

Sulphonylureas and risk of cardiovascular disease: systematic review and meta-analysis

CRITICAL APPRAISAL

Autori:

A
Battaglia¹

A
Donzelli²

M Font³

1. MMG,
Verona

2. Specialista
in Igiene e in
Scienza della
Alimentazione
Esperto di
Sanità
Pubblica.
Milano.

3. Farmacista
ULSS 20
Verona.

METANALISI DI STUDI OSSERVAZIONALI: RICHIEDONO POOLING SEPARATI

La table 1 dell' articolo riporta in misura aggregata sotto forma di Risk Ratio

- (1) i risultati ottenuti metanalizzando tutti gli studi (RCT + NRS)
- (2) i risultati ottenuti metanalizzando solo i RCT
- (3) i risultati ottenuti metanalizzando gli studi osservazionali (NRS)
- (4) i risultati ottenuti metanalizzando solo gli studi osservazionali di coorte (NRS cohort)
- (5) i risultati ottenuti metanalizzando solo gli studi osservazionali caso-controllo (NRS case-control)

La scelta degli autori di considerare nello stesso pooling studi osservazionali e sperimentali (punto 1) è assai discutibile [Higgins 2011].

La Cochrane Collaboration sconsiglia infatti in modo esplicito di sottoporre a “pooling misti” dati osservazionali e dati sperimentali, perché sono rispettivamente generati da ricerche caratterizzate da rischi di bias assai diversi [Higgins 2011]. Un bias infatti può condizionare sia la direzione che la dimensione dei risultati, rendendo di conseguenza inopportuni i pooling misti.

È da notare tra l' altro che sono proprio i risultati di queste analisi a dar maggior peso alle conclusioni di Pungh e coll.. Il pooling di tutti gli studi (RCT + cohort + case-control) e il pooling di tutti gli studi osservazionali (cohort + case-control) hanno fornito infatti risultati significativi o al limite della significatività statistica a sfavore dell' intervento per tutti gli end-point ad eccezione dello stroke (table 1 della metanalisi). Il pooling delle evidenze sperimentali (RCT) ha invece prodotto risultati non significativi per tutti gli end-point, con direzione dell' effetto a sfavore dell' intervento solo per la morte cvd e per lo stroke (table 1 della metanalisi).

Negli studi NRS la mancanza di randomizzazione comporta sbilanciamento tra i due bracci in caratteristiche note e ignote (=fattori prognostici, confounder) in grado di influire sull' outcome in modo indipendente dall' intervento studiato. Opportune tecniche statistiche (matching, regressione multivariata, propensity score..) vengono di solito utilizzate negli studi NRS per eseguire adeguati aggiustamenti ma ciò vale solo –ovviamente– per le fonti di bias di cui sia nota l' esistenza. Un RCT presenta al contrario l' enorme vantaggio di non richiedere alcun aggiustamento perché tutti i confounders noti e ignoti vengono equamente bilanciati al momento della randomizzazione. Data l' enorme differenza nel rischio di bias gli esperti Cochrane sconsigliano fermamente di considerare nello stesso pooling studi randomizzati e non randomizzati per guidare i risultati verso un'unica misura di efficacia [Higgins 2011].

La scelta degli autori di considerare nello stesso pooling studi osservazionali con disegno diverso (punto 3) è, parimenti, assai discutibile [Higgins 2011].

L'opportunità di eseguire o no metanalisi di studi osservazionali è stata oggetto di lunghi dibattiti.

Alcuni autori sconsigliano di sottoporre a pooling studi diversi dagli RCT, perché la mancanza di randomizzazione comporta elevato rischio di distorsioni e possibilità di risultati aggregati ‘biased’.

La Cochrane Collaboration non sconsiglia in modo esplicito l'esecuzione di metanalisi di studi osservazionali, ma vincola l' opportunità di queste alla mancata disponibilità in letteratura di dati sperimentali (RCT) o a situazioni in cui per ragioni da soppesare di volta in volta (esempio: qualora l'outcome sia rappresentato da eventi avversi rari) i disegni sperimentali non siano in grado di offrire informazioni sufficienti [Higgins 2011].

Una metanalisi dovrebbe presentare in queste particolari circostanze risultati osservazionali e sperimentali in analisi rigorosamente separate [Higgins 2011]. Un problema legato alle metanalisi di studi osservazionali è costituito inoltre dalla eterogeneità di questi disegni e a una loro diversa propensione al bias. Uno studio di coorte è strutturato ad esempio in modo prospettico sulla rilevazione di un determinato end-point in soggetti esposti o non esposti ad un determinato fattore di rischio, mentre uno studio caso controllo è strutturato in modo retrospettivo sulla rilevazione di una pregressa esposizione al rischio in soggetti affetti o non affetti da un determinato end-point.

Una metanalisi di studi osservazionali dovrebbe pertanto presentare analisi rigorosamente separate per ciascuno di questi disegni [Higgins 2011]: vuoi per la loro diversità strutturale, vuoi per la diversità di approccio, vuoi per una propensione al bias che può essere molto diversa da disegno a disegno. Per esempio -almeno per gli studi di etiologia- uno studio caso-controllo si considera molto più esposto al bias rispetto ad uno studio di coorte e, d'altra parte, uno studio di coorte basato sull' analisi di dati storici risulta molto più esposto al bias rispetto ad uno studio di coorte basato sull' analisi di dati incidenti.

Una metanalisi dovrebbe considerare, in generale, gli studi non randomizzati solo in queste circostanze [Higgins 2011]

- a) in assenza di evidenze di qualità migliore (RCT)
- b) per studiare eventi difficilmente intercettati da uno studio RCT
 - a. effetti collaterali rari
 - b. effetti inattesi del trattamento
 - c. effetti del trattamento che richiedono lunghi periodi di latenza
- c) presentando metanalisi separate qualora siano disponibili anche RCT
- d) presentando metanalisi separate in presenza di NRS caratterizzati da disegni diversi

In quest'ottica abbiamo sufficienti ragioni di considerare potenzialmente 'biased' i risultati metanalitici presentati dagli autori ai punti (1) e (3).

Phung e coll non hanno tenuto conto delle direttive Cochrane :infatti non solo hanno inserito nello stesso pooling studi osservazionali e studi sperimentali (pooling "all studies") ma hanno anche inserito nello stesso pooling studi osservazionali caratterizzati da disegni diversi (pooling "observational only") .

In quest'ottica non terremo in alcuna considerazione i risultati presentati da Phung e coll come frutto di pooling "misti" ma limiteremo la nostra analisi ai risultati stratificati correttamente secondo la tipologia del disegno ("randomized controlled trial only", "cohort only", "case-control only" della tabella 1).dalla table 1.

Occorre dire che il vero obiettivo degli autori era, in ogni caso, inserire nello stesso pooling studi non randomizzati (NRS) e studi sperimentali (RCT).

Nelle metanalisi di dati aggregati (dove l'unità di analisi è il trial) quando l'endpoint è un evento i calcoli possono essere eseguiti in due diverse maniere.

Il primo metodo (di prima scelta nelle metanalisi di RCT) segue un approccio frequentistico e utilizza i dati grezzi riportati dai trial : vale a dire il numero di pazienti che hanno e che non hanno subito l' evento in ciascuno dei due bracci .

Il secondo metodo (utile quando gli autori non riportano i dati grezzi) utilizza invece le misure di efficacia riportate dai trial assieme al loro errore standard (calcolabile , se non pubblicato, dagli intervalli di confidenza della misura di efficacia o dal suo valore di P.).

In una metanalisi di studi osservazionali utilizzare il primo metodo (quello basato sull' analisi dei dati grezzi) significa rinunciare nel pooling ad ogni aggiustamento messo in essere dai singoli autori. Utilizzare il secondo metodo implica invece considerare misure di efficacia già in qualche modo "aggiustate" e ha un suo razionale (=attenuare per quanto possibile le distorsioni dei risultati del pooling inevitabilmente connesse con la natura osservazionale dei trial inseriti nel pool). Ovviamente questa scelta ha comunque dei limiti, in quanto non è detto in presenza di rischio simile di bias che due autori abbiano aggiustato i propri risultati utilizzando la stessa tecnica e programmando lo stesso tipo di errore.

Phung e coll hanno adottato il secondo metodo, ma al solo scopo di poter sottoporre a pooling tutti i risultati prodotti dalla letteratura esaminata.

MISURE DI EFFICACIA RELATIVA: IL LORO SIGNIFICATO È DIVERSO

Gli autori hanno eseguito "pooling misti" metanalizzando risultati di efficacia espressi indifferentemente in Hazard Ratio, Odds Ratio o Risk Ratio e presupponendo quindi , implicitamente, l'equivalenza clinica di queste tre misure. La **web tabella A** illustra il significato clinico di Risk Ratio (RR), Odds Ratio (OR) e Hazard Ratio (HR): tutti sono costruite su rapporti, e in tal senso rappresentano sicuramente misure relative.

Però mentre il calcolo di Risk Ratio e di Odds Ratio rientra in un approccio di analisi frequentistico (vale a dire considera solo il numero dei pazienti che alla fine dello studio hanno o non hanno subito un end-point), l'Hazard Ratio introduce un' altra dimensione , quella temporale, in quanto confronta la pendenza di due curve di sopravvivenza : l' approccio è completamente diverso, come è completamente diversa la natura di ciò che viene confrontato tra i bracci.

Per esempio

Un RR=2 indica che in base ai risultati della sperimentazione la probabilità di subire l'evento nei braccio di intervento è risultata doppia rispetto alla probabilità rilevata nel braccio di controllo.

Un OR=2 indica che in base ai risultati della sperimentazione l'odds di subire l' evento nei braccio di intervento è risultato doppio rispetto all'odds rilevato nel braccio di controllo.

Un HR=2 in un determinato punto del follow-up dice che nel tempuscolo successivo la probabilità di avere l'evento tra i pazienti che non l' hanno ancora avuto è doppia nei pazienti assegnati al braccio di intervento rispetto al braccio di controllo. Il confronto tra due curve presuppone proporzionalità dei rischi, ossia che i tassi istantanei dell' evento

varino nei due bracci con il tempo in modo proporzionale; in ogni punto del follow-up pertanto i valori di Hazard Ratio rimarranno costanti.

Considerando solo le due misure frequentistiche (Risk Ratio e Odds Ratio) la somiglianza o meno dei loro valori risulta tra l'altro fortemente condizionata dalla probabilità dell'evento: solo in presenza di eventi molto rari (es: nelle sperimentazioni sui vaccini) l'Odds Ratio ha lo stesso valore di un Risk Ratio.

Nella maggior parte dei trial che considerano endpoint cardiovascolari l'Odds Ratio tende invece a produrre risultati molto diversi dal Risk Ratio in quanto l'evento studiato non è rarissimo: se si attribuisse ad un Odds Ratio il significato clinico di un Risk Ratio si incorrerebbe in una sicura sovrastima della differenza rilevata tra i due bracci.

Anche le proprietà matematiche di queste misure risultano profondamente diverse.

Odds Ratio può assumere infatti qualunque valore compreso tra zero e infinito mentre Risk Ratio può essere compreso solo tra zero e $1/[\text{Rischio Assoluto dell'evento nel braccio di controllo}]$ [Higgins 2011].

Odds Ratio presenta, ancora, un comportamento coerente con le modalità con cui viene espresso l'end-point. Per esempio la mortalità rilevata in un trial può essere espressa indifferentemente sia come odds di morte che come odds di vita: invertendo l'end-point potrebbe sembrare legittimo aspettarsi dalla misura di efficacia un comportamento clinicamente coerente 'vale a dire l'inversione dei suoi valori.

Questo è vero solo per l'Odds Ratio ma non per il Risk Ratio.

Un $OR=0.7$ indica infatti che nel braccio di intervento l'odds di morte è il 70% rispetto all'odds di morte del braccio di controllo: il corrispondente l'Odds Ratio di vita corrisponde esattamente all'inverso dell'Odds Ratio di morte: ossia a $OR=1/0.7=1.42$.

Se questi calcoli venissero fatti confrontando i due bracci con un Risk Ratio anziché con un Odds Ratio l'inversione delle modalità con cui viene definito l'end-point comporterebbe comportamenti incoerenti e imprevedibili di questa misura di efficacia.

La scelta degli autori di Phung e coll., vale a dire considerare le tre misure relative di efficacia clinicamente equivalenti e adatte ad essere sottoposte ad un unico pooling appare pertanto assolutamente discutibile e questo può essere chiarito anche con un esempio operativo.

Una delle ricerche considerate dalla metanalisi è appunto lo studio UKPDS 33 [UKPDS 1998]. Questo RCT ha confrontato in una casistica di diabetici in un follow-up di 10 anni l'efficacia di un trattamento ipoglicemizzante intensivo ($n=2729$) con quella di un trattamento convenzionale ($n=1138$) per l'end-point composito "any diabetes related end-point". Nel braccio di intervento sono stati registrati 563 eventi, nel braccio di controllo 438. I due bracci sono stati confrontati utilizzando un Hazard Ratio e il lavoro pubblicato riporta i valori $HR=0.88$ (0.79-0.99).

Esprimendo l'efficacia rilevata per lo stesso end-point con approccio frequentistico possiamo calcolare valori di Risk Ratio corrispondenti a $RR=0.54$ (0.48-0.59) e valori di Odds Ratio corrispondenti a $OR=0.42$ (0.36-0.49). Per un identico fenomeno i valori delle tre misure relative di efficacia risultano quindi drammaticamente diversi l'uno dall'altro. Nella condizione estrema in cui tre trial assolutamente uguali ma con risultati di efficacia espressi in modo diverso venissero sottoposti a metanalisi, considerare 'clinicamente equivalenti' tre misure di efficacia diverse esprimenti lo stesso fenomeno porterebbe a risultati imprevedibili e disastrosi.

Dai dati pubblicati da Phung e coll. nelle etable 4 ed etable 5 risulta che 45 confronti sono stati espressi in Risk Ratio, 24 in Odds Ratio e 74 in Hazard Ratio (**web tabella B**).

I risultati espressi in Risk Ratio sono stati quasi tutti calcolati partendo dai dati grezzi degli studi RCT (un solo risultato è stato espresso in adjusted Risk Ratio in uno studio di coorte).

Calcolare in sede metanalitica una misura di efficacia partendo dai dati crudi è una opzione legittima quando si tratti di studi randomizzati e controllati, perché la randomizzazione offre garanzia assoluta sul bilanciamento iniziale dei bracci a confronto e -almeno in teoria- non dovrebbe comportare necessità di aggiustamenti per confounders.

