentazione orale [64].

7.3 Monte-Carlo Sensitivity Analysis

7.3.1 Bifosfonati e Fratture, prevenzione primaria

Questo studio aveva come obiettivo lo studio dell'efficacia dei bifosfonati orali nel ridurre il rischio di frattura, in prevenzione primaria.

Fonte dei dati

I dati utilizzati per condurre questo studio sono stati estratti dai database amministrativi del Progetto AIFA (Contesto A, pag. 62).

Selezione della coorte

Sono stati selezionati tutti i pazienti con età superiore ai 55 anni con almeno una prescrizione di bifosfonati orali (alendronato e risedronato) nel periodo 1 Luglio 2003- 31 Dicembre 2005. la data della prima prescrizione di BF è stata definita come *data di ingresso*.

Da questi pazienti, sono stati esclusi coloro che: (i) erano già stati ricoverati per frattura osteoporotica nei sei mesi prima della data di ingresso, al fine di concentrarsi sulla prevenzione primaria delle fratture osteoporotiche; (ii) avevano già ricevuto almeno una prescrizione di BF (o di altri agenti per il turnover osseo) nei sei mesi prima dell'ingresso, in modo da privilegiare l'inclusione degli utilizzatori incidenti; (iii) avevano sperimentato un ricovero per morbo di Paget, osteomalacia, sindrome di Cushing, ipertiroidismo, iperparatiroidismo, malattia celiaca nei sei mesi prima dell'ingresso, al fine di escludere i pazienti a più alto rischio; (iv) erano stati ricoverati per tumore maligno, registrato durante l'intero periodo di osservazione.

Selezione di casi e controlli

Il database delle SDO è stato utilizzato per identificare i membri della coorte che sono stati ricoverati in ospedale, durante il follow-up, con una diagnosi di fratture osteoporotica. La data del primo ricovero ospedaliero è stata definita come *data indice*. Per ogni caso, sono stati selezionati da 1 a 20 controlli, tra i membri della coorte ancora a rischio di sperimentare l'outcome, e appaiati

per unità territoriale di reclutamento, sesso, età all'ingresso di coorte, e la data di ingresso.

Esposizione e covariate

Per ogni soggetto incluso, sono state identificate tutte le prescrizioni di BF rimborsabili dal SSN (alendronato, da solo o in combinazione con la vitamina D, e risedronato) dal database delle prescrizioni farmaceutiche.

L'esposizione a BF è stata valutata in base alla durata della terapia, sommando i giorni di copertura di ciascuna prescrizione. Successivamente, la durata è stata classificata come <1 anno, tra 1-2 anni, e ≥ 2 anni.

Per ogni caso e controllo, sono stati identificati i potenziali confondenti durante il follow-up, quali l'uso concomitante di farmaci (antidepressivi, anticoagulanti, inibitori della pompa protonica e antagonisti dei recettori H2, corticosteroidi, statine, antidiabetici, antiepilettici, raloxifene, calcio e vitamina D) e diverse comorbidità, quali malattia di Parkinson, ipotensione ortostatica, epilessia, cecità, demenza, malattie cerebrovascolari, identificate tramite la presenza di un ricovero ospedaliero. nello stesso periodo, è stato inoltre calcolato il *Charlson Comorbidity Index* (CCI) e classificato come pari a 0, 1, e ≥ 2 .

Analisi statistica

Per stimare gli ORs di frattura associati alle categorie di durata della terapia con BF (e i relativi intervalli di confidenza al 95%), è stato implementato un modello di regressione logistica condizionata.

Controllo del confondimento non misurato

Per quantificare l'effetto del confondimento non misurato, è stata implementata una **Monte-Carlo Sensitivity Analysis** (MCSA) (Sezione 4.1 a pag. 35).

Per semplificare l'interpretazione dei risultati, la durata della terapia con BF così come la gravità dell'osteoporosi sono state dicotomizzate nel modo seguente: ≥ 2 anni vs. <2 anni e severa vs. moderata, rispettivamente.

Inoltre, la prevalenza di osteoporosi severa, tra i pazienti con un trattamento

più prolungato, è stata fatta variare tra 0.2 e 0.8, mentre la prevalenza di pazienti severi nei soggetti in terapia da meno di 2 anni, è stata fissata a 0.2. Infine, sulla base di studi precedenti, è stata assunta una forza di associazione tra gravità della malattia e rischio di frattura che variava tra 1.2 e 3.5.

Risultati principali

Per i soggetti in terapia con BF per almeno 2 anni, si è osservata una riduzione del rischio di frattura del 21% (7%-33%), rispetto ai pazienti in terapia da meno di 1 anno.

La riduzione del rischio di frattura associato con l'uso prolungato di BF, non è stata annullata in seguito all'aggiustamento per il confondimento non misurato. Al contrario, aumentando la prevalenza di esposizione a tale confondente, così come la forza dell'associazione confondente-outcome, si è osservata un'azione protettiva progressivamente più pronunciata.

La Figura 7.4 mostra gli ORs di frattura, associati ai soggetti in trattamento da almeno 2 anni, aggiustati per la severità della malattia (come esempio di confondente non misurato) tramite la MCSA. Tanto più la prevalenza del confondente e la forza tra confondente-outcome aumentano, quanto più la riduzione del rischio di frattura diventa più marcata.

Ad esempio, assumendo una prevalenza del 40% di malati severi, tra coloro che sono in terapia da più tempo, le stime degli ORs si riducono progressivamente dalla stima osservata pari a 0.79 (0.67-0.93) a 0.72 (0.59-0.77), 0.68 (0.59-0.77), e 0.64 (0.56-0.74), assumendo una forza dell'associazione tra la severità della malattia e il rischio di frattura pari a 1.5, 2.0 e 2.5, rispettivamente.

Conclusioni

Questo studio mostra evidenze scientifiche che l'utilizzo prolungato di BF riduce il rischio di frattura del 21%. Tale risultato è stato ulteriormente confermato dall'applicazione della MCSA.

Questo studio è stato pubblicato su *European Journal of Clinical Pharmacology* nel 2014 [65].