Simile approccio non vale per i dati osservazionali, per i quali in sede metanalitica vale la regola di non ricalcolare le misure di efficacia trial specifiche partendo dai dati grezzi perché si annullerebbero con questa procedura gli aggiustamenti messi in essere dagli stessi autori dei trial.

Piuttosto (vedi sopra) in una metanalisi di studi osservazionali occorre sottoporre a pooling le misure di efficacia già aggiustate riportate dai singoli autori [Higgins 2011] sfruttando opportune tecniche di pooling che prevedono l'utilizzo dell'errore standard delle misure aggiustate.

Gli autori della metanalisi hanno seguito correttamente questo percorso ma solo al fine di sottoporre (erroneamente) ad un unico pooling dati osservazionali e dati sperimentali (RCT).

Se avessero scelto di esprimere i risultati sperimentali in odds ratio, cosa facilmente consentita dalla disponibilità dei dati grezzi degli RCT, si sarebbe posto il problema di come confrontare i risultati espressi in Odds ratio con quelli degli studi di coorte, tutti tranne uno espressi in Hazard Ratio.

Per bypassare il problema gli autori l' hanno semplicemente ignorato e per la metanalisi hanno utilizzato gli Odds Ratio e gli Hazard Ratio come fossero Risk Ratio.

Considerare il significato clinico di un Hazard Ratio equivalente a quello di un Risk Ratio è una assunzione forte, ma da alcuni viene tollerata [Borenstein 2009]. Il problema più grosso è invece considerare un Odds Ratio alla stessa stregua di un Risk Ratio : questo assunto come si diceva non vale in presenza di analisi caratterizzate da eventi non rari (come appunto quelle pertinenti end-point cardiovascolari), in cui per uno stesso risultato i valori delle due misure sono molto diverse. In queste situazioni infatti i valori di Odds ratio sopravvalutano costantemente la dimensione dell' effetto espressa da un Risk Ratio [Borenstein 2009, Higgins 2011].

Gli autori riconoscono questi limiti ma riportano: “most studies reported hazard ratios, with very few reporting odds ratios, so our results would not likely be overestimated”.

Questa affermazione è però falsa, in quanto se è vero che gli studi che riportano i risultati in Odds Ratio sono solo quattro studi caso-controllo è altrettanto vero che gli autori hanno completamente ignorato il problema della interdipendenza tra i singoli confronti (spesso multipli entro un singolo trial) utilizzandoli nella metanalisi come fossero originati da trial indipendenti (per i commenti su questa procedura vedi oltre).

Infatti le analisi usate nel pooling in veste di Odds Ratio non sono solo quattro, ma risultano in realtà 24 (vale a dire rappresentano il 17% del totale).

Ciò può aver ovviamente contribuito ad ipertrofizzare i risultati pertinenti sia il pool “all studies” che il pool “observational only” illustrati nella metanalisi

STUDI OSSERVAZIONALI: NUMEROSITÀ CAMPIONARIA E LUNGHEZZA DEL FOLLOWUP NON GIUSTIFICANO DA SOLI DARE PRIORITÀ AI LORO RISULTATI

Phung e coll sottolineano i vantaggi della letteratura osservazionale.

La grande numerosità campionaria di questi studi (di solito più elevata rispetto a quella degli RCT) apporta infatti all' analisi una maggior potenza statistica, consentendo di rilevare con maggior sicurezza differenze tra i due bracci non dovute al caso.

Allo stesso tempo i follow-up di più lunga durata degli studi osservazionali consente di intercettare con maggior efficacia endpoint che non sarebbero rilevabili nei periodi di follow-up solitamente più brevi degli studi RCT.

Un terzo vantaggio è rappresentato dalla diversità del setting: gli studi sperimentali sono caratterizzati da casistiche selezionate e motivate, mentre gli studi osservazionali da casistiche miste e più pertinenti condizioni ‘real life’.

Gli svantaggi della letteratura osservazionale (vedi sopra) consistono però in una maggior propensione al bias e in un elevato rischio di interpretazioni distorte.

Vediamo i primi due vantaggi

Circa la maggior potenza statistica dell' analisi, Phung e coll. avrebbero buoni motivi per enfatizzare la virtù degli studi osservazionali solo se gli studi sperimentali avessero dimostrato risultati nella stessa direzione ma in assenza della significatività statistica rilevata nelle analisi osservazionali.

Infatti abbiamo calcolato che in questo scenario teorico la letteratura sperimentale avrebbe offerto risultati gravati da una quantità non accettabile errore beta (=oltre 20%) (**web tabella C**) .

In realtà grazie a importanti lacune della metanalisi (vedi punto successivo) i risultati riportati dagli autori per gli studi sperimentali non sono quelli veri (vedi oltre) : qui basti ricordare che in tutti i casi i risultati sperimentali sono molto meno soggetti a bias e sono pertanto decisamente più attendibili rispetto ai dati offerti dalla letteratura osservazionale sugli stessi end-point.

Circa la seconda virtù osservazionale enfatizzata dagli autori (la lunghezza del periodo di osservazione di queste ricerche), ci sarebbero stati fondati motivi per accentuarne l' importanza qualora la lunghezza del follow-up negli RCT considerati risultasse palesemente insufficiente a intercettare gli eventi cardiovascolari studiati dalla metanalisi. Occorre allora ricordare che nella maggior parte degli studi randomizzati e controllati di intervento dove sono stati studiati end-point cardiovascolari in presenza di differenze significative del confronto la pendenza delle curve di

sopravvivenza che esprimevano l' andamento del fenomeno nei due bracci cominciava a divergere con significatività statistica già dopo i primi 24 mesi di follow-up (vedi ad esempio lo studio HOPE [HOPE 2000]).

Abbiamo pertanto condotto una analisi per sottogruppi atta a verificare per ogni end-point l' esistenza o meno di differenze importanti tra i risultati aggregati degli studi RCT caratterizzati da durata di follow-up inferiore a 2 anni e i risultati degli studi RCT caratterizzati da durata di follow-up superiore o uguale a 2 anni. In questo tipo di analisi l' eventuale differenza tra i risultati rilevati in sottogruppi diversi viene espressa da un "test for interaction", in cui viene testata l' ipotesi nulla "i risultati dei sottogruppi sono uguali".

Nel nostro caso la non significatività del test for interaction, dimostrata per tutti gli end-point, ha ragionevolmente escluso la presenza di un 'effetto sottogruppo', ossia che la durata del follow-up sia stata in grado di condizionare il risultato dell' analisi (web tabella D, grafici 1-5).

GLI AUTORI HANNO TRASCURATO O SOTTORIPORTATO IMPORTANTI ENDPOINT

Un' altra considerazione riguarda l' incompleta raccolta di dati messa in essere dagli autori della metanalisi (**web tabelle L->V**).

Phung e coll non hanno infatti, prima di tutto, considerato l' outcome "morte da tutte le cause", che era in realtà accessibile in tutti i dati pubblicati dagli RCT considerati dalla metanalisi. La mortalità generale rappresenta invece in una analisi di safety una importantissima proxy di Serious Adverse Events (SAEs) [Therapeutic Initiative 2003] e quindi di "pericolosità" associata al trattamento.

Se lo scopo della metanalisi era mettere in guardia sugli eventi avversi associati all' utilizzo delle sulfaniluree, questo end-point costituiva sicuramente il più importante outcome da analizzare.

Abbiamo raccolto pertanto autonomamente i dati di mortalità riportati dai trial (**web tabella N**): la nostra metaanalisi ha dimostrato assoluta mancanza di differenze tra i due bracci in termini di mortalità generale [RR=0.98 (0.86-1.12)] (**webtabella I, web grafico 2**).

Phung e coll. hanno inoltre ignorato i dati pubblicati di mortalità cardiovascolare (**web tabella M**) riportati dallo studio di Kahan, che riporta palesemente il numero di morti per infarto [Kahan 2006] e, soprattutto, i dati di mortalità cardiovascolare riportati dallo studio UKPDS 33 [UKPDS 1998]. Gli autori dello studio UKPDS (un trial con follow-up di 10 anni) hanno in realtà pubblicato in forma disaggregata il numero di decessi per infarto miocardico, morte improvvisa, stroke e arteriopatia periferica rilevati nella loro ricerca. Essendo i decessi, notoriamente, end-point mutuamente esclusivi gli autori avrebbero potuto seguire la nostra procedura essere sommando tranquillamente il loro numero per costruire l' outcome composito "cvd death".

L' inserimento nella metanalisi dei dati di Kahan e soprattutto dello studio UKPDS ha rovesciato la direzione dei risultati riportati da Phung e coll: dall' incremento del rischio segnalato dalla metanalisi (+ 22%) si passa a una riduzione del rischio pari al 5%. (**webtabella I, web grafico 1**)

Infine gli autori non hanno riportato i dati sulle ospedalizzazioni pubblicate da Giles [Giles 2008]: anche questa omissione ha condizionato l' esito del pooling: inserendo nel pooling i dati di Giles dalla riduzione non significativa del rischio segnalata dagli autori (-21%) si passa a una *riduzione significativa* del rischio pari a - 36% (**web tabella I, web grafico 5**).

L' END POINT "CARDIOVASCULAR COMPOSITE" NON HA SENSO CLINICO

La definizione di questo outcome varia da studio a studio e alcuni trial hanno considerato più di un end-point: in tal caso la scelta degli autori è caduta arbitrariamente su uno solo, senza alcuna giustificazione addotta a sostegno della singola scelta (**web tabella H**).

Pertanto consideriamo assolutamente inaffidabile l' interpretazione clinica dei risultati pertinenti questo end-point, che non verrà più considerato nella nostra discussione.

GLI AUTORI NON HANNO CORRETTO LE ANALISI PER CONFRONTI INTERDIPENDENTI

In diversi modi (di seguito sinteticamente descritti) gli autori hanno erroneamente trattato alla stregua di analisi indipendenti analisi che riguardavano in realtà -almeno in parte- gli stessi pazienti. Tale procedura comporta il rischio di un eccesso di errore alfa, ossia aumenta il rischio di considerare erroneamente 'reali' risultati in realtà dovuti al caso.

Innanzitutto gli autori hanno costruito a tavolino i casi di infarto e stroke sommando gli eventi fatali e non fatali.

Questa procedura è scorretta e implica il rischio di analisi non indipendenti perché casi di infarto o di stroke 'fatali' e 'non fatali' non rappresentano end-point mutuamente esclusivi: un paziente ad esempio può subire due infarti o due stroke, il primo non fatale e il secondo fatale.

Abbiamo quindi scorporato tutti i dati raccolti rispettivamente da trial che riportavano per infarto e stroke i dati aggregati (= casi fatali + casi non fatali) (**web tabella O**, **web tabella R**) o, rispettivamente, i casi fatali (**web tabella P**, **web tabella S**) o i casi non fatali (**web tabella Q**, **web tabella T**) dei singoli due end-point.

Un altro esempio è rappresentato dalle modalità con cui gli autori hanno gestito i dati estratti da trial caratterizzati da più bracci a confronto: ciascun confronto è stato considerato nel pooling alla stessa stregua del risultato di una ricerca indipendente (**web tabella E**).

Un problema frequente in una metanalisi è appunto rappresentato da come organizzare il pooling in presenza di trial caratterizzati da più di due bracci (es: più interventi versus un identico comparator).

La **web tabella E** illustra per ogni endpoint il rapporto tra i trial coinvolti dall'analisi e il numero di analisi in inserite nel pooling.

Questa anomalia riguarda nel lavoro di Phung et al anche e soprattutto le analisi eseguite sui risultati degli studi osservazionali

Non è corretto inserire nel pooling i singoli confronti come se fossero singoli trial (come è stato fatto invece nella metanalisi di Phung) perchè in tal caso i calcoli sono gravati da “unit of analysis error”, con un eccesso di errore alfa associato alla mancata interdipendenza tra un confronto e l'altro [Higgins 2011].

Per esempio nel trial UKPDS esistono tre bracci due bracci: clorpropamide, glibenclamide e cure abituali.

Cio' permette di analizzare in coerenza all'obiettivo della metanalisi due possibili confronti: il confronto clorpropamide versus cure abituali e il confronto glibenclamide versus cure abituali.

Se la metanalisi viene eseguita inserendo nel pooling ciascuno di questi confronti come se si trattasse di dati recuperati da due trial diversi, l'errore prodotto sarebbe collegato al fatto che i due confronti non sono analisi indipendenti in quanto prevedono come controlli gli stessi soggetti (vale a dire i pazienti allocati alle cure abituali).

Il grado di associazione tra diversi confronti è statisticamente misurabile e viene espresso da un coefficiente di correlazione (web tabella F).

In caso di outcome rappresentati da misure relative (Risk Ratio, Odds ratio, ecc) per ragioni matematiche occorre calcolare r dalla versione logaritmica di queste misure.

Nei trial Kahn e UKPDS (i due RCT caratterizzati da più bracci a confronto) il valore di r da noi calcolati per i vari end-point sono illustrati nella **web tabella G**

Non tener conto del grado di interdipendenza tra le varie analisi in presenza di più bracci entro uno stesso trial equivale implicitamente a considerare indipendenti i singoli confronti assumendo $r=0$.

Se in condizioni reali esiste una correlazione positiva tra un confronto e l'altro ($r>0$) l'errore prodotto consiste in una sottostima della varianza (con risultati imprecisi).

Se invece esiste una correlazione positiva ($r<0$) errore prodotto consiste in una sovrastima della varianza [Borenstein 2009].

Per superare in una metanalisi il problema dell'interdipendenza dei confronti multipli nel contesto di uno stesso trial una delle soluzioni proposte [higgins 2011, Borenstein 2009] è calcolare per il trial una misura media dell'efficacia rilevata entro i singoli confronti e una varianza intorno alla media che tenga conto non solo della varianza tipica di ciascun confronto ma anche della misura in cui singoli confronti risultano correlati [Borenstein 2009 Higgins 2011]. Questi risultati trial specifici potranno essere quindi utilizzati per i calcoli metanalitici.

Abbiamo pertanto ripetuto i calcoli metanalitici che si riferiscono ai trial randomizzati e controllati seguendo questo approccio (**web tabella G**, **web grafici 1→5**).

Nei trial dove esistevano confronti multipli (esprimevano analisi non indipendenti perché caratterizzate da un braccio in comune), vale a dire i trial UKPDS e Kahan, abbiamo espresso i risultati per ogni end-point in un pooled log Risk Ratio trial-specifico il cui Errore Standard è stato aggiustato per il grado di interdipendenza tra i confronti non indipendenti considerati nel trial.

Per due confronti non indipendenti abbiamo seguito queste tappe.

In primis abbiamo espresso il grado di dipendenza tra i confronti calcolando un coefficiente di correlazione (r).

Quindi abbiamo calcolato un risultato aggregato sottoponendo a pooling i due confronti interdipendenti calcolando la varianza in modo da tener conto sia della varianza di ciascun confronto che del grado di interdipendenza tra i due confronti.

Per calcolare il coefficiente di correlazione (r) abbiamo calcolato un log-odds ratio che rappresentasse il confronto tra i due confronti calcolandolo dalla differenza dei rispettivi log-odds ratio [Glenny 2005]: quindi abbiamo convertito il log-odds ratio in effect size (d) e, successivamente, in coefficiente di correlazione (r) [Borenstein 2009].

Per esprimere in forma aggregata il risultato pooled dei due confronti abbiamo calcolato in primis per ciascun confronto un effect size (d) partendo dal rispettivo log-odds ratio [Borenstein 2009] quindi abbiamo calcolato la media dei due effect size: quest' ultima esprimeva il risultato 'pooled' dei due confronti.

La varianza intorno a questa media è stata calcolata tenendo conto delle varianze di ciascun confronto e del grado di interdipendenza tra i due confronti (espresso da r) utilizzando la tecnica descritta da Borenstein [Borenstein 2009]. Le misure aggregate (d e varianza di d) espresse in effect size sono state quindi convertite in log-odds ratio e in standard error del log-odds ratio [Borenstein 2009]. Abbiamo quindi calcolato i valori esponenziali di log-odds ratio e di standard error del log-odds ratio per esprimere i risultati in Odds Ratio, ottenendo la stima puntuale e calcolando gli intervalli di confidenza in forma non logaritmica. Cio' ha consentito nella tappa successiva di calcolare un Risk Ratio con i rispettivi intervalli di confidenza attraverso la formula $RR = OR/[1-CER*(1-OR)]$ [Higgins 2011]. Il CER (Control Event rate) è stato calcolato in forma 'pooled' sottoponendo a metanalisi fixed effect based i valori di CER dei due confronti interdipendenti. Abbiamo quindi calcolato lo standard error del Risk Ratio dagli intervalli di confidenza. Infine, il logaritmo del Risk Ratio e il logaritmo del suo errore standard sono stati utilizzati per esprimere nella metanalisi finale (in forma quindi aggregata e aggiustata per il grado di interdipendenza tra i due confronti) il risultato pooled dei due confronti interdipendenti.

La metanalisi finale è stata eseguita con tecnica di Wolff (Inverse variance method) utilizzando (come Phung e coll) un modello a effetti random [Higgins 2011, Borenstein 2009].

Abbiamo in questo modo corretto gli errori commessi dagli autori nel pooling degli studi gli studi sperimentali aggiustando i calcoli per il grado di interdipendenza tra i vari confronti .

I nostri risultati (**web tabella I, web grafici 1→5**) dimostrano che per nessun end-point l' esposizione a sulfaniluree comporta un aumento significativo del rischio. Inoltre per tutti gli end-point (ad eccezione dello stroke) la direzione dell' effetto è sempre a favore del trattamento. Circa i risultati sullo stroke segnaliamo un aumento non significativo del rischio pari più o meno a un terzo (+5%) rispetto a quello segnalato dalla metanalisi (+14%) ma per questo end-point qualsiasi considerazione è resa inaffidabile dalla presenza di publication bias, non dimostrato per tutti gli altri outcome considerati dalla ricerca (**web grafico 9**) .

Il pooling degli studi di coorte ha evidenziato un incremento significativo del rischio di morte cardiovascolare [RR=1.26 (1.18-1.34)] e di ospedalizzazioni [RR=1.31 (1.08-1.58)], mentre gli studi caso-controllo hanno evidenziato un incremento significativo del rischio di infarto [RR=1.30 (1.09-1.54)].

Questi risultati osservazionali aggregati sono inquietanti e meriterebbero un approfondimento. Tuttavia i risultati osservazionali soffrono sicuramente delle stesse conseguenze degli errori di metodo sopra segnalati.

Per esempio (web tabella E), per l'outcome cvd death gli autori hanno estratto i dati da 9 studi di coorte, ma hanno inserito nel pooling un numero di confronti più che doppio (20) perchè alcuni di questi studi avevano previsto più bracci.

Non abbiamo potuto correggere i risultati osservazionali per il grado di interdipendenza tra i vari confronti in quanto i dati pubblicati dai singoli trial non erano sufficienti a questa analisi.

Una correzione dei risultati per i confronti multipli sarebbe stata possibile -almeno approssimativamente- qualora tutte le ricerche avessero pubblicato i tassi degli eventi rilevati nei singoli bracci.

Purtroppo solo una parte degli studi di coorte ha reso disponibili queste informazioni: per eseguire un pooling metodologicamente corretto occorrerebbe quindi contattare i singoli autori delle ricerche osservazionali per recuperare le informazioni mancanti, cosa assolutamente normale in una metanalisi ma che Phung e coll non hanno fatto.

GLI AUTORI NON HANNO SEGUITO L' APPROCCIO INTENSION TO TREAT

Abbiamo verificato che per molte analisi gli autori non hanno raccolto i dati secondo il principio Intention To Treat, in quanto nelle tabelle pubblicate in web hanno riportato non il numero di pazienti randomizzati al trattamento, ma piuttosto il numero di pazienti che hanno iniziato il trattamento (web tabella L)

Nelle nostre rielaborazioni dei risultati sperimentali (**web tabella G, web grafici 1→5**) abbiamo corretto anche questo errore.

CONCLUSIONI

La metanalisi di Phung e coll è gravata da errori di metodo che rendono inaffidabili le conclusioni degli autori. Le nostre conclusioni si basano sui risultati dell'analisi eseguita dopo le correzioni degli errori metodologici e delle lacune della metanalisi originale.

Questo studio ha mescolato dati osservazionali e dati sperimentali; ha considerato clinicamente equivalenti misure di efficacia caratterizzate in realtà da significati diversi; ha sottoposto a pooling confronti caratterizzati da identico comparator alla stessa stregua di ricerche indipendenti; ha ignorato importanti dati di mortalità generale e cardiovascolare e ha riportato in modo incompleto anche il numero dei ricoveri.

La nostra rielaborazione dei risultati sperimentali ha permesso di dimostrare

- I soggetti esposti alle sulfaniluree non incorrono in un maggior rischio di morte cardiovascolare (abbiamo anzi rilevato una riduzione non significativa del rischio pari a -5% mentre il pooling degli RCT eseguito dagli autori segnala per questo end-point un aumento non significativo pari a +22%)
- I soggetti esposti alle sulfaniluree non incorrono in un maggior rischio di morte da ogni causa (abbiamo rilevato per questo end-point una riduzione non significativa di rischio pari a -2%)
- I soggetti esposti alle sulfaniluree presentano un rischio di infarto miocardico *minore* rispetto a quello rilevabile nei soggetti non esposti (-15%) con risultati *al limite della significatività statistica*
- I soggetti esposti alle sulfaniluree presentano un rischio di ospedalizzazioni da causa cardiovascolare *significativamente minore* rispetto a quello rilevabile nei soggetti non esposti (-36%)

Non consideriamo in alcun modo utilizzabili le informazioni prodotte dalla metanalisi sull'endpoint composito cardiovascolare data la polimorfa definizione di questo end-point, che rende clinicamente non interpretabili i risultati del pooling.

La metanalisi riporta per gli studi osservazionali risultati aggregati apertamente divergenti rispetto a quelli prodotti dalla correzione dei dati sperimentali.

Il pooling degli studi di coorte riporta infatti aumento significativi del rischio sia per la mortalità cardiovascolare (+26%), per le ospedalizzazioni (+31%).

Il pooling degli studi caso-controllo riporta incremento significativi anche del rischio di infarto (+30%).

I dati pubblicati da Phung e coll ci hanno permesso di ricostruire e correggere i dati prodotti per i risultati sperimentali ma la stessa procedura non è applicabile all'analisi critica dei risultati osservazionali.

Infatti i dati pubblicati dagli autori delle singole ricerche non sono sufficienti a calcolare il grado di interdipendenza tra i confronti multipli largamente utilizzati nella metanalisi di Phung e coll.

Una metanalisi ben condotta dei contributi di questa letteratura osservazionale dovrebbe coinvolgere anche gli autori di queste ricerche finalizzata a conoscere anche i dati non pubblicati necessari all'analisi.

In attesa di ulteriori sviluppi di questo filone di analisi a nostro parere il lettore dovrebbe limitarsi a prendere atto dei risultati sperimentali, che non sono a sostegno di un incremento di rischio cardiovascolare associato all'uso delle sulfaniluree

BIBLIOGRAFIA

- Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to Meta-Analysis. Wiley 2009 ISBN 978-0-470-05724-7
- Duerden 2009. What are hazard ratios? http://www.medicines.ox.ac.uk/bandolier/painres/download/whatis/what_are_haz_ratios.pdf
- Higgins JPT and Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 <http://www.cochrane.org/training/cochrane-handbook>
- Giles TD et al. Pioglitazone and heart failure: results from a controlled study in patients with type 2 diabetes mellitus and systolic dysfunction. J Card Fail 2008; 14: 445-452.
- Glenny AM et al. Indirect comparisons of competing interventions. Health Technology Assessment 2005 Vol 9 n.26
- HOPE The Heart Outcomes Prevention Evaluation Study Investigators
Effects of an Angiotensin-Converting-Enzyme Inhibitor, Ramipril, on Cardiovascular Events in High-Risk Patients
N Engl J Med 2000; 342:145-153
- Phung OJ et al. Sulphonylureas and risk of cardiovascular disease: systematic review and meta-analysis. Diabetic Medicine 2013
DOI: 10.1111/dme.12232
- Therapeutics Initiative – Evidence Based drug Therapy. Do Statins have a Role in Primary Prevention? April - May - June 2003
- UK Prospective Diabetes Study (UKPDS) Group* Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33)
Lancet 1998; 352: 837-53

TABELLE E GRAFICI

Web tabella A				
MISURE DI EFFICACIA RELATIVA E LORO SIGNIFICATO				
Misura di efficacia	Come vengono confrontati i due bracci	Che cosa si misura in ciascun braccio	Che cosa esprime clinicamente?	Variabili coinvolte nella costruzione della misura di efficacia
Risk Ratio	Calcolando il rapporto tra Rischio assoluto dell' evento nel braccio di intervento e Rischio assoluto dell' evento nel braccio di controllo	Rischio Assoluto dell' evento = probabilita' dell' evento nel braccio= numero pazienti che hanno subito l' evento / totale pazienti nel braccio	Quanto è diverso il rischio nel braccio di intervento rispetto al braccio di controllo : esprime la frazione di rischio basale rimasta dopo l' intervento	Numero di pazienti
Odds Ratio	Calcolando il rapporto tra odds dell' evento nel braccio di intervento e odds dell' evento nel braccio di controllo	Odds dell' evento nel braccio = numero pazienti che hanno subito l'evento / numero pazienti che non hanno subito l'evento nel braccio	Quanto è diverso l' odds nel braccio di intervento rispetto al braccio di controllo : esprime la frazione di odds basale rimasta dopo l' intervento	Numero di pazienti
Hazard Ratio	Calcolando il rapporto tra tasso istantaneo dell' evento nel braccio di intervento e tasso istantaneo dell' evento nel braccio di controllo. Presuppone che i tassi istantanei delle due curve varino con il tempo in modo costantemente proporzionale	Hazard Rate = tasso istantaneo dell' evento= numero di eventi / numero di persone a rischio dell' evento nel tempuscolo (t) nel braccio.	Quanto è diversa la pendenza della curva di sopravvivenza del braccio di intervento rispetto al braccio di controllo: esprime la frazione di hazard rate basale rimasta dopo l' intervento	Numero di pazienti Tempo
<p>Quando il confronto tra due bracci viene espresso da una misura di efficacia relativa (ossia come Risk Ratio, Odds ratio o Hazard Ratio) un valore superiore a 1 indica che nel braccio di intervento la frequenza dell' evento è maggiore rispetto a quella rilevata nel braccio di controllo mentre un valore inferiore a 1 indica l' opposto. Gli intervalli di confidenza esprimono il livello di precisione di queste stime: intervalli ampi sono associati a una stima imprecisa mentre intervalli stretti esprimono una stima precisa. Per queste misure di efficacia relativa quando l' intervallo di confidenza comprende il valore di 1 il risultato va definito "non significativo" in quanto non può essere esclusa in tal caso l' ipotesi nulla caratterizzata da identica frequenza dell' end-point nei due bracci .</p>				

Web tabella B						
MISURE DI EFFICACIA RELATIVA UTILIZZATE NELLA METANALISI						
	Cvd Death	CVD composite	All IM	All Stroke	CVD hospitalizations	tot
Numero di confronti espressi in Risk Ratio	7 calcolati da RCT	11 calcolati da RCT	10 calcolati da RCT 1 cohort studies	9 calcolati da RCT	4 calcolati da RCT 3 cohort studies	45
Numero di confronti espressi in Odds Ratio	-	4 case-control studies	20 case-control studies	-	-	24
Numero di confronti espressi in Hazard Ratio	26 cohort studies	34 cohort studies	10 cohort studies	4 cohort studies	-	74
TOTALE	33	49	41	13	7	143

Web tabella C

POTENZA STATISTICA POSSEDUTA DAL POOL DI STUDI SPERIMENTALI QUALORA AVESSE DOVUTO CONFERMARE I RISULTATI DEGLI STUDI OSSERVAZIONALI

	n RCT	Risultato rilevato dal pooling Observational only *=significativo	Potenza statistica del pooling degli RCT nel confermare il risultato rilevato per gli studi osservazionali
Cardiovascular death	7	RR=1.26*	Power RCT=0.592
Myocardial Infarction	10	RR=1.20*	Power RCT=0.384
Stroke	9	RR=1.05	Power RCT=0.053
Hospitalization for CHD causes	4	RR=1.31*	Power RCT=0.249

Si definisce errore beta la probabilità di sostenere erroneamente 'dovuto al caso' un risultato che invece è 'reale'. L'errore beta è complementare alla potenza statistica (power), che consiste nella capacità dell'analisi di dimostrare l'esistenza di risultati 'reali' quando questi effettivamente esistono (errore beta = 1-power).