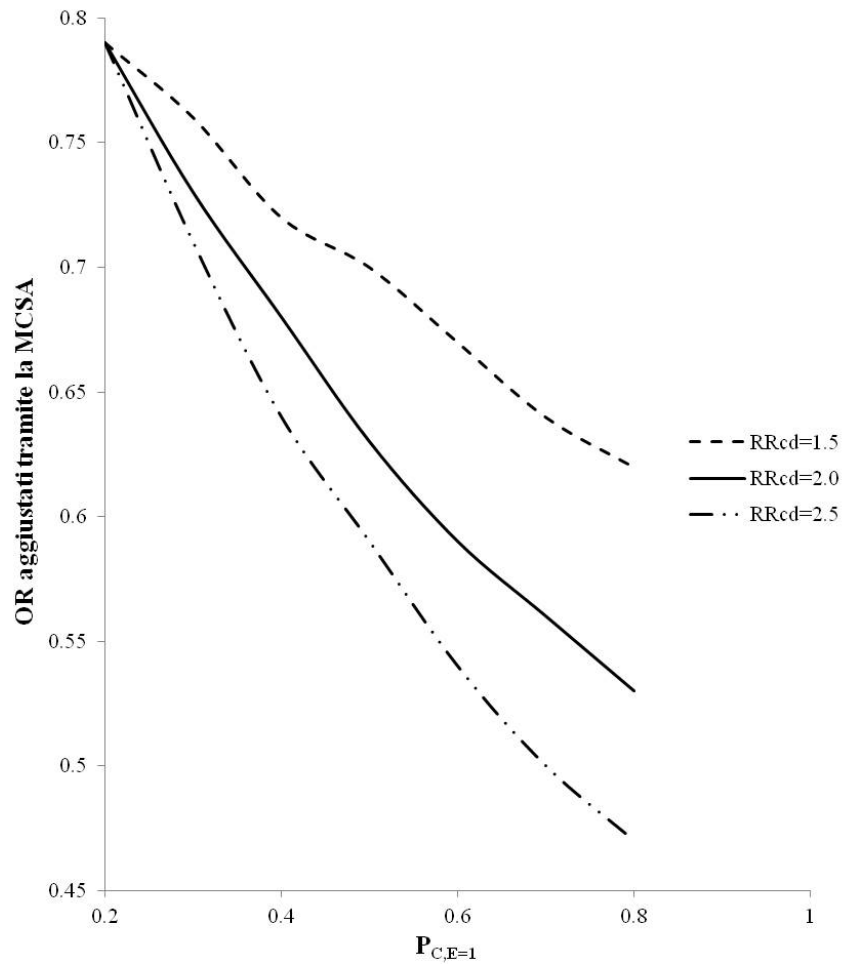


Figura 7.4: Cambiamenti negli ORs di frattura associati alle categorie di durata della terapia con BF, in seguito all'aggiustamento per il confondimento non misurato (gravità dell'osteoporosi). L'aggiustamento è stato effettuato tramite la Monte-Carlo Sensitivity Analysis.

7.4 Correzione Immeasurable Time Bias

7.4.1 Bifosfonati e Complicazioni gastrointestinali

L'obiettivo di questo studio era quello di valutare l'associazione tra utilizzo di bifosfonati orali (BF) e rischio di complicazioni gastrointestinali (UGIC) mediante un approccio caso-controllo innestato in una coorte di utilizzatori di BF.

Fonte dei dati

I dati utilizzati per questo studio sono stati estratti dai DB amministrativi del Progetto AIFA (Contesto A, pag. 62).

Selezione della coorte

Dai DB amministrativi di ciascuna unità territoriale sono stati estratti tutti i pazienti con almeno 45 anni di età che avevano ricevuto almeno una prescrizione di BF tra l'1 Luglio 2003 e il 31 Dicembre 2005.

Da questa popolazione, sono stati esclusi tutti coloro che nei sei mesi precedenti l'ingresso avevano ricevuto una o più prescrizioni di BF, così come coloro che in tale periodo erano stati ricoverati per frattura osteoporotica, complicazioni gastrointestinali, problemi di coagulazione, abuso di alcol e malattia cronica del fegato, e tumore maligno (in qualsiasi momento del periodo di osservazione). Infine, i pazienti con meno di 60 giorni di follow-up non sono stati inclusi nelle analisi. Ogni membro della coorte è stato seguito dalla data di ingresso fino al primo tra i seguenti eventi: ricovero per UGIC, morte o fine del follow-up (31 Dicembre 2007).

Selezione di casi e controlli

I casi erano i pazienti che, durante il follow-up, erano stati ricoverati per UGIC. La *data indice* è stata definita come la data di ricovero per UGIC. Per ogni caso sono stati selezionati fino a 20 controlli (tra i membri della coorte ancora a rischio di sviluppare l'evento al momento dell'insorgenza del caso) e appaiati per unità territoriale di appartenenza, età, genere e data di

ingresso nello studio.

Esposizione

Durante il periodo di studio è stata valutata l'esposizione a BF ed è stata poi classificata come "corrente" quando l'ultima pastiglia assunta dal paziente cadeva nei 30 giorni precedenti la data indice o "passata" quando l'ultimo utilizzo di BF era avvenuto nei 31 giorni (o più) prima della data indice.

Analisi statistica

Per stimare (a livello puntuale e intervallare) l'Odds Ratio (OR) di UGIC associato all'uso corrente di BF (rispetto all'uso passato) è stato implementato un modello di regressione logistica condizionata aggiustato per diverse covariate (trattamenti concomitanti).

Correzione dell'errore di misura

Dal momento che nei DB amministrativi non è contenuta alcuna informazione riguardo i farmaci prescritti durante un periodo di ricovero, non è possibile identificare e, pertanto, classificare l'esposizione a BF in modo corretto qualora i pazienti fossero stati ricoverati, per cause diverse da UGIC, durante il follow-up. Per correggere la distorsione derivante da **immeasurable time bias** (si veda Sezione 6.1.3 a pag. 54) sono state proposte diverse tecniche di analisi.

Risultati principali

La stima dell'OR di UGIC associata all'uso corrente di BF (con riferimento ad un uso passato), aggiustata per il solo effetto dell'età e dell'utilizzo di farmaci concomitanti, è risultata pari a 0.75 (95% IC 0.65-0.94).

L'applicazione di alcune tecniche per l'aggiustamento dell'immeasurable time bias ha prodotto i seguenti risultati (Figura 7.5):

- (i) Escludendo i giorni di ricovero dal follow-up di casi e controlli, la stima dell'OR di UGIC associata all'utilizzo corrente di BF è risultata pari a 0.94 (IC 95% 0.81-1.08).