Convenzionalmente si considerano 'accettabili' valori di errore beta non superiori al 20% (0.20) e, coerentemente, valori di potenza statistica non inferiori all'80% (1-0.20=0.8).

Web tabella D

ANALISI PER SOTTOGRUPPI STRATIFICATI PER LUNGHEZZA DEL FOLLOW-UP

	Follow up < 2 anni	Follow-up ≥ 2 anni	Test for interaction (subgroup effect)
Cardiovascular death	RR=1.05 (0.47-2.32)	Rr=0.94 (0.78-1.14)	P=0.79
Death	RR=1.03 (0.54-1.95)	Rr=0.98 (0.86-1.12)	P=0.89
Myocardial Infarction	RR=0.67 (0.20-2.21)	RR=0.84 (0.71-1.00)	P=0.70
Stroke	RR=0.19 (0.02-1.67)	RR=1.13 (0.89-1.44)	P=0.11
Hospitalization for CHD causes	RR=0.64 (0.42-0.98)	-	-

Web tabella E

CONFRONTO TRA IL NUMERO DI TRIAL E IL NUMERO DI ANALISI PERTINENTI I SINGOLI POOLING

	Cvd Death	CVD composite	All IM	All Stroke	CVD hospitalizations
Numero di trial pertinenti il pooling	7 (RCT) 9 (cohort) 0 (case-control)	9 (RCT) 11 (cohort) 1 (case-control)	6 (RCT) 5 (cohort) 3 (case-control)	5 (RCT) 2 (cohort) 0 (case-control)	4 (RCT) 2 (cohort) 0 (case-control)
Numero di confronti inseriti dagli autori nel pooling come fossero analisi indipendenti	7 (RCT) 20 (cohort) 0 (case-control)	11 (RCT) 28 (cohort) 4 (case-control)	10 (RCT) 11 (cohort) 4 (case-control)	9 (RCT) 4 (cohort) 0 (case-control)	4 (RCT) 3 (cohort) 0 (case-control)

Web tabella F

CALCOLO DEL COEFFICIENTE DI CORRELAZIONE TRA ANALISI NON INDIPENDENTI

La procedura per il calcolo del coefficiente di correlazione e'

$$d = \logor * \left(\frac{\sqrt{a}}{n} \right)$$

$$r = \frac{d}{\sqrt{(d^2 + a)}}$$

$$a = \frac{(n1 + n2)^2}{n1 * n2}$$

dove

logor = log odds ratio

d = effect size

a = fattore di correzione

r = coefficiente di correlazione

La procedura per il calcolo di un pooled odds ratio da due analisi interdipendenti e'

$$\text{pooled } d = \frac{\sum d}{2}$$

$$V \logor = 1/a + 1/b + 1/c + 1/d \quad (a \ b \ c \ d \text{ sono i numeri della tabella di contingenza})$$

$$Vd = V \logor * \frac{a}{n^2} = V1 \text{ e } V2 \text{ nei due sottogruppi interdipendenti}$$

$$\text{pooled } Vd = \frac{1}{4} * (V1 + V2 + 2r * \sqrt{V1} * \sqrt{V2})$$

$$\text{pooled } \logor = \text{pooled } d * \frac{n}{\sqrt{a}}$$

$$V \text{ pooled } \logor = Vd * \frac{n^2}{a}$$

$$SE \text{ pooled } \logor = \sqrt{(V \text{ pooled } \logor)}$$

dove

pooled d = media dei due effect size

V logor = varianza del log odds ratio

Vd = varianza effect size = *V1, V2* = varianza effect size gruppo 1 e gruppo 2

pooled Vd = varianza della media dei due effect size corretta per il grado di correlazione tra i due gruppi

V pooled logor = varianza del pooled logor

SE pooled logor = standard error del pooled logor

Web Tabella G

**GRADO DI INTERDIPENDENZA TRA DIVERSI TIPI DI CONFRONTO RILEVATO NEGLI RCT
CARATTERIZZATI DA PIÙ DI DUE BRACCI**

trial	confronti	Cvd Death	Death	All IM	Fatal IM	Non Fatal IM	All Stroke	Fatal Stroke	Non Fatal Stroke
kahn	Glibenclamide vs metformina	r= -0.0003	r= 0.0256	nd	r= -0.00037	r= 0.0483	r=-.0483	nd	nd
	Glibenclamide vs rosiglitazone								
UKP DS 33	Clorpropamide vs conventional	r=0 .0067	r=0.0377	r= 0.03165	r=-.0560	r= 0.03979	r=-.0919	r= -0.1629	r= -.07848
	Glibenclamide vs conventional								

Web tabella H

DEFINIZIONI DI CARDIOVASCULAR COMPOSITE ENDPOINT

* = scelto dagli autori della metanalisi

Gerstein 2010	-morte + infarto non fatale + stroke non fatale + rivascolarizzazioni coronariche + ospedalizzazioni per ischemia cardiaca (*) -morte + infarto non fatale + stroke non fatale
Seino 2010	-disturbi cardiaci (*)
Giles 2008	-morte cardiovascolare + ospedalizzazione per scompenso cardiaco + visite al pronto soccorso per scompenso cardiaco (*)
Niessen 2008	-morte cardiovascolare + infarto non fatale + stroke non fatale -morte cardiovascolare + infarto non fatale + stroke non fatale + ospedalizzazione per angina instabile o scompenso cardiaco (*) -morte cardiovascolare + infarto non fatale + stroke non fatale + rivascolarizzazioni carotidee + rivascolarizzazioni coronariche + ospedalizzazione per angina instabile o scompenso cardiaco
Jain 2006	-coronaropatia (non ulteriormente definita) o+ infarto (*)
Mazzone 2006	-morte cardiovascolare + infarto non fatale + stroke non fatale (*) -morte cardiovascolare + infarto non fatale + stroke non fatale + rivascolarizzazione coronarica + endoarteriectomia + stenting carotideo + ospedalizzazione per angina instabile + ospedalizzazione per scompenso cardiaco
Kahan 2006	Infarto miocardico + scompenso cardiaco + stroke
StJohn Sutton 2002	-eventi avversi associati al cuore (*)
Hermann 1994	-Sintomi cardiovascolari severi (*)

Web tabella I

CONFRONTO TRA I RISULTATI PUBBLICATI PER IL POOLING DEGLI RCT E I NOSTRI RISULTATI: MOTIVI DI DIVERGENZA

	Riportati dalla metanalisi	Ricalcolati da noi	Motivi della divergenza
Cardiovascular death	RR=1.22 (0.63-2.39)	RR=0.95 (0.78-1.15)	Non hanno inserito i dati UKPDS 1998 e Kahn 2006
Death	-	RR=0.98 (0.86-1.12)	Outcome importante non considerato dalla metanalisi
Myocardial Infarction	RR=0.88 (0.74-1.05)	RR=0.84 (0.71-1.00)	Non hanno corretto per confronti multipli
Stroke	RR=1.14 (0.86-1.51)	RR=1.05 (0.73-1.52)	Non hanno corretto per confronti multipli
Hospitalization for CHD causes	RR=0.79 (0.48-1.29)	RR=0.64 (0.42-0.98)	Non hanno inserito i dati Giles 2008

ANALISI DEI DATI GREZZI PERTINENTI GLI STUDI SPERIMENTALI: DETTAGLI DELLE CORREZIONI DA NOI APPORTATE

Le prossime tabelle illustrano le correzioni da noi effettuate paragonando i dati pubblicati dai singoli trial con i dati pubblicati dalle web tabelle della metanalisi di Phung & coll

Web Tabella L

NUMERO DI PAZIENTI ASSEGNATI AI BRACCI DELLE RICERCHE: DATI CORRETTI IN BASE AL PRINCIPIO INTENTION TO TREAT

trial	anno	sulfanilurea	comparator	randomizzati a sulfanilurea	randomizzati a comparator	note
gerstein	2010	glipizide	rosiglitazone	339	333	
giles_1	2010	glibenclamide	pioglitazone	149	151	
seino	2010	glibenclamide	liraglutide	139	272	nella metanalisi riportano 132/268 (il n. pazienti che hanno ricevuto l' intervento dopo la randomizzazione)
seck	2010	glipizide	sitagliptin	584	588	nella metanalisi riportano 248/256 (numero di pazienti inseriti in una analisi finale per protocollo)
giles_2	2008	glibenclamide	pioglitazone	256	262	
niessen	2008	glimepiride	pioglitazone	273	274	nella metanalisi riportano 272/270 (il n. pazienti che hanno ricevuto l' intervento dopo la randomizzazione)
jain	2006	glibenclamide	pioglitazone	251	251	
mazzone	2006	glimepiride	pioglitazone	230	232	nella metanalisi riportano 228/230 (il n. pazienti che hanno ricevuto l' intervento dopo la randomizzazione)
kahan_a	2006	glibenclamide	metformina	1441	1454	
kahan_b	2006	glibenclamide	rosiglitazone	1441	1456	
st_john_sutton	2002	glibenclamide	rosiglitazone	99	104	
hermann_a	1994	glibenclamide	metformina	34	38	Nella metanalisi riportano 19/19 Non è chiaro da dove abbiano ricavati questi dati (non risultano nel lavoro pubblicato). Una nota di richiamo (c) rimanda a uno script incomprensibile "Baseline results reported per study year, none reported for overall cohort"
hermann_b	1994	glibenclamide	glibenclamide_metformina	34	72	Nella metanalisi riportano 19/46 Non è chiaro da dove abbiano ricavati questi dati (non risultano nel lavoro pubblicato). Una nota di richiamo (c) rimanda a uno script incomprensibile "Baseline results reported per study year, none reported for overall cohort"
ukpds33_A	1998	chlorpropamide	conventional	619	896	
ukpds33_B	1998	glibenclamide	conventional	615	896	

Web tabella M						
MORTALITÀ CARDIOVASCOLARE: DATI CORRETTI						
trial	anno	Morti cvd sulfanilurea	N sulfanilurea	Morti cvd comparator	N comparator	note
gerstein	2010	3	339	4	333	
giles_1	2010	2	149	0	151	
seck	2010	2	584	0	588	Nell' analisi viene utilizzato dagli autori un approccio ITT (contrariamente alla scelta indicata nella web tabella L)
giles_2	2008	6	256	5	262	
niessen	2008	1	273	3	274	Nell' analisi riportano 1/273 e 3/270 (coerentemente alla scelta indicata dalla web tabella L)
jain	2006	1	251	0	251	
mazzone	2006	0	230	0	232	Nell' analisi riportano 0/228 e 0/230 (coerentemente alla scelta indicata dalla web tabella L)
kahan_a	2006	3	1441	2	1454	Per il trial Khan non viene riportato alcun evento, mentre in realtà il lavoro pubblicato riporta la mortalità per infarto miocardico (quella indicata a fianco)
kahan_b	2006	3	1441	2	1456	idem
ukpds33_A	1998	71	619	110	896	UKPDS non riporta le morti cardiovascolari in forma aggregata ma il loro numero è facilmente calcolabile sommando i decessi per infarto miocardico, per morte improvvisa, per stroke e per arteriopatia periferica, che sono end-point mutuamente esclusivi: per dlorpropamide la somma è: $56+6+9+0=71$
ukpds33_B	1998	69	615	110	896	UKPDS non riporta le morti cardiovascolari in forma aggregata ma il loro numero è facilmente calcolabile sommando i decessi per infarto miocardico, per morte improvvisa, per stroke e per arteriopatia periferica, che sono end-point mutuamente esclusivi: per glibenclamide la somma è: $46+7+16+0=69$

Per il trial Khan non viene riportato alcun evento, mentre in realtà il lavoro pubblicato riporta la mortalità per infarto miocardico.

Ma il problema più grande è senz'altro rappresentato dalla omissione nella metanalisi dei dati di mortalità cardiovascolare dello studio UKPDS. Questo trial non riporta -in effetti- le morti cardiovascolari ma il loro numero è facilmente calcolabile sommando i decessi per infarto miocardico, per morte improvvisa, per stroke e per arteriopatia periferica, che lo studio UKPDS riporta e che ovviamente sono end-point mutuamente esclusivi.

Web tabella N						
MORTALITÀ DA OGNI CAUSA: DATI CORRETTI						
trial	anno	Morti sulfanilurea	N sulfanilurea	Morti comparator	N comparator	note
gerstein	2010	7	339	8	333	
giles_1	2010	2	149	0	151	Il trial riporta i morti da causa CV (non fa cenno ad altre cause di decesso)
seino	2010	0	139	1	272	
seck	2010	8	584	1	588	
giles_2	2008	6	256	5	262	Il trial riporta i morti da causa CV (non fa cenno ad altre cause di decesso)
niessen	2008	2	273	3	274	
jain	2006	2	251	0	251	
mazzone	2006	0	230	1	232	
kahan_a	2006	31	1441	31	1454	
kahan_b	2006	31	1441	34	1456	
ukpds33_A	1998	136	619	190	896	
ukpds33_B	1998	121	615	190	896	

La metanalisi non riporta questo importante end-point, nonostante molti trial contengano l'informazione.

La mortalità cardiovascolare è una ottima proxy di SAES ed è un end-point difficilmente misclassificabile, talchè viene raccomandata la sua analisi nelle revisioni sistematiche della letteratura pertinenti end point cardiovascolari [Therapeutics Initiative 2003].

Web tabella O						
INFARTO (FATALE O NON FATALE, DATI AGGREGATI) : DATI CORRETTI						
trial	anno	Infarto (fatale o non n fatale) sulfanilurea	N sulfanilurea	Infarto (fatale o non n fatale) comparator	N comparator	note
Gerstein	2010	nd	339	nd	333	La metanalisi riporta la somma degli eventi fatali e non fatali nei due bracci (7/339 e 8/333). Questo approccio è errato perchè gli eventi non sono mutuamente esclusivi
seino	2010	0	139	1	272	La metanalisi non riporta infarti in forma aggregata; in realtà il lavoro originale riporta un infarto miocardico nel braccio liraglutide
giles_2	2008	2	256	4	262	
jain	2006	2	251	2	251	la metanalisi non riporta infarti in forma aggregata; in realtà la tabella 3 del lavoro originale riporta i casi di infarto segnalati in tabella
ukpds33_A	1998	100	619	162	896	la metanalisi non riporta infarti in forma aggregata; in realtà la tabella 3 del lavoro originale riporta i casi di infarto segnalati in tabella. Nella metanalisi gli autori sottopongono a pooling come fossero trial separati i casi di infarto fatale associati alla sulfanilurea e al comparator nel confronto clorpropamide/conventional, i casi di infarto non fatale associati alla sulfanilurea e al comparator nel confronto clorpropamide/conventional, i casi di infarto fatale associati alla sulfanilurea e al comparator nel confronto glibenclamide/conventional, i casi di infarto non fatale associati alla sulfanilurea e al comparator nel confronto glibenclamide/conventional.
ukpds33_B	1998	90	615	162	896	idem

Nella dati riportati dalla etable 4 del materiale web della metanalisi (Cardiovascular Endpoints) esistono parecchie incongruenze.

Abbiamo considerato innanzitutto i trial che riportavano l' outcome come infarto miocardico fatale o non fatale in forma aggregata.

Per lo studio di Gerstein gli autori hanno sommato gli infarti fatali e quelli non fatali (è una incongruenza perchè non necessariamente i due end-point sono mutuamente esclusivi).

Per lo studio di Jain gli autori non riportano eventi, mentre la tabella 3 del lavoro originale riporta i casi di infarto

Per lo studio UKPDS gli autori metanalizzano come trial separati i casi di infarto fatale associati alla sulfanilurea e al comparator nel confronto clorpropamide/conventional, i casi di infarto non fatale associati alla sulfanilurea e al comparator nel confronto clorpropamide/conventional, i casi di infarto fatale associati alla sulfanilurea e al comparator nel confronto glibenclamide/conventional, i casi di infarto non fatale associati alla sulfanilurea e al comparator nel confronto glibenclamide/conventional.

Tale procedura non è corretta per varie ragioni:

a) incrementa il rischio di errore alfa associato a confronti multipli (è un rischio inevitabile nelle metanalisi di trial con più bracci , ma con la procedura seguita dagli autori il trial viene inserito nel pooling quadruplicato e il rischio è pertanto doppio rispetto a quello che si incorre di solito duplicando i dati di un trial a tre bracci con un unico comparator). Il pooling non è stato fatto (vedi testo) tenendo conto degli aggiustamenti necessari alla correzione della interdipendenza tra i confronti

b) l' infarto fatale e l' infarto non fatale non sono end-point mutuamente esclusivi e questa scelta comporta ancor più il rischio di coinvolgere più volte nei calcoli gli stessi pazienti

c) una attenta lettura dello studio UKPDS dimostra che i dati aggregati sull' infarto (fatale + non fatale) in realtà sono a disposizione in forma aggregata già nel lavoro originale.

Web tabella P

INFARTO FATALE: DATI CORRETTI

trial	anno	Infarto fatale sulfanilurea	N sulfanilurea	Infarto fatale comparator	N comparator	note
gerstein	2010	1	339	1	333	
seck	2010	1	584	0	588	la metanalisi non riporta eventi fatali in realtà vengono riportate nel trial due morti (una improvvisa, l' altra attribuita a infarto nel gruppo sulfanilurea)
kahan_a	2006	3	1441	2	1454	
kahan_b	2006	3	1441	2	1456	
ukpds33_A	1998	56	619	80	896	
ukpds33_B	1998	46	615	80	896	

Web tabella Q

INFARTO NON FATALE : DATI CORRETTI

trial	anno	Infarto non fatale sulfanilurea	N sulfanilurea	Infarto non fatale comparator	N comparator	note
gerstein	2010	6	339	7	333	
niessen	2008	4	273	2	274	La metanalisi riporta 4/273 e 2/270, coerentemente alla scelta descritta alla web tabella L 1
mazzone	2006	1	230	0	232	La metanalisi riporta 1/228 e 0/230, coerentemente alla scelta descritta alla web tabella L
kahan_a	2006	15	1441	21	1454	
kahan_b	2006	15	1441	25	1456	
ukpds33_A	1998	53	619	87	896	La metanalisi riporta un valore errato nel braccio sulfaniluree (59/619)
ukpds33_A	1998	46	615	87	896	

Web tabella R

STROKE FATALE E NON FATALE : DATI CORRETTI

trial	anno	stroke sulfanilurea	N sulfanilurea	stroke comparator	N comparator	Note
gerstein	2010	1	339	5	333	
kahan_a	2006	17	1441	19	1454	
kahan_b	2006	17	1441	16	1456	
						la metanalisi non riporta stroke in forma aggregata ; in realtà la tabella 3 del lavoro originale riposta i casi segnalati in tabella. Nella metanalisi gli autori sottopongono a pooling come fossero trial separati i casi di strohe fatale associati alla sulfanilurea e al comparator nel confronto clorpropamide/conventional, i casi di stroke non fatale associati alla sulfanilurea e al comparator nel confronto clorpropamide/conventional, i casi di sttoke fatale associati alla sulfanilurea e al comparator nel confronto glibenclamide/conventional, i casi di stroke non fatale associati alla sulfanilurea e al comparator nel confronto glibenclamide/conventional.
ukpds33_A	1998	33	619	47	896	glibenclamide/conventional.
ukpds33_B	1998	45	615	47	896	idem

Solo pochi trial riportano in forma aggregata i dati sullo stroke fatale e non fatale. Riportiamo nella tabella i trial che riportano l' outcome in forma aggregata .

Gli autori hanno seguito per questo end-point la stessa scelta adottata per l' outcome infarto (quadruplicando i dati del trial UKPDS, vedi sopra). Tale scelta è discutibile per incremento di errore alfa e per misclassificazione dell' outcome (vedi sopra) in quanto non necessariamente i casi fatali o non fatali sono mutuamente esclusivi

Web tabella S

STROKE FATALE: DATI CORRETTI

trial	anno	Stroke fatale sulfanilurea	N sulfanilurea	Stroke fatale comparator	N comparator	Note
gerstein	2010	0	339	0	333	
ukpds33_A	1998	9	619	12	896	
ukpds33_B	1998	16	615	12	896	

L' unico trial che riporta eventi per questo end-point è l' UKPDS, dove il risultato pooled dei due sottogruppi glibenclamide versus conventional therapy e rispettivamente clorpropamide versus conventional therapy corretto per l' interdipendenza dei due confronti (vedi sopra) produce un risultato caratterizzati da un incremento non significativo del 45% del rischio di questo end-point. (**web grafico 4bis**)

Queste conclusioni non sono generalizzabili in quanto sono prodotte da un solo trial e non è possibile pertanto una analisi formale del rischio di publication bias.

Web tabella T						
STROKE NON FATALE: DATI CORRETTI						
trial	anno	Stroke non fatale sulfanilurea	N sulfanilurea	Stroke non fatale comparator	N comparator	Note
gerstein	2010	1	339	5	333	
niessen	2008	1	273	0	274	La metanalisi riporta 1/273 e 0/270, coerentemente alla scelta descritta in web tabella L
mazzone	2006	1	230	0	232	La metanalisi riporta 1/228 e 0/230, coerentemente alla scelta descritta in web tabella L
ukpds33_A	1998	26	619	38	896	
ukpds33_B	1998	34	615	38	896	La metanalisi riporta un dato errato nel braccio sulfanilurea (64/615)

Web tabella U	
OUTCOME OSPEDALIZZAZIONE PER MOTIVI CARDIOVASCOLARI	
*Endpoint scelto arbitrariamente dagli autori per la metanalisi	
trial	Definizioni delle cause di ospedalizzazione per motivi CV
Gerstein 2010	-Ricoveri per ischemia miocardica *
Giles 2010	-Ricoveri per eventi cardiovascolari *
Giles 2008	-Ricoveri per scompenso cardiaco
Niessen 2008	-Ricoveri per angina instabile * -Ricoveri per scompenso cardiaco
Mazzone 2006	-Ricoveri per angina instabile * -Ricoveri per scompenso cardiaco

La definizione di questo end-point è eterogenea.

Nei trial in cui esistevano dati per più di un end-point (diverse cause CV di ospedalizzazione) gli autori hanno optato per le cause ischemiche; si tratta di una scelta arbitraria e non coerente con la definizione dell' end-point. Abbiamo pertanto eseguito la metanalisi inserendo anche i dati di Giles 2008 (**web grafico 10**)

Web tabella V						
OSPEDALIZZAZIONI DA CAUSE CV: DATI CORRETTI						
trial	anno	ospedalizzazioni sulfanilurea	N sulfanilurea	ospedalizzazioni comparator	N comparator	Note
gerstein	2010	7	339	11	333	
giles_1	2010	13	149	13	151	
giles_2	2008	12	256	26	262	La metanalisi non riporta 12 + 26 ricoveri per CHF che invece il lavoro originale pubblica
niessen	2008	2	273	4	274	La metanalisi riporta 2/273 e 4/270, coerentemente con la scelta illustrata alla web tabella L per il denominatore
mazzone	2006	0	230	0	232	La metanalisi riporta 1/228 e 0/230, coerentemente con la scelta illustrata alla web tabella L per il denominatore

Web tabella Z(a)

DIVERGENZE TRA I DATI RIPORTATI DA PHUNG & COLL PER I RCT E I RISULTATI DEL NOSTRO CRITICAL APPRAISAL

	Morte cardiovascolare	Morte da ogni causa	Infarto miocardico (tutti I tipi)	Infarto miocardico fatale	Infarto miocardico Non fatale
Conclusioni della metanalisi di Phung	Aumento non significativo del 22% del rischio di morte CVD RR=1.22 (0.63-2.39)	Endpoint non valutato	Riduzione non significativa del 12% del rischio di infarto RR=0.88 (0.74-1.05)	Endpoint non valutato	Endpoint non valutato
Conclusioni del nostro critical appraisal	Riduzione non significativa del 5% del rischio di morte CVD 0.95 (0.78-1.15)	Riduzione non significativa del 2% rischio di morte da ogni causa RR=0.98 (0.86-1.12)	Riduzione del rischio di infarto del 16% al limite della significativita' RR=0.84 (0.71-1.00)	Riduzione del rischio di infarto fatale del 6% in assenza di significativita' RR=0.94 (0.74-1.20)	Riduzione del rischio di infarto non fatale del del20% in assenza di significativita' RR=0.80 (0.65-0.99)
Note	Phung & coll non hanno considerato i dati dello studio UKPDS 33 (non riportati dal lavoro originale ma calcolabili) e non hanno considerato nei calcoli l' interdipendenza dei confronti nei trial Kahan e UKPDS Per questo end-point non è stato dimostrato publication bias	Per questo end-point non è stato dimostrato publication bias	Phung & coll hanno considerato come indipendenti eventi non mutuamente esclusivi (infarto fatale e infarto non fatale) e non hanno considerato nei calcoli l' interdipendenza dei confronti nel trial UKPDS Per questo end-point non è stato dimostrato publication bias	Per questo end-point non è stato dimostrato publication bias	Per questo end-point non è stato dimostrato publication bias

Web Tabella Z (b)

DIVERGENZE TRA I DATI RIPORTATI DA PHUNG & COLL PER I RCT E I RISULTATI DEL NOSTRO CRITICAL APPRAISAL

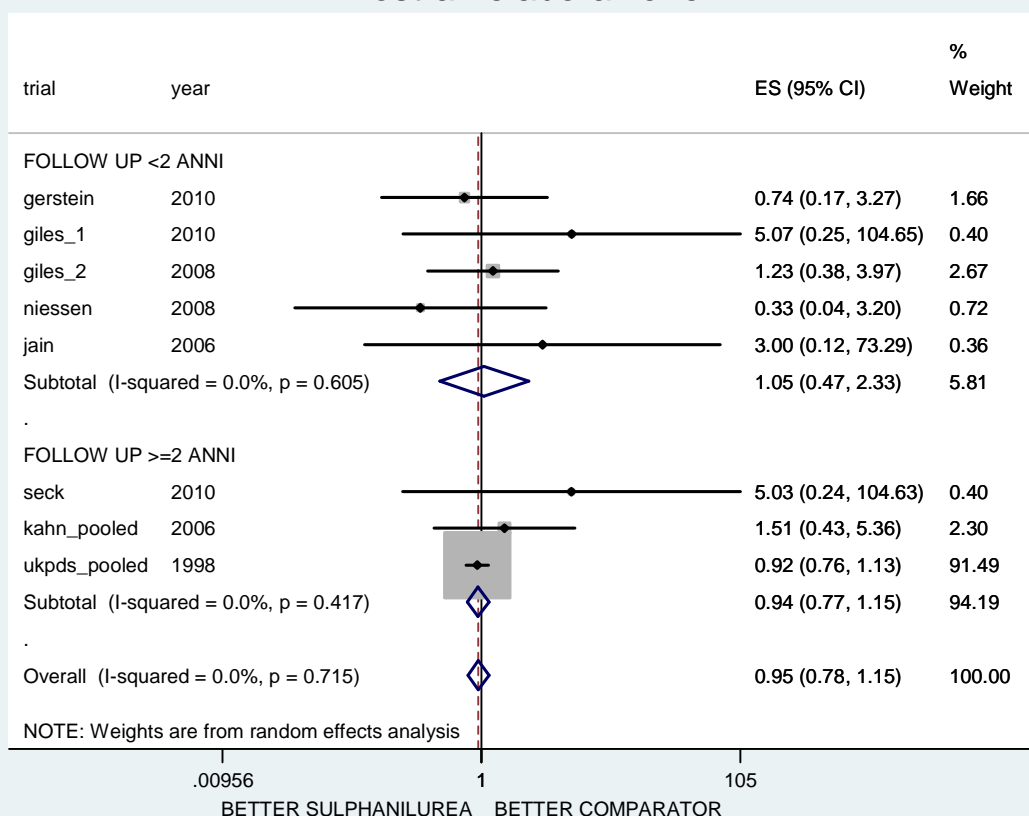
	Stroke (tutti i tipi)	Stroke fatale	Stroke non fatale	Cardiovascular composite endpoint	Ospedalizzazioni da cause CV
Conclusioni della metanalisi di Phung	Aumento non significativo del rischio di stroke RR=1.14 (0.86-1.51)	Endpoint non valutato	Endpoint non valutato	RR=0.98 (0.73-1.32)	Riduzione non significativa del rischio di ospedalizzazione RR=0.79 (0.48-1.29)
Conclusioni del nostro critical appraisal	Aumento non significativo del 7% del rischio di stroke RR=1.05 (0.73-1.52)	Solo l' UKPDS riporta eventi per questo endpoint RR=1.45 (0.87-2.44)	Aumento non significativo del 6% del rischio di stroke non fatale RR=1.06 (0.57-1.98)	Non l' abbiamo considerato	Riduzione significativa del 35% del rischio di ospedalizzazione RR=0.64 (0.42-0.98)
Note	Phung & coll hanno considerato come indipendenti eventi non mutuamente esclusivi (stroke fatale e stroke non fatale) Per questo endpoint è stato dimostrato publication bias (Egger p =0.046) per cui le informazioni non possono essere conclusive	I risultati riportati da un solo trial non consentono una analisi formale della presenza di publication bias per cui le informazioni non possono essere conclusive	Per questo endpoint non è stato dimostrato publication bias	L' analisi di questo endpoint non fornisce informazioni clinicamente utili in quanto le definizioni variano enormemente da trial e trial. Per enorme presenza di misclassificazione consideriamo fuorvianti le conclusioni di questa analisi.	Phung & coll non hanno considerato i dati dello studio Giles 2008 Per questo endpoint non è stato dimostrato publication bias