- (ii) Restringendo l'analisi ai soggetti senza ospedalizzazioni nei 30 giorni precedenti la data indice o aggiustando per la presenza di un ricovero in tale periodo la stima dell'OR è risultata, rispettivamente, pari a 0.79 (95% IC 0.65-0.96) e a 0.80 (95% IC 0.66-0.96).
- (iii) Utilizzando invece il tempo misurabile nei 30 giorni precedenti la data indice per calcolare il tasso di esposizione, l'OR di UGIC è risultato pari a 0.87 (95% IC 0.75-1.00) mentre attraverso l'incidenza cumulativa di esposizione nei 30 giorni precedenti l'evento, calcolata tramite lo stimatore di Kaplan-Meier, si è ottenuto un OR pari a 1.03 (95% IC 1.00-1.07).

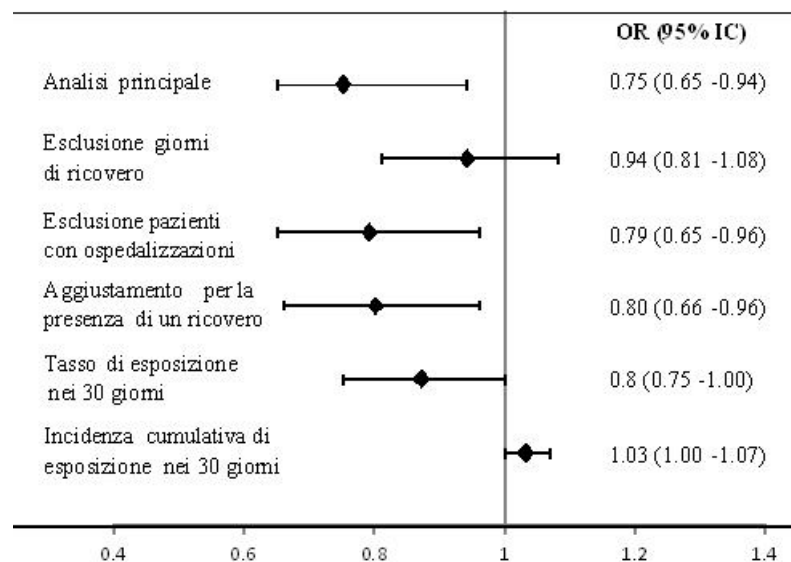


Figura 7.5: Stime degli ORs e dei relativi intervalli di confidenza di UGIC associati all'uso corrente di BF ottenute dall'analisi principale e da alcune tecniche per il controllo dell'immeasurable time bias.

Supponendo che l'assunzione di BF aumenti il rischio di UGIC e, alla luce del fatto che l'UGIC sia un evento acuto, ci si attenderebbe che il rischio di complicazioni gastrointestinali a seguito di un'esposizione corrente sia superiore a quello relativo ad un'esposizione passata, ipotesi che, tuttavia, non

viene confermata dall'analisi principale.

Conclusioni

I risultati dello studio suggeriscono che l'effetto dell'uso corrente di BF sul rischio di UGIC sia verosimilmente condizionato da immeasurable time bias. L'effetto protettivo riscontrato senza alcun aggiustamento, infatti, non viene confermato dai risultati ottenuti a seguito dell'applicazione delle tecniche per il controllo della distorsione presentata.

Una delle tecniche per la correzione dell'immeasurable time bias, in particolare l'esclusione dei giorni di ricovero dal follow-up di casi e controlli, è stata inclusa nell'analisi principale del lavoro pubblicato su *PLOS One* nel 2013 [66].

7.5 Approccio Rule-Out

7.5.1 Bifosfonati e Complicazioni gastrointestinali

Si faccia riferimento allo studio della Sezione precedente.

In seguito alla correzione per l'immeasurable time bias, è stata effettuata una sensitivity analysis per valutare la presenza di confondimento non misurato.

Controllo del confondimento non misurato

Per quantificare l'impatto del confondimento non misurato, è stato applicato l'**approccio Rule-Out, RO** (si veda la Sezione 4.2 a pag. 40). Lo scopo del RO è quello di quantificare la forza di un confondente non misurato necessaria per spostare l'associazione osservata verso l'unità, ovvero il valore atteso sotto l'ipotesi di mancanza di associazione.

Risultati principali

La stima dell'OR di UGIC associata all'uso corrente di BF (con riferimento ad un uso passato), aggiustata per l'effetto dell'età, dell'utilizzo di farmaci concomitanti e per la presenza di immeasurable time bias, è risultata pari a 0.94 (IC 95% 0.81-1.08).

Rule-Out

Nella nostra applicazione, l'utilizzo del RO è stato generalizzato alle situazioni in cui l'associazione osservata non raggiunge la significatività statistica. L'interesse diventa perciò, quello di identificare la forza di un confondente non misurato, necessaria per rendere positiva e, statisticamente significativa, l'associazione osservata. A tal fine, è stato applicato il metodo utilizzando il valore del limite inferiore dell'intervallo di confidenza (al posto dell'OR osservato), in modo tale da determinare la forza necessaria per "spostare" tale valore al di là dell'unità.

Per fare questo, sono stati considerati due casi ipotetici: (i) il confondente non misurato aumenta fortemente il rischio di UGIC, con un RR che può variare da 1 a 10 (cioè gli esposti a tale fattore hanno un rischio di UGIC fino a 10 volte superiore rispetto ai non esposti a tale confondente); (ii) il con-

fondente è fortemente associato con l'esposizione (uso corrente di BF), con un OR che può variare da 1 a 10 (cioè gli utilizzatori correnti di BF possono essere esposti al confondente fino a 10 volte in più rispetto agli utilizzatori passati).

In entrambi i casi si è assunto che il confondente non misurato sia presente nella popolazione in studio con una prevalenza del 10%, 25% o 50%. I risultati dell'approccio RO sono rappresentati in Figura 7.6.