FOREST PLOT

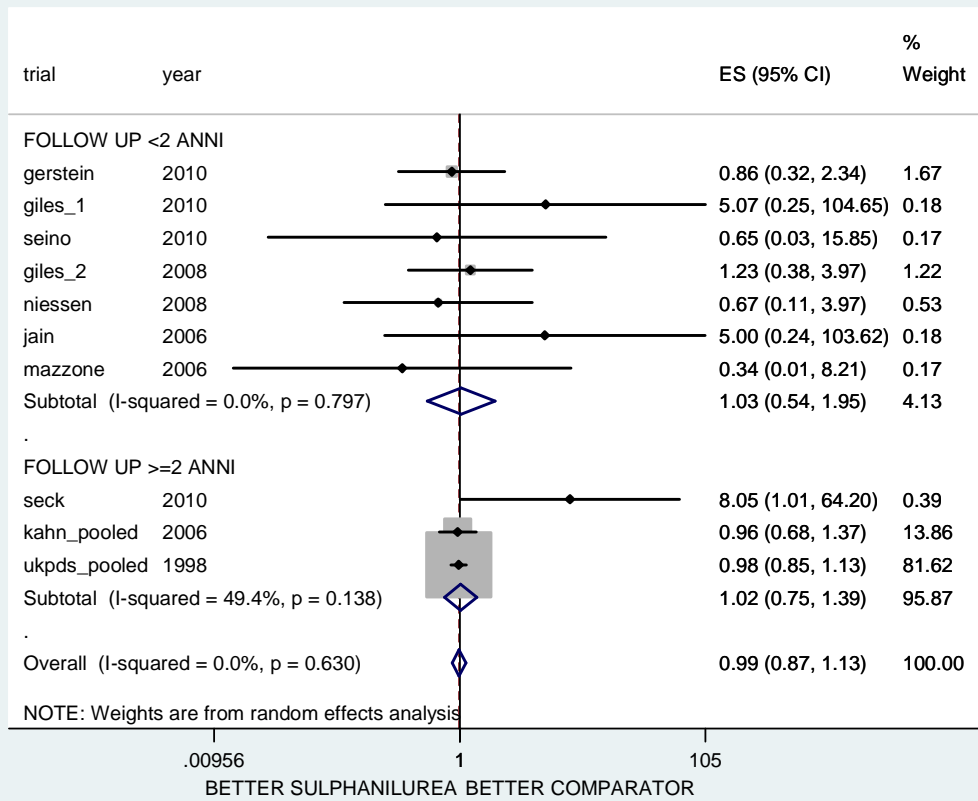
I Grafici 1-5 rappresentano le metanalisi da noi eseguite utilizzando i dati grezzi della letteratura sperimentale sottoposti alle nostre correzioni (vedi sopra web tabelle L-Z). Abbiamo anche eseguito una analisi per sottogruppi stratificati per lunghezza del follow-up; ulteriori dettagli sono riportati nella web tabella D

La tecnica di pooling da noi utilizzata è random effec based come quella usata dagli autori della metanalisi (tecnica di Wolff)

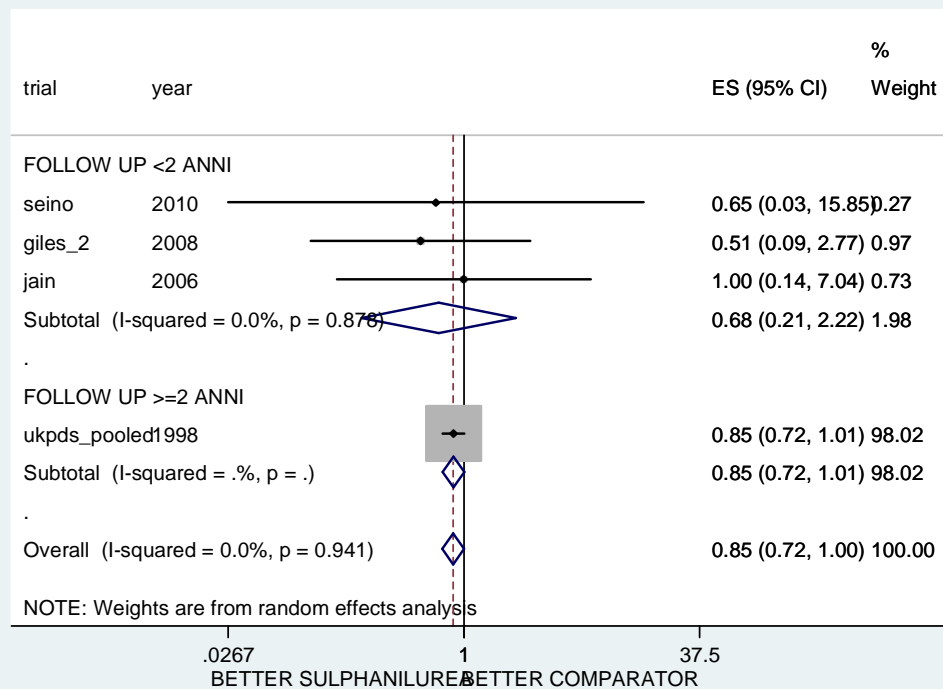
WEB GRAFICO 1: MORTE CARDIOVASCOLARE - nostra rielaborazione -



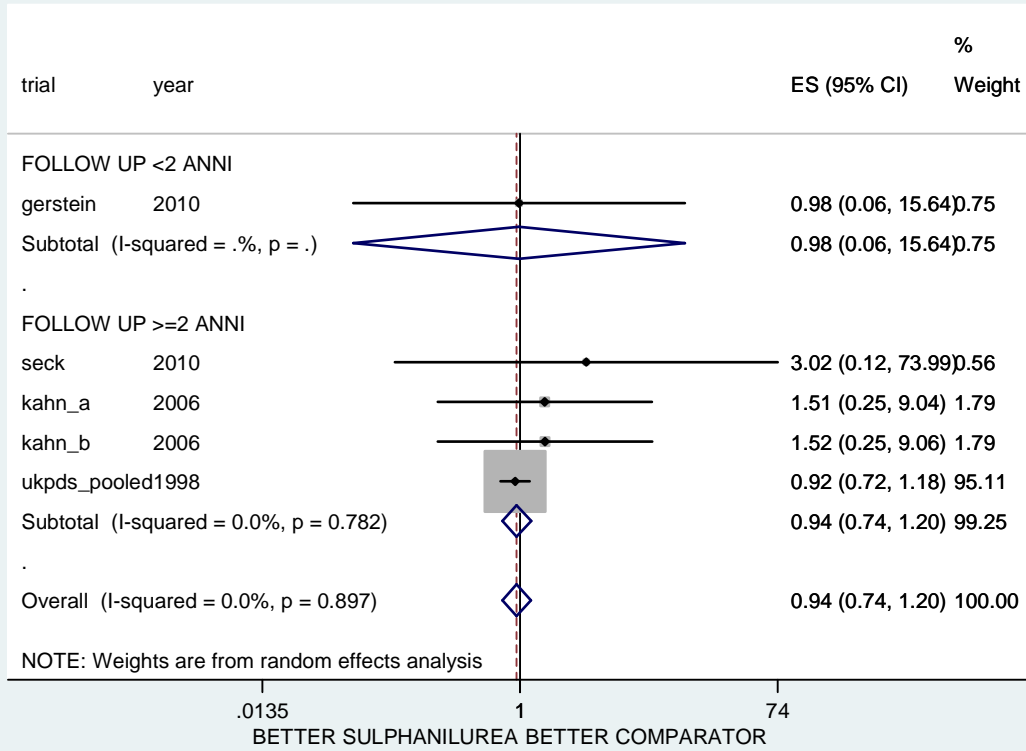
WEB GRAFICO 2: MORTE DA OGNI CAUSA - nostra rielaborazione -



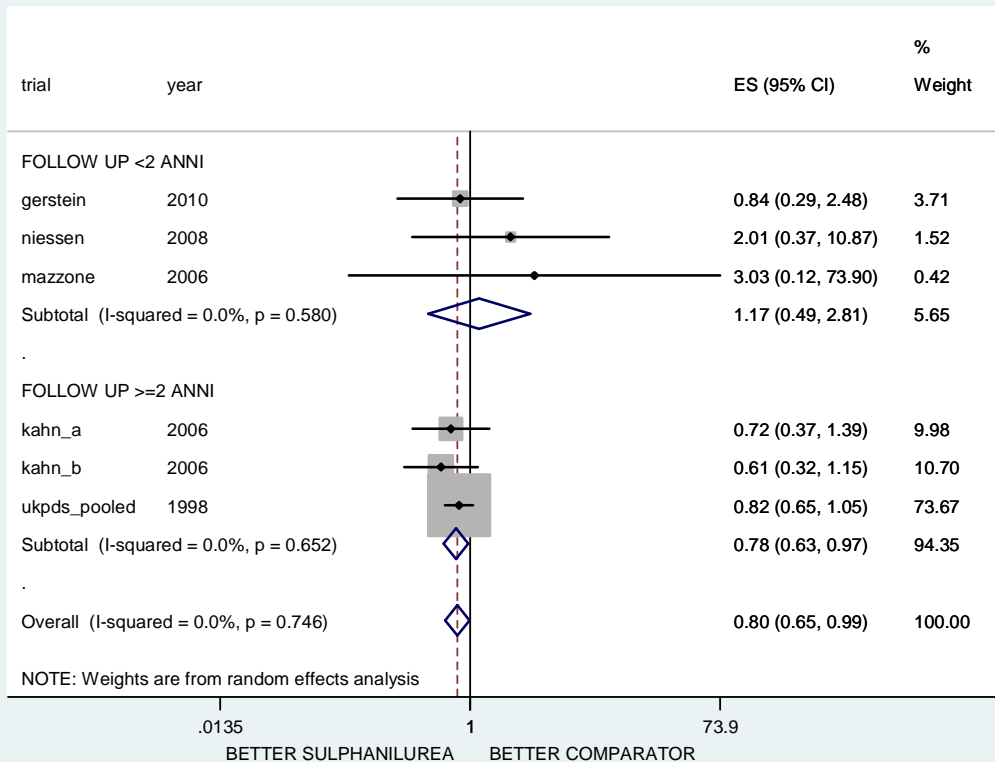
WEB GRAFICO 3: INFARTO MIOCARDICO - nostra rielaborazione -



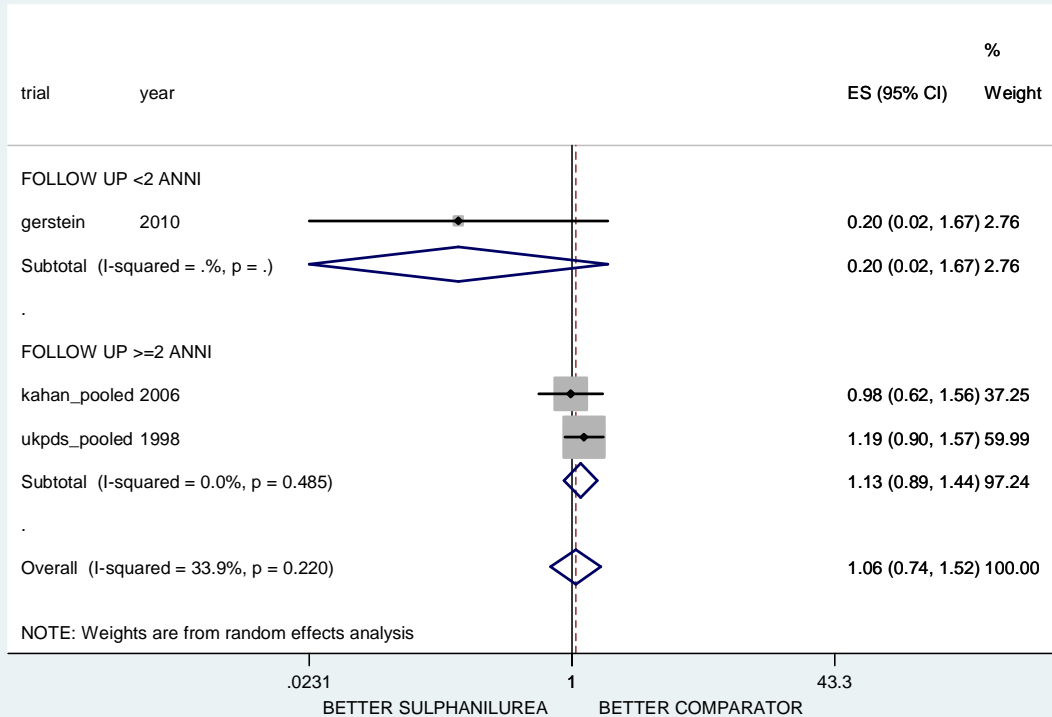
WEB GRAFICO 3bis: INFARTO MIOCARDICO FATALE
- nostra rielaborazione -



WEB GRAFICO 3tris: INFARTO MIOCARDICO NON FATALE
- nostra rielaborazione -

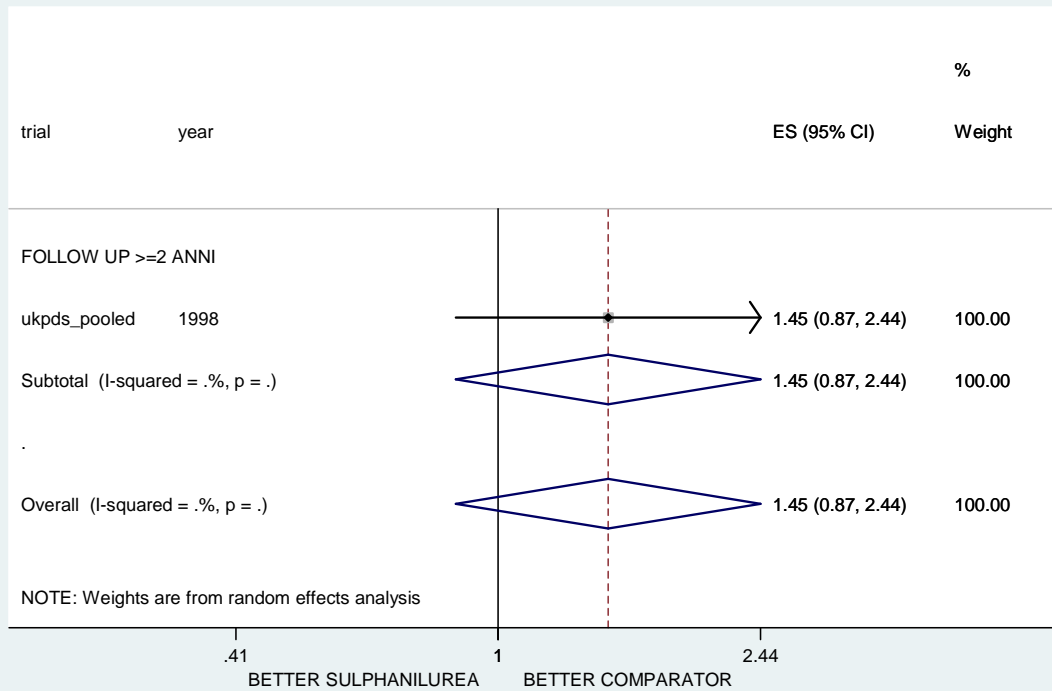


WEB GRAFICO 4: STROKE
- nostra rielaborazione -

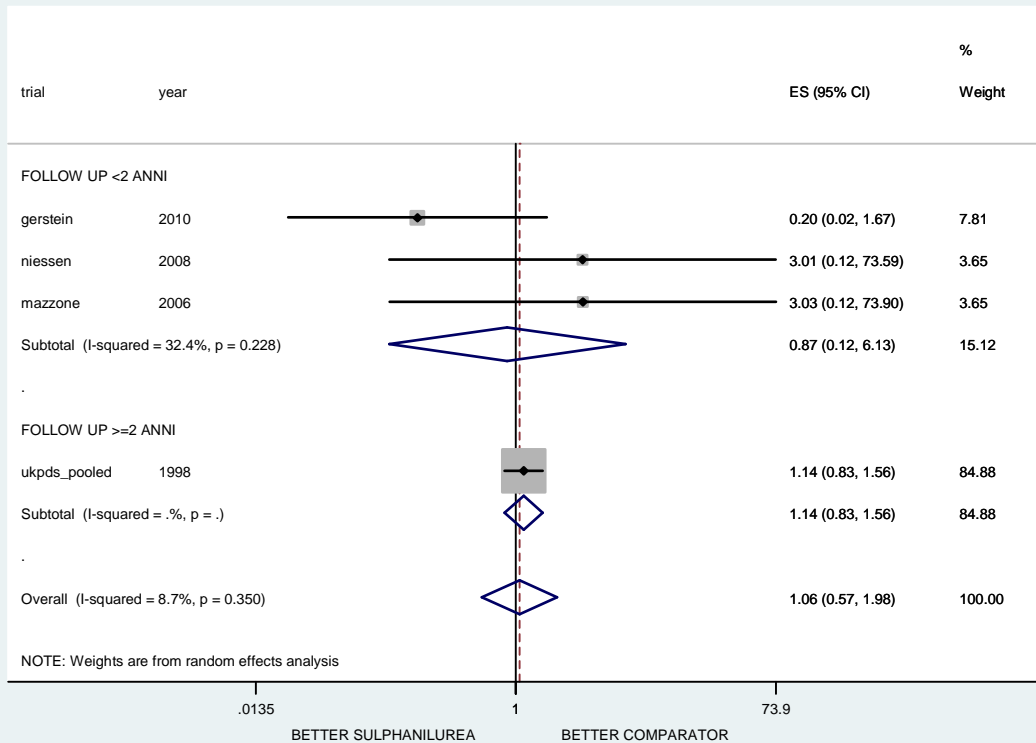


Lo stroke è l' unico end-point per cui la nostra rielaborazione dei risultati sperimentali ha rilevato una direzione dell' effetto a sfavore del trattamento . Tuttavia il test di Egger dimostra presenza di publication bias (**web grafico 9**) , per cui per questo end-point le informazioni non possono essere considerate esaustive e qualsiasi considerazione clinica risulterebbe fuorviante.

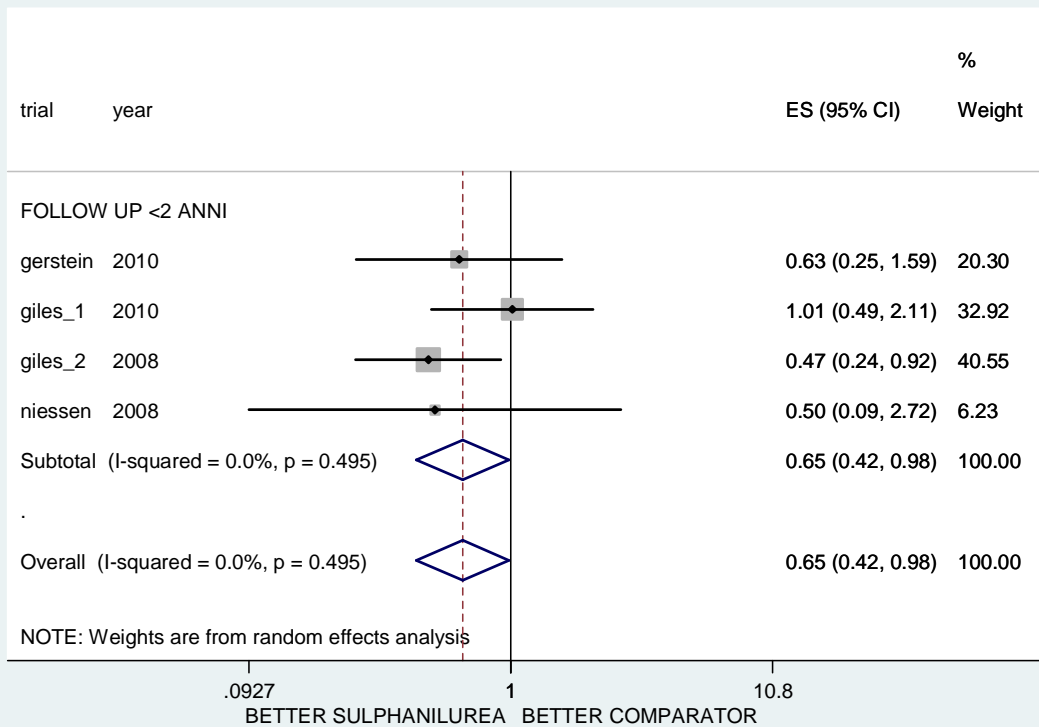
WEB GRAFICO 4bis: STROKE FATALE
nostra rielaborazione



WEB GRAFICO 4tris: STROKE NON FATALE
- nostra rielaborazione -



WEB GRAFICO 5: OSPEDALIZZAZIONI DA CAUSA CVD
- nostra rielaborazione -



FUNNEL PLOT

I Grafici 6-10 (Funnel Plot) rappresentano tecniche grafiche atte a valutare la presenza di publication bias per i singoli end-point considerati metanalisi. Si definisce publication bias la tendenza da parte di autori, editori e/o sponsor a pubblicare soprattutto risultati a favore dell' intervento e/o dotati di significatività statistica. La presenza di publication bias inficerebbe qualsiasi interpretazione clinica dei risultati della metanalisi, che sarebbe ovviamente mistificatoria. La presenza di publication bias è indagabile statisticamente sia con tecniche grafiche che con test formali. La tecnica grafica più utilizzata (Funnel Plot) presuppone che i risultati di efficacia delle ricerche prodotte su un determinato argomento presentino una distribuzione gaussiana.

Plottando allora i risultati di efficacia dei singoli trial versus la precisione delle rispettive stime (a propria volta inversamente proporzionale alla numerosità del campione e quindi inversamente proporzionale alla varianza della misura di efficacia) si otterrà in assenza di publication bias una distribuzione grossolanamente gaussiana.

In tale situazione ideale i trial di dimensioni più grandi forniranno stime più precise e saranno pertanto posizionati nella curva vicino al valore che esprime il risultato della metanalisi e nella parte apicale.

I trial di dimensioni più piccole, che forniranno stime meno precise e più soggette ad errore random, saranno invece posizionati a distanza più o meno variabile dalla linea verticale che esprime il risultato della metanalisi e occuperanno prevalentemente la base della gaussiana, perchè caratterizzati da minor precisione.

In presenza di publication bias la simmetria della curva tende a scomparire.

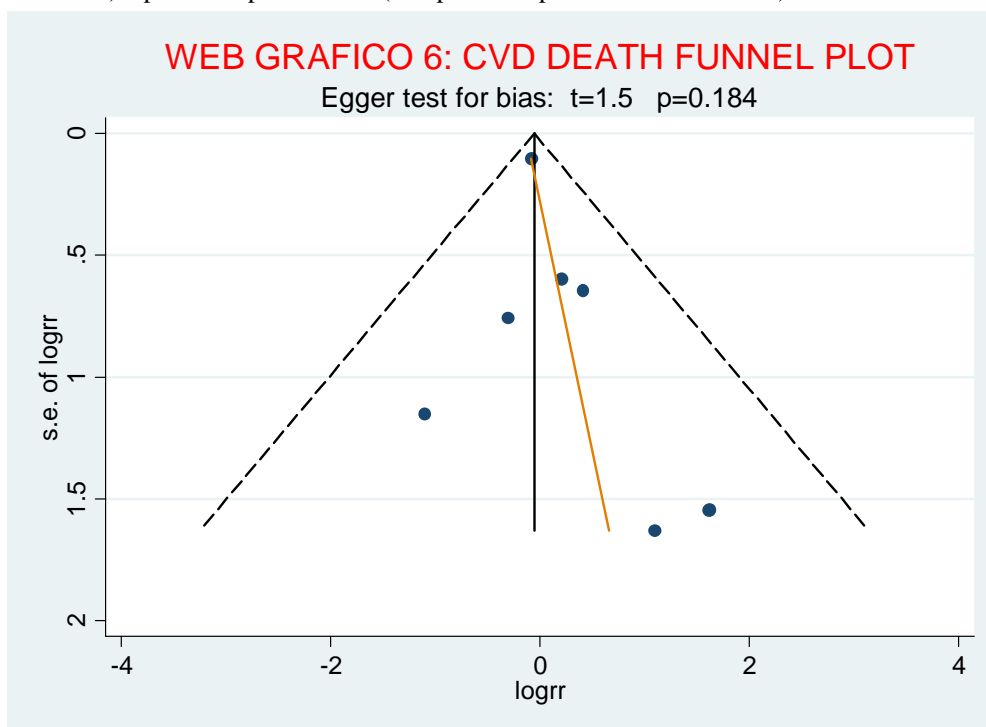
Per fare un esempio: se una metanalisi ha recuperato ricerche rivolte a dimostrare l' efficacia di un farmaco nei confronti della mortalità da ogni causa, valori di RR inferiori a 1 esprimeranno efficacia e valori di RR superiori a 1 esprimeranno dannosità. In presenza di Publication bias il funnel plot sarà asimmetrico per carenza di risultati caratterizzati da valore di RR superiore a 1 e abbondanza di risultati caratterizzati da valore di RR inferiore a 1.

Nei funnel plot sotto riportati la misura di efficacia è espressa dal logaritmo del Risk Ratio (logrr) e la precisione della stima è espressa dall' errore standard della misura di efficacia (standard error of log rr).

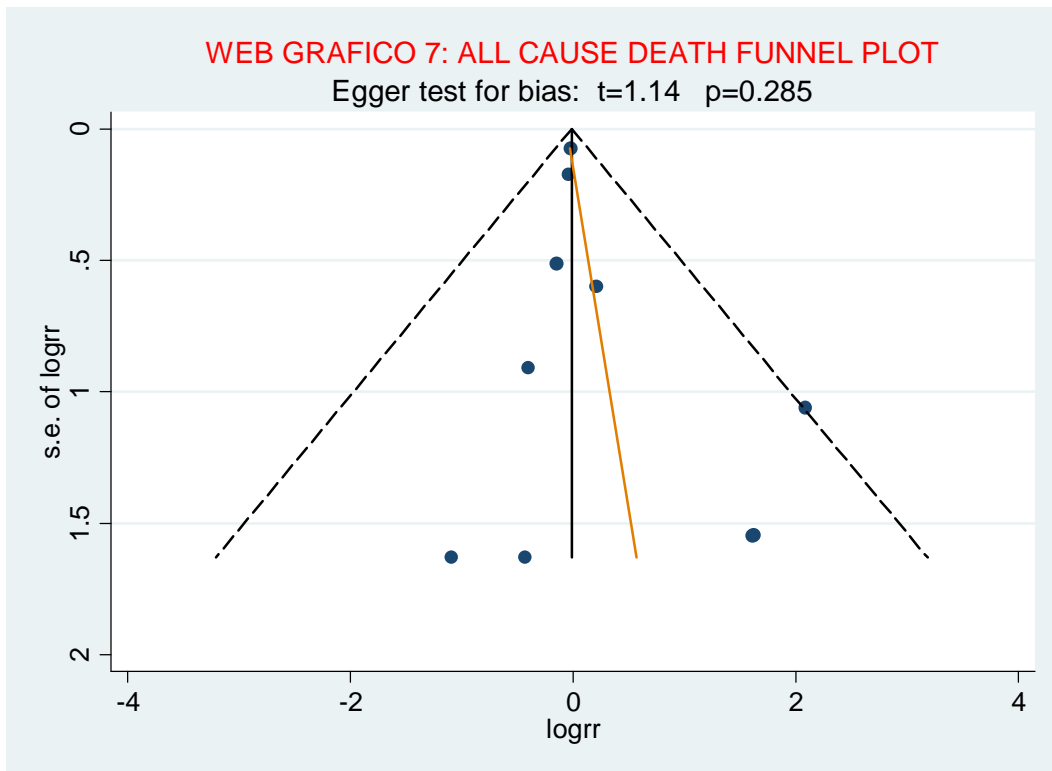
Il publication bias viene testato anche da test statistici formali; ciascun grafico è qui accompagnato dai risultati di uno dei test più usati, il test di Egger: valori di P inferiori a 0.10 sono suggestivi per esistenza di publication bias.