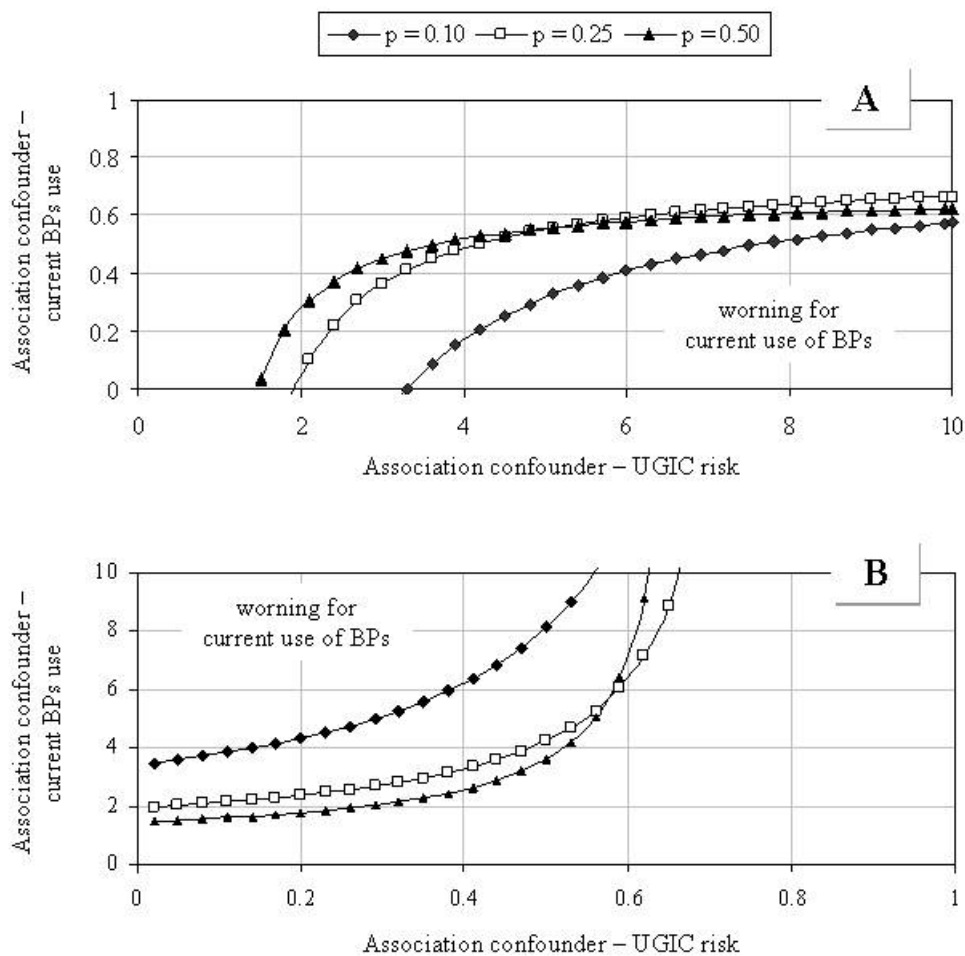


Figura 7.6: Combinazioni delle associazioni confondente – uso corrente di BF (sulle ordinate) e confondente – UGIC (sulle ascisse) necessarie per invertire l'associazione osservata nel caso di un confondente non misurato associato positivamente con l'outcome, per diversi valori della prevalenza del confondente nella popolazione

Si supponga che la prevalenza del confondente non misurato (U) nella popolazione sia del 10%. Se i pazienti esposti a U avessero un rischio di UGIB 4 volte superiore rispetto a quello dei pazienti non esposti a U (Figura 7.6) si dovrebbe osservare un numero di soggetti esposti a U inferiore di 5.8 volte durante l'esposizione corrente a BF per poter invertire il verso dell'associazione osservata (e dunque trovare un'associazione positiva tra uso corrente di BF e UGIC). Una riduzione di circa 2 volte del numero di esposti a U sarebbe necessaria: (i) per un confondente con una prevalenza del 25% o 50%, (ii) per un'associazione confondente-outcome 8-volte più forte. Se questi pazienti avessero un rischio di UGIB 4 volte inferiore rispetto a quello dei pazienti non esposti a U (Figura 7.6) si dovrebbe osservare un incremento di oltre 4.7 volte degli esposti ad U durante l'esposizione corrente a BF per poter invertire il verso dell'associazione osservata. Un leggero aumento nell'esposizione sarebbe richiesto se la prevalenza di U fosse più alta (25% o 50%).

Conclusioni

Da questo studio è emerso che, il rischio di UGIC tra coloro che avevano assunto il farmaco poco prima dell'evento, non era superiore a quello di coloro che l'avevano assunto tempo prima.

Tale risultato, viene confermato anche in seguito all'applicazione del Rule-Out. Infatti, da quest'analisi emerge che, per poter invertire la relazione osservata, dovrebbe esistere un forte fattore di rischio per l'UGIC decisamente sbilanciato all'interno dei livelli di esposizione.

Questo studio è stato pubblicato su *PLOS one* nel 2013 [66].

7.6 Regression Calibration

7.6.1 Antiinfiammatori e Sanguinamento gastrointestinali

L'obiettivo di questo studio era quello di valutare l'associazione tra la durata della terapia con farmaci antiinfiammatori non steroidei (FANS) e insorgenza di Infarto Acuto al Miocardio (AMI) a un anno dall'ingresso dello studio, mediante un approccio di coorte.

Fonte dei dati

I dati utilizzati per questo studio sono stati estratti dal database olandese PHARMO (Contesto C, pag. 62).

Selezione della coorte

Tutti i pazienti che avevano ricevuto almeno una prescrizione di FANS nel periodo 1 Gennaio 2004-31 Dicembre 2008 sono stati selezionati. Da questi pazienti sono stati esclusi tutti coloro che (i) avevano ricevuto una o più prescrizioni di FANS nell'anno precedente l'ingresso, per restringere l'analisi ai soli utilizzatori incidenti, (ii) erano stati ricoverati con diagnosi di AMI o tumore maligno nell'anno precedente l'ingresso, per escludere i soggetti a più alto rischio di AMI e (iii) avevano meno di 18 anni di età. Ogni membro della coorte è stato seguito dalla data di ingresso fino al primo tra i seguenti eventi: ricovero per AMI, morte o fine del follow-up (1 anno dall'ingresso).

Selezione dei casi

I casi sono stati identificati come tutti i pazienti con almeno un ricovero per infarto acuto al miocardio entro 1 anno dall'ingresso nello studio. La data indice è stata definita come la data del primo ricovero per AMI.

Esposizione

Per ogni paziente è stata calcolata la durata della terapia con uno specifico FANS durante il follow-up mediante due modalità: attraverso (i) la Defined

Daily Dose (DDD), quantità generalmente riportata nei DB amministrativi che rappresenta la dose media di un farmaco assunta giornalmente per la sua indicazione principale nell'adulto (in base alla definizione della World Trade Organisation, WHO) e la (ii) Prescribed Daily Dose (PDD), quantità non presente nei DB amministrativi, che rappresenta la dose giornaliera prescritta dal medico. Entrambe le variabili erano disponibili in PHARMO.

Analisi statistica

Per stimare (a livello puntuale e intervallare) l'Odds Ratio (OR) di AMI entro 1 anno associato alla durata della terapia con FANS è stato implementato un modello di regressione logistica (dal momento che tutti i pazienti sono stati seguiti per 1 anno).

Correzione dell'errore di misura

Dal momento che le DDD consentono solamente di approssimare la durata reale della terapia, è possibile che la durata della terapia con FANS, misurata con questo approccio, sia affetta da errore. Per correggere le stime del parametro relativo alla durata dell'esposizione è stata utilizzata la **Regression Calibration** (Sezione 6.2 a pag. 58).

In particolare, l'aggiustamento delle stime è stato effettuato implementando un modello di regressione lineare (nel DB PHARMO) per modellare la relazione tra la durata della terapia con ciascun FANS, calcolata tramite le PDD (vera esposizione), e la durata della terapia calcolata tramite le DDD (quantità affette da errore di misura) ($E(PDD|DDD) = \alpha + \gamma DDD$).

Il coefficiente γ , ottenuto da tale modello, è stato successivamente utilizzato per correggere le stime di interesse, ottenuto implementando il modello di regressione logistica nello studio principale.

Risultati principali

In Tabella 7.2 sono riportate le stime degli OR di AMI (a 1 anno dall'ingresso) associati ad incrementi unitari (in giorni) nella terapia con uno specifico FANS, ottenute da modelli di regressione logistica. Sono inoltre riportate le stime degli OR e dei relativi intervalli di confidenza aggiustati per mezzo

della Regression Calibration.

Dalla Tabella 7.2, si può osservare come la durata della terapia calcolata

FANS	γ	OR	IC 95%	OR _{agg}	IC 95%
Indometacin	1.26	0.93	(0.87-1.00)	0.95	(0.89-1.00)
Sulindac	1.64	0.96	(0.86-1.07)	0.97	(0.91-1.04)
Diclofenac	0.98	0.99	(0.99-1.00)	0.99	(0.99-1.00)
Aceclofenac	1.44	1.04	(1.00-1.07)	1.03	(1.00-1.05)
Piroxicam	1.00	0.99	(0.96-1.02)	0.99	(0.96-1.02)
Meloxicam	1.48	0.99	(0.98-1.01)	1.00	(0.98-1.01)
Ibuprofen	1.22	1.00	(0.99-1.00)	1.00	(0.99-1.00)
Naproxen	0.71	0.99	(0.99-1.00)	0.99	(0.98-1.00)
Ketoprofen	0.94	1.00	(0.97-1.02)	1.00	(0.97-1.022)
Tiaprofenic acid	1.55	0.93	(0.79-1.08)	0.95	(0.86-1.05)
Celecoxib	1.09	1.02	(0.99-1.05)	1.02	(0.99-1.04)
Rofecoxib	1.34	1.00	(0.99-1.02)	1.00	(0.99-1.01)
Etoricoxib	0.88	0.87	(0.63-1.19)	0.85	(0.60-1.22)
Nabumetone	0.95	0.98	(0.96-1.00)	0.98	(0.96-1.00)

Tabella 7.2: Stime degli effetti dei FANS considerati, sul rischio di AMI entro 1 anno dall'ingresso, senza e con correzione per l'errore di misura.

tramite le DDD sottostimi la durata della terapia calcolata con le PDD (ad eccezione dei casi di Diclofenac, Naproxen, Ketoprofen, Etoricoxib e Nabumetone).

In questo caso, tuttavia, l'aggiustamento non produce sostanziali cambiamenti nella stima puntuale e intervallare degli ORs.

Conclusioni

Questo studio mostra come i risultati di studi basati su database amministrativi possano essere distorti a causa dell'errore di misura.

In particolare, nel calcolo della durata di una terapia farmacologica utilizzando le DDD, è importante tenere presente che la reale durata (tramite le

PDD) potrebbe essere molto diversa.

Questo studio è stato presentato al *11^o Consortium Meeting* del Progetto SOS, tenutosi a Milano il 3 Dicembre 2012.

Parte V

Considerazioni
finali

Discussione degli approcci considerati

Questa Tesi aveva come obiettivo l'approfondimento delle diverse tecniche per il controllo del confondimento non misurato e dell'errore di misura, nell'ambito degli studi condotti tramite database amministrativi.

Il controllo del confondimento non misurato è stato affrontato nel Capitolo 3, attraverso l'approfondimento di strategie che permettono di minimizzare tale distorsione mediante il disegno stesso come:

- (a) i disegni user-only (pag. 14)
- (b) i disegni case-only (pag. 16)

e nel Capitolo 4, tramite la descrizione di alcune tecniche analitiche appositamente sviluppate, quali:

- (c) la Monte-Carlo Sensitivity Analysis, MCSA (pag. 35)
- (d) il metodo Rule-Out, RO (pag. 40)
- (e) la Propensity Score Calibration, PSC (pag. 42).

L'applicazione di alcuni tra i metodi sopra-citati a diversi contesti della farmacoepidemiologia, ha reso possibile un'analisi critica delle potenzialità e dei limiti di ciascun approccio utilizzato.

Controllo del confondimento tramite disegno

Il controllo del confondimento tramite il disegno dello studio è una buona scelta qualora non si abbia a disposizione alcuna informazione in merito ai fattori non misurati dal potenziale effetto confondente.

L'implementazione di un disegno user-only o di un case-only, infatti, non richiede assunzioni in merito alla forza dei confondenti o la presenza di uno studio di validazione.

Tale scelta, tuttavia, rappresenta un valido approccio nelle situazioni in cui il campione a disposizione sia sufficientemente numeroso.

Questo poiché, se si decidesse di optare per questa soluzione, sarebbe necessario restringere lo studio a particolari gruppi di pazienti (ad esempio gli esposti al farmaco o i soggetti che sviluppano l'outcome in studio) e, di conseguenza, la potenza dello studio potrebbe risentirne.

Fermo restando che il campione sia numeroso, l'implementazione di un disegno user-only è consigliabile ogniqualevolta si voglia valutare l'effetto (benefico o meno) di un farmaco nella reale pratica clinica e si supponga che il gruppo degli esposti e quello dei non esposti al farmaco non siano confrontabili a causa del confondimento da indicazione, che, come visto nella Sezione 3.1 si genera quando l'indicazione a un trattamento è essa stessa un confondente. Chiaramente, l'indicazione al trattamento non può essere identificata tramite database amministrativi.

I disegni case-only invece, sono da preferire nelle situazioni in cui si suppone possa esistere un forte confondimento da parte di fattori soggetto-specifici non misurati, quali ad esempio lo stile di vita, l'abitudine al fumo, il consumo di alcol, caratteristiche genetiche, ecc...

Bisogna sottolineare che, per poter ottenere la massima efficienza dalle stime ottenute da un case-only, è necessario che l'esposizione di interesse sia di natura transitoria. In caso contrario, è comunque possibile applicare tali disegni ma in questi contesti è necessario prestare maggiore attenzione.

Infatti, se l'esposizione fosse di tipo prolungato ma non cronico, è sufficiente ampliare le finestre a confronto, tenendo conto però che così facendo vi è il

rischio di introdurre qualche distorsione derivata da fattori che variano nel tempo (ad esempio la severità della malattia).

Se invece l'esposizione fosse di tipo cronico allora bisognerebbe accertarsi che vi sia un numero sufficiente di soggetti che risultano esposti solamente in una delle finestre a confronto, dal momento che nella stima dell'OR vengono coinvolte solamente le coppie "discordanti" .

Controllo del confondimento tramite tecniche analitiche

Nelle situazioni in cui si hanno informazioni in merito alle caratteristiche di un ipotetico confondente non misurato (quali ad esempio la prevalenza nella popolazione, la sua forza di associazione con l'outcome in studio, ecc.) o si disponga di uno studio di validazione in cui tale confondente è misurato, la scelta di una tecnica quali i metodi (c)-(e) può rappresentare la scelta migliore.

Tali strategie consentono di aggiustare la stima di interesse per la presenza di confondimento non misurato o di quantificare la forza che tale confondente dovrebbe avere per spiegare completamente l'associazione osservata.

Nel caso in cui si possa assumere che vi sia solamente un confondente non misurato, la MCSA è preferibile alla PSC, in quanto quest'ultima tecnica richiede assunzioni più stringenti, e spesso, difficili da rispettare.

La scelta di applicare una MCSA piuttosto che un RO, dipende dalla tipologia del confondente non misurato che si vuole controllare. Se il confondente presenta più di due modalità allora necessariamente si deve ricorrere a una MCSA, in quanto il RO permette di quantificare solamente la forza di un confondente binario.

Se, al contrario, il confondente non misurato avesse due modalità la scelta tra i due metodi dipende dall'obiettivo che si vuole raggiungere. Il RO non consente di aggiustare la stima osservata (che si presume essere distorta a causa dell'effetto del confondente non misurato), ma di quantificare la forza che questo confondente dovrebbe avere per annullare l'associazione osservata. Viceversa, la MCSA permette di aggiustare la stima di interesse.

D'altro canto, è bene tenere a mente che l'applicazione del RO richiede so-

lamente la conoscenza della prevalenza dell'esposizione nella popolazione P_E (quantità facilmente reperibile in letteratura), della prevalenza del confondente nella popolazione generale P_C (anch'essa facilmente reperibile) e della forza di associazione confondente-outcome RR_{CD} .

L'applicazione della MCSA invece, richiede la conoscenza, oltre di RR_{CD} , della prevalenza del confondente in entrambi i livelli dell'esposizione (P_{C_1} e P_{C_0}). E' facile intuire che questo non è sempre un compito facile.

Infine, se invece si suppone che la stima di interesse ottenuta sia influenzata dall'effetto di più di un confondente non misurato, il metodo più adatto, tra quelli visti, è la PSC.

Affinché venga utilizzata nella maniera più accurata, è indispensabile la presenza di uno studio di validazione che contenga, oltre a tutti i confondenti misurati presenti nello studio principale anche il maggior numero possibile di confondenti non misurati.

Il Capitolo 6, invece, è stato dedicato ai metodi per il controllo dell'errore di misura, in particolare a:

- (e) tecniche per la correzione dei Time-Related Bias (pag. 47)
- (f) la Regression Calibration (pag. 58).

Controllo dell'errore di misura tramite tecniche analitiche

I database amministrativi sono particolarmente vulnerabili ai time-related bias, particolari distorsioni che derivano da un'errata valutazione statica dell'esposizione e dall'impossibilità di identificarla in particolari periodi di tempo. In questi casi, il problema non sta nella scelta del disegno o nella mancanza di variabili ciniche, il problema risiede nella modalità in cui si analizza l'esposizione.

Le tecniche viste nelle Sezione 6.1 rappresentano validi approcci per minimizzare la distorsione derivante da questi contesti.

Inoltre, se si avesse il sospetto che la variabile di esposizione possa essere stata misurata con errore, e si disponesse di uno studio di validazione in cui

tale variabile è misurata correttamente, la scelta della Regression-Calibration dovrebbe essere presa in considerazione.

E' importante sottolineare come questo metodo sia fortemente legato alla natura della variabile di esposizione, in quanto, la sua formulazione originale prevede che l'esposizione sia di tipo continuo.

Conclusione

L'utilizzo di database amministrativi a fini epidemiologici è un fenomeno in continua crescita. Tali database rappresentano uno strumento prezioso e versatile, data l'enorme quantità di informazioni sanitarie contenute al loro interno, relative ad altrettanto ampie popolazioni.

Nonostante tali vantaggi, vi sono alcune limitazioni che possono minare la validità dei risultati ottenuti da essi.

Pertanto, un'accurata valutazione dei potenziali confondenti non misurati o di eventuali errori di misura, e la successiva scelta di un'opportuna strategia di correzione dovrebbero entrare nella routine della pratica del ricercatore che si affaccia al mondo dei database amministrativi.

Bibliografia

- [1] Dieppe P, Bartlett C, Davey P, Doyal L, and Ebrahim S. Balancing benefits and harms: the example of non-steroidal anti-inflammatory drugs. *Br Med J*, 2004;329:31–4.
- [2] Schneeweiss S and Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol*, 2005;58:323–37.
- [3] Fitz-Simon N, Bennett K, and Feely J. A review of studies of adherence with antihypertensive drugs using prescription databases. *Therapeutics and Clinical Risk Management*, 2005;1:93–106.
- [4] Mazzaglia G, Mantovani LG, Sturkenboom MC, Filippi A, Trifirò G, Cricelli C, Brignoli O, and Caputi AP. Patterns of persistence with antihypertensive medications in newly diagnosed hypertensive patients in italy: a retrospective cohort study in primary care. *J Hypertens*, 2005;23:2093–100.
- [5] Van Wijk BLG, Klungel OH, Heerdink ER, and de Boer A. Rate and determinants of 10-year persistence with anti-hypertensive drugs. *J Hypertens*, 2005;23:2101–7.
- [6] Corrao G, Zambon A, Parodi A, Poluzzi E, Baldi I, and Merlino L. Discontinuation of and changes in drug therapy for hypertension among newly treated patients: a population-based study in italy. *J Hypertens*, 2008;26:819–24.

-
- [7] Vrijens B, Vincze G, Kristanto P, Urquhart J, and Burnier M. Adherence to prescribed antihypertensive drug treatments: longitudinal study of electronically compiled dosing histories. *Br Med J*, 2008;336:1114–7.
- [8] Black N. Why we need observational studies to evaluate the effectiveness of health care. *Br Med J*, 1996;312:1215–8.
- [9] Strom BL. *Overview of automated databases in pharmacoepidemiology*. In: *Pharmacoepidemiology (4th edn)*. Wiley, 2005;219–222.
- [10] Muhajarine N, Mustard C, Roos LL, Young K, and Gelskey DE. Comparison of survey and physician claims data for detecting hypertension. *J Clin Epidemiol*, 1997;50:711–8.
- [11] Zhang J, Yun H, Wright NC, Kilgore M, Saag KG, and Delzell E. Potential and pitfalls of using large administrative claims data to study the safety of osteoporosis therapies. *Curr Rheumatol Rep*, 2011;13:273–82.
- [12] Corrao G and Mancia G. Generating evidence from computerized healthcare utilization databases. *Hypertension*, 2015;65:490–8.
- [13] Schneeweiss S and Wang P. Association between ssri use and hip fractures and the effect of residual confounding bias in claims database studies. *J Clin Psychopharmacol*, 2004;24:632–8.
- [14] Hennessy S. Use of health care databases in pharmacoepidemiology. *Basic Clin Pharmacol Toxicol*, 2006;98:311–3.
- [15] Kleinbaum DG, Kupper LL, and Morgenstern H. *Epidemiologic Research: Principles and Quantitative Methods*. Lifetime Learning Publications, 1982.
- [16] Jurek AM, Maldonado G, Greenland S, and Church TR. Exposure-measurement error is frequently ignored when interpreting epidemiologic study results. *Eur J Epidemiol*, 2006;21:871–6.
- [17] Schneeweiss S, Glynn RJ, Tsai EH, Avorn J, and Solomon DH. Adjusting for unmeasured confounders in pharmacoepidemiologic claims

- data using external information: the example of cox2 inhibitors and myocardial infarction. *Epidemiology*, 2005;16:17–24.
- [18] Walker AM. *Observation and Inference*. Epidemiology Resources Inc, 1991;120-124.
- [19] Corrao G, Ghirardi A, Segafredo G, Zambon A, Della Vedova G, Lapi F, Cipriani F, Caputi A, Vaccheri A, Gregori D, Gesuita R, Vestri A, Staniscia T, Mazzaglia G, and Di Bari M. User-only design to assess drug effectiveness in clinical practice: application to bisphosphonates and secondary prevention of fractures. *Pharmacoepidemiol Drug Saf*, 2014;23:859–67.
- [20] Miettinen OS. *Theoretical epidemiology: principles of occurrence research in medicine*. John Wiley and Sons, 1985.
- [21] Nordmann S, Biard L, Ravaud P, Esposito-Farèse M, and Tubach F. Case-only designs in pharmacoepidemiology: a systematic review. *PLoS One*, 2012;7:e49444.
- [22] Maclure M. The case-crossover design: a method for studying transient effects on the risk of acute events. *Am J Epidemiol*, 1991;133:144–53.
- [23] Mittleman MA, Maclure M, and Robins JM. Control sampling strategies for case-crossover studies: an assessment of relative efficiency. *Am J Epidemiol*, 1995;142:91–8.
- [24] Breslow NE and Day NE. Statistical methods in cancer research. vol 2. the design and analysis of cohort studies. *IARC Sci Publ*, 1995:1–406.
- [25] Wang PS, Schneeweiss S, Glynn RJ, Mogun H, and Avorn J. Use of the case-crossover design to study prolonged drug exposures and insidious outcomes. *Ann Epidemiol*, 2004;14:296–303.
- [26] Suissa S. The case-time-control design. *Epidemiology*, 1995;6:248–253.
- [27] Greenland S. Confounding and exposure trends in case-crossover and case-time-control designs. *Epidemiology*, 1996;7:231–9.

- [28] Wang S, Linkletter C, Maclure M, Dore D, Mor V, Buka S, and Wellenius GA. Future cases as present controls to adjust for exposure trend bias in case-only studies. *Epidemiology*, 2011;22:568–74.
- [29] Farrington CP. Relative incidence estimation from case series for vaccine safety evaluation. *Biometrics*, 1995;51:228–35.
- [30] Farrington CP, Nash J, and Miller E. Case series analysis of adverse reactions to vaccines: a comparative evaluation. *Am J Epidemiol*, 1996;143:1165–73.
- [31] Whitaker HJ, Farrington CP, Spiessens B, and Musonda P. Tutorial in biostatistics: the self-controlled case series method. *Stat Med*, 2006;30:1768–97.
- [32] Smeeth L, Donnan PT, and Cook DG. The use of primary care databases: case-control and case-only designs. *Fam Pract*, 2006;23:597–604.
- [33] Farrington P and Hocine MNP. Within-individual dependence in self-controlled case series models for recurrent events. *Appl Statist*, 2010;59:457–75.
- [34] Whitaker HJ, Hocine MN, and Farrington CP. The methodology of self-controlled case series studies. *Stat Methods Med Res*, 2009;18:7–26.
- [35] Farrington CP, Whitaker HJ, and Hocine MN. Case series analysis for censored, perturbed, or curtailed post-event exposures. *Biostatistics*, 2009;10:3–16.
- [36] Farrington CP, Anaya-Izquierdo K, Whitaker HJ, Hocine MN, Douglas I, and Smeeth L. Self-controlled case series analysis with event-dependent observation periods. *J Am Stat Assoc*, 2011;106:417–26.
- [37] Schneeweiss S, Glynn RJ, Tsai EH, Avorn J, and Solomon DH. Adjusting for unmeasured confounders in pharmacoepidemiologic claims data using external information: the example of cox2 inhibitors and myocardial infarction. *Epidemiology*, 2005;16:17–24.

- [38] Cornfield J, Haenszel W, Hammond EC, Lilienfeld AM, Shimkin MB, and Wynder EL. Smoking and lung cancer: recent evidence and a discussion of some questions. *J Natl Cancer Inst*, 1959;22:173–203.
- [39] Steenland K and Greenland S. Monte carlo sensitivity analysis and bayesian analysis of smoking as an unmeasured confounder in a study of silica and lung cancer. *Am J Epidemiol*, 2004;160:384–92.
- [40] Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiol Drug Saf*, 2006;15:291–303.
- [41] Psaty BM, Koepsell TD, Lin D, Weiss NS, Siscovick DS, Rosendaal FR, Pahor M, and Furberg CD. Assessment and control for confounding by indication in observational studies. *J Am Geriatr Soc*, 1999;47:749–54.
- [42] Stürmer T, Schneeweiss S, Avorn J, and Glynn RJ. Adjusting effect estimates for unmeasured confounding with validation data using propensity score calibration. *Am J Epidemiol*, 2005;162:279–89.
- [43] Rosenbaum PR and Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 1983;70:41–55.
- [44] D’Agostino RB Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med*, 1998;17:2265–81.
- [45] Rosner B, Spiegelman D, and Willett WC. Correction of logistic regression relative risk estimates and confidence intervals for measurement error: the case of multiple covariates measured with error. *Am J Epidemiol*, 1990;132:734–45.
- [46] Suissa S and Azoulay L. Metformin and the risk of cancer: time-related biases in observational studies. *Diabetes Care*, 2012;35:2665–73.
- [47] Suissa S. Effectiveness of inhaled corticosteroids in chronic obstructive pulmonary disease: immortal time bias in observational studies. *Am J Respir Crit Care Med*, 2003;168:49–53.

- [48] Samet JM. Measuring the effectiveness of inhaled corticosteroids for copd is not easy! *Am J Respir Crit Care Med*, 2003;168:1–2.
- [49] Suissa S. Immortal time bias in observational studies of drug effects. *Pharmacoepidemiol Drug Saf*, 2007;16:241–9.
- [50] Fisher LD and Lin DY. Time-dependent covariates in the cox proportional-hazards regression model. *Annu Rev Public Health*, 1999;20:145–57.
- [51] Khurana V, Bejjanki HR, Caldito G, and Owens MW. Statins reduce the risk of lung cancer in humans: a large case-control study of us veterans. *Chest*, 2007;131:1282–8.
- [52] Suissa S, Dell’aniello S, Vahey S, and Renoux C. Time-window bias in case-control studies: statins and lung cancer. *Epidemiology*, 2011;22:228–31.
- [53] Suissa S. Immeasurable time bias in observational studies of drug effects on mortality. *Am J Epidemiol*, 2008;168:329–35.
- [54] Nikitovic M, Solomon DH, and Cadarette SM. Methods to examine the impact of compliance to osteoporosis pharmacotherapy on fracture risk: systematic review and recommendations. *Ther Adv Chronic Dis*, 2010;1:149–162.
- [55] Rosner B, Willett WC, and Spiegelman D. Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. *Stat Med*, 1989;8:1051–69.
- [56] Carroll RJ, Ruppert D, Stefanski LA, and Crainiceanu CM. *Measurement Error in Nonlinear Models: A Modern Perspective*. CRC Press, 2006.
- [57] Spiegelman D, McDermott A, and Rosner B. Regression calibration method for correcting measurement-error bias in nutritional epidemiology. *Am J Clin Nutr*, 1997;65:1179S–1186S.

- [58] Carroll RJ and Stefanski LA. Approximate quasilielihood estimation in models with surrogate predictors. *JAMA*, 1990;85:652–663.
- [59] Spiegelman D, Carroll RJ, and Kipnis V. Efficient regression calibration for logistic regression in main study/internal validation study designs with an imperfect reference instrument. *Stat Med*, 2001;20:139–160.
- [60] Freedman LS, Fainberg V, Kipnis V, Midthune D, and Carroll RJ. A new method for dealing with measurement error in explanatory variables of regression models. *Biometrics*, 2004;60:172–81.
- [61] Wang CY, Hsu L, Feng ZD, and Prentice RL. Regression calibration in failure time regression. *Biometrics*, 1997;53:131–45.
- [62] Rudemo M, Ruppert D, and Streibig JC. Random effect models in nonlinear regression with applications to bioassay. *Biometrics*, 1989;45:349–362.
- [63] Brown JP and Josse RG. Scientific advisory council of the osteoporosis society of canada. 2002 clinical practice guidelines for the diagnosis and management of osteoporosis in canada. *CMAJ*, 2002;167:S1–34.
- [64] Ghirardi A, Segafredo G, Zambon A, Arfè A, Ibrahim I, Nicotra F, Romio S, Scotti L, Soranna D, and Corrao G. Utilizzo di antidepressivi e rischio di frattura. *VII Congresso Nazionale della Società Italiana di Statistica Medica ed Epidemiologia Clinica (SISMEC)*, Roma, 25-28 Settembre 2013.
- [65] Ghirardi A, Di Bari M, Zambon A, Scotti L, Della Vedova G, Lapi F, Cipriani F, Caputi AP, Vaccheri A, Gregori D, Gesuita R, Vestri A, Staniscia T, Mazzaglia G, and Corrao G; BEST investigators. Effectiveness of oral bisphosphonates for primary prevention of osteoporotic fractures: evidence from the aifa-best observational study. *Eur J Clin Pharmacol*, 2014;70:1129–37.
- [66] Ghirardi A, Scotti L, Zambon A, Della Vedova G, Cavalieri D'oro L, Lapi F, Cipriani F, Caputi AP, Vaccheri A, Gregori D, Gesuita R, Vestri A,

Staniscia T, Mazzaglia G, and Corrao G; BEST Investigators. Risk of severe upper gastrointestinal complications among oral bisphosphonate users. *PLoS One*, 2013;8:e73159.