È in ogni caso da notare che queste tecniche presentano una alta probabilità di falsi negativi in quanto caratterizzate da bassa potenza statistica. La loro bassa potenza statistica è legata allo scarso numero di trial che in genere viene recensito da una metanalisi. È per correggere questo limite che il cut off di P del test di Egger viene interpretato 'suggestivo per publication bias' a valori <0.10 (e non <0.05). In ogni caso, sempre seguendo l' ottica di questo ragionamento, le conclusioni che si possono trarre dalle tecniche atte a indagare sul publication bias sono più attendibili quando queste analisi forniscono risultati positivi (ossia a sostegno dell' esistenza di publication bias)

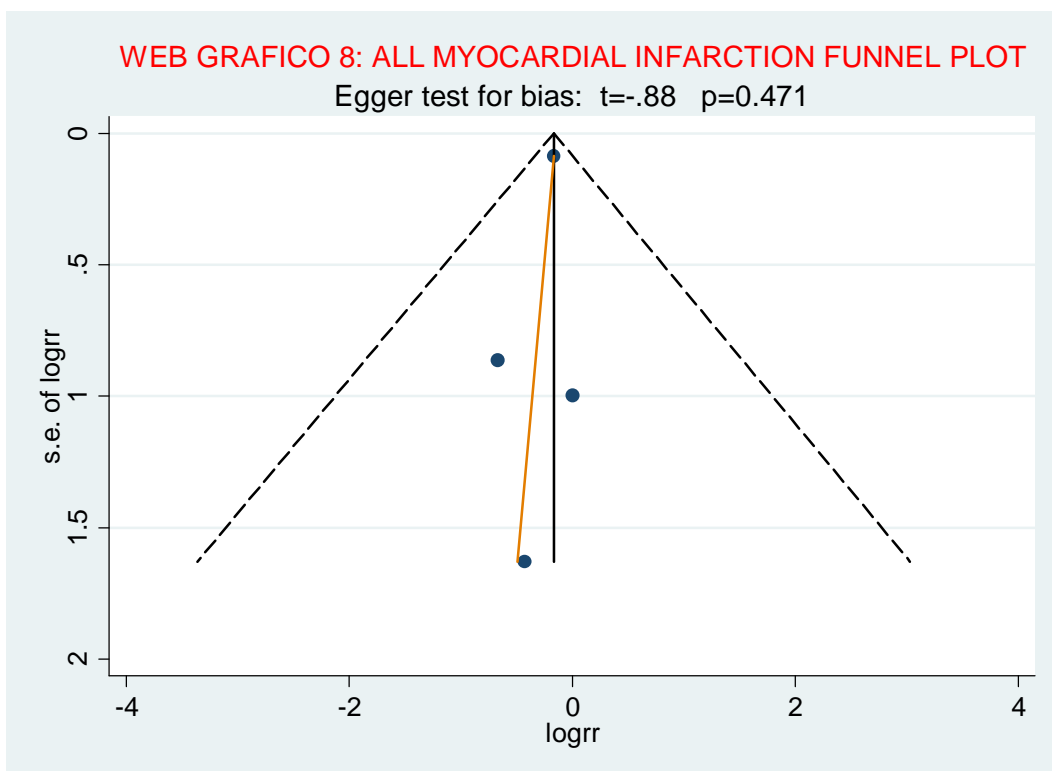
La presenza di publication bias è stata evidenziata dalle nostre analisi per l' end-point infarto fatale (non per l' endpoint infarto da tutte le cause) e per l' end-point stroke (non per l' endpoint stroke non fatale)



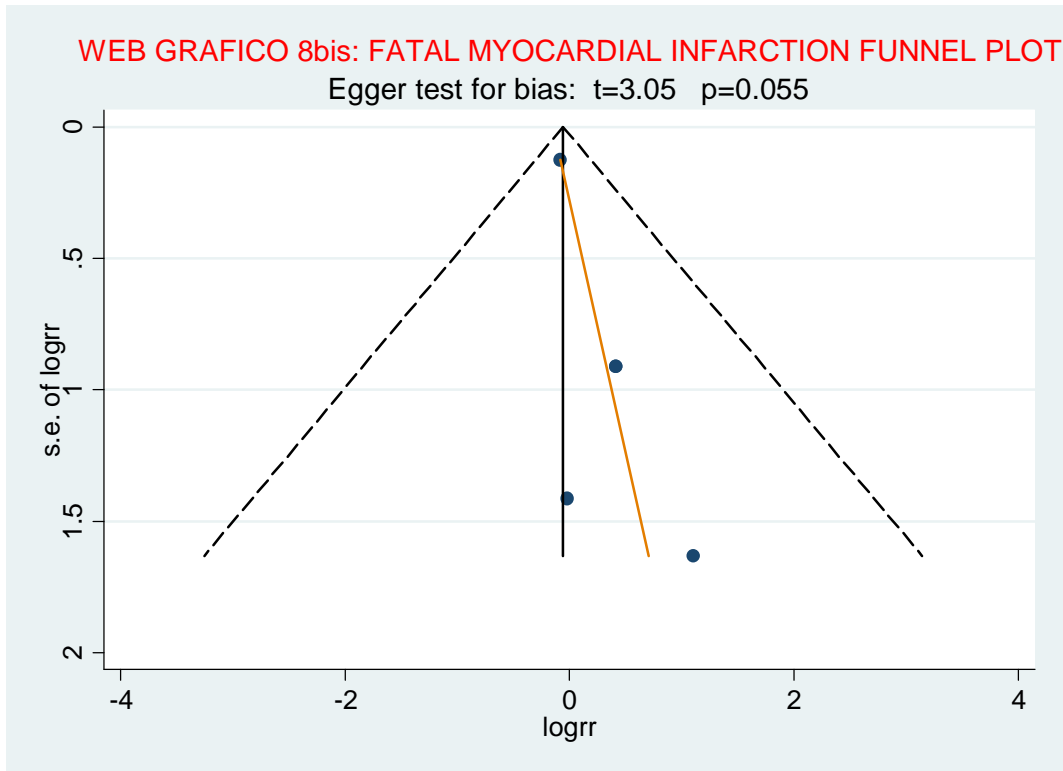
Il Funnel plot e l' Egger test non confermano esistenza di publication bias



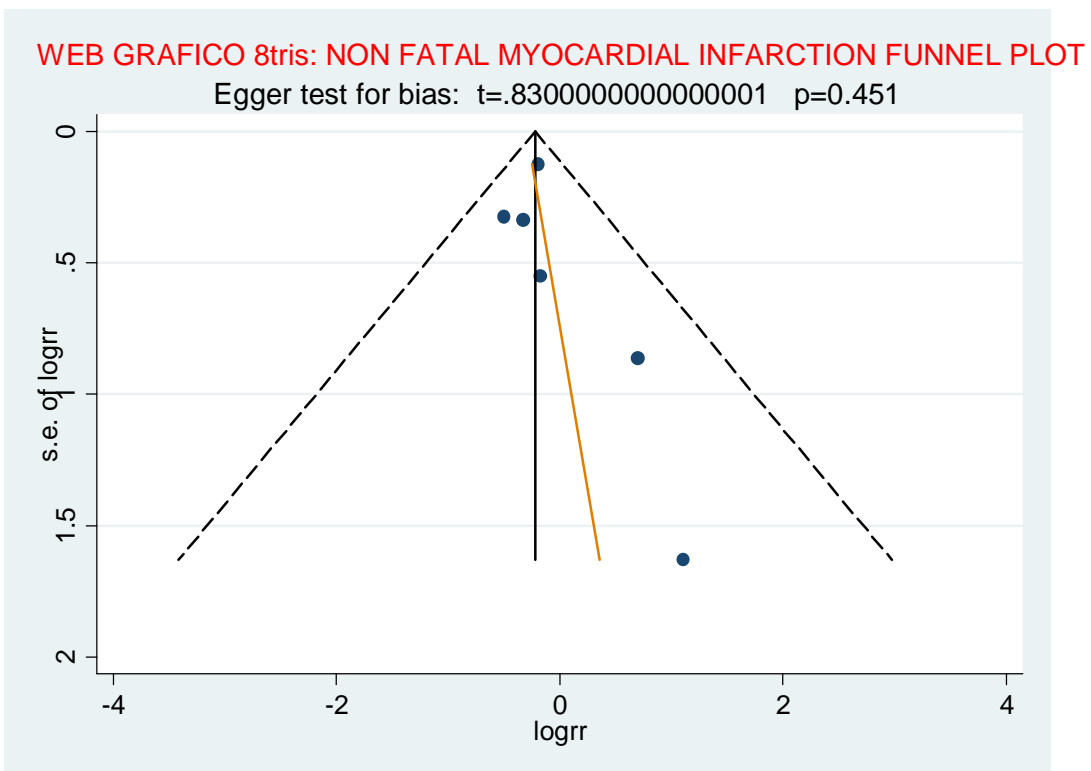
Il Funnel plot e l' Egger test non confermano esistenza di publication bias



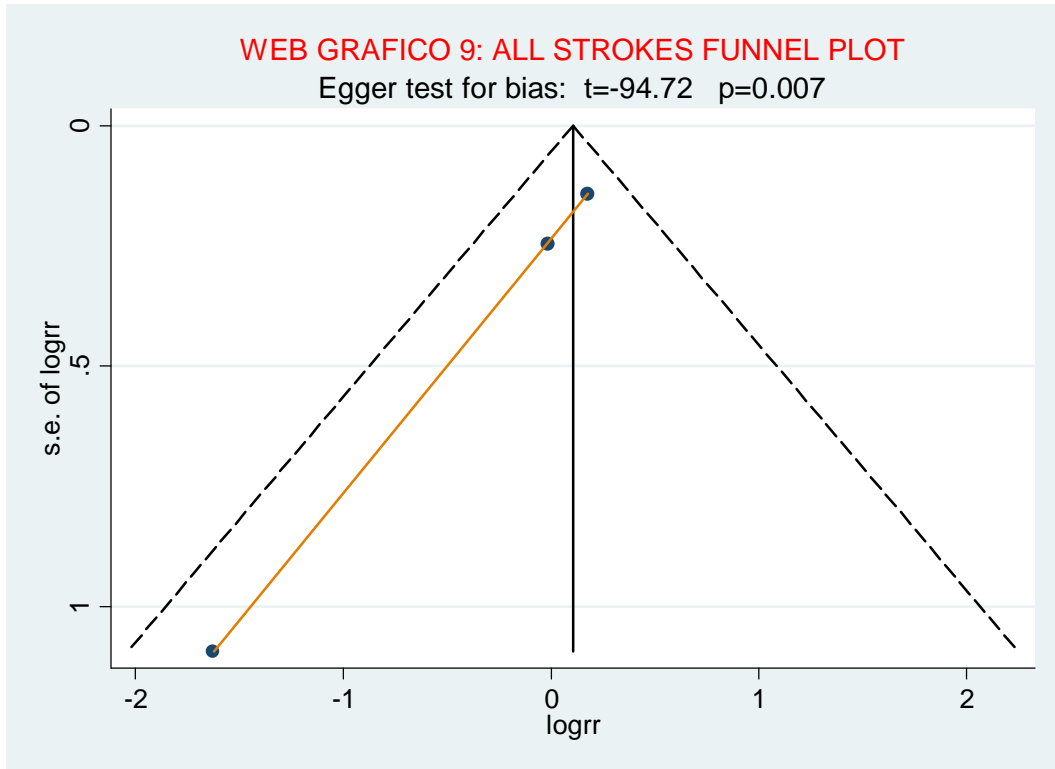
Il Funnel plot e l' Egger test non confermano esistenza di publication bias



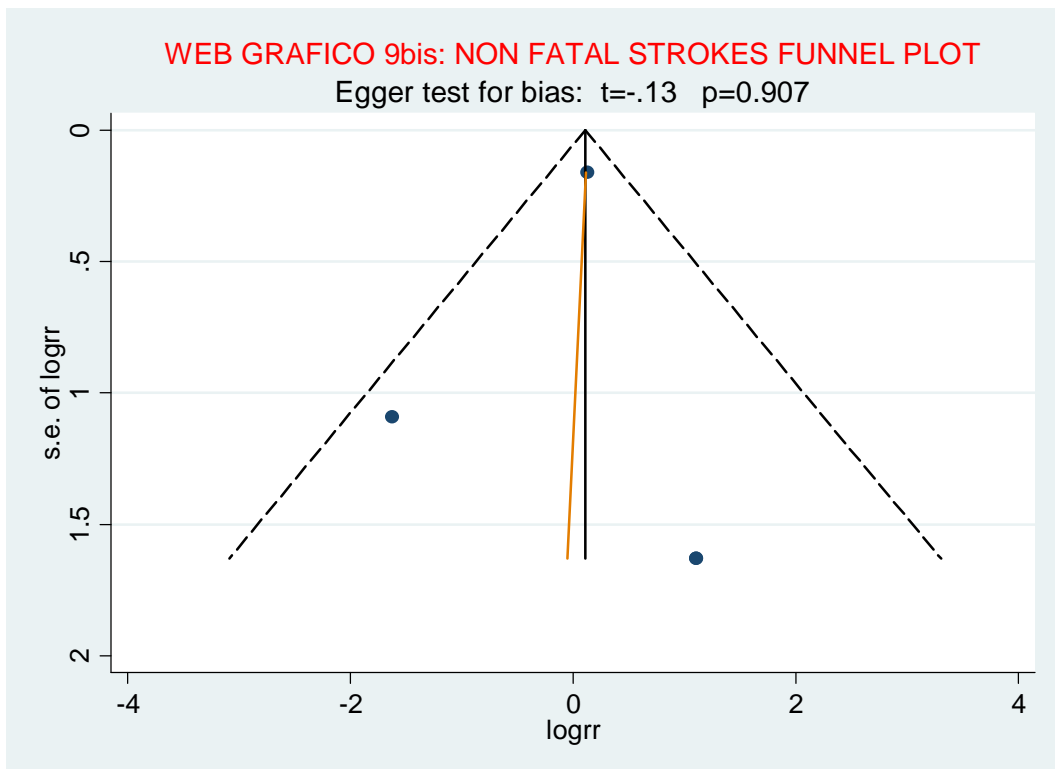
Il Funnel plot e l' Egger test suggeriscono esistenza di publication bias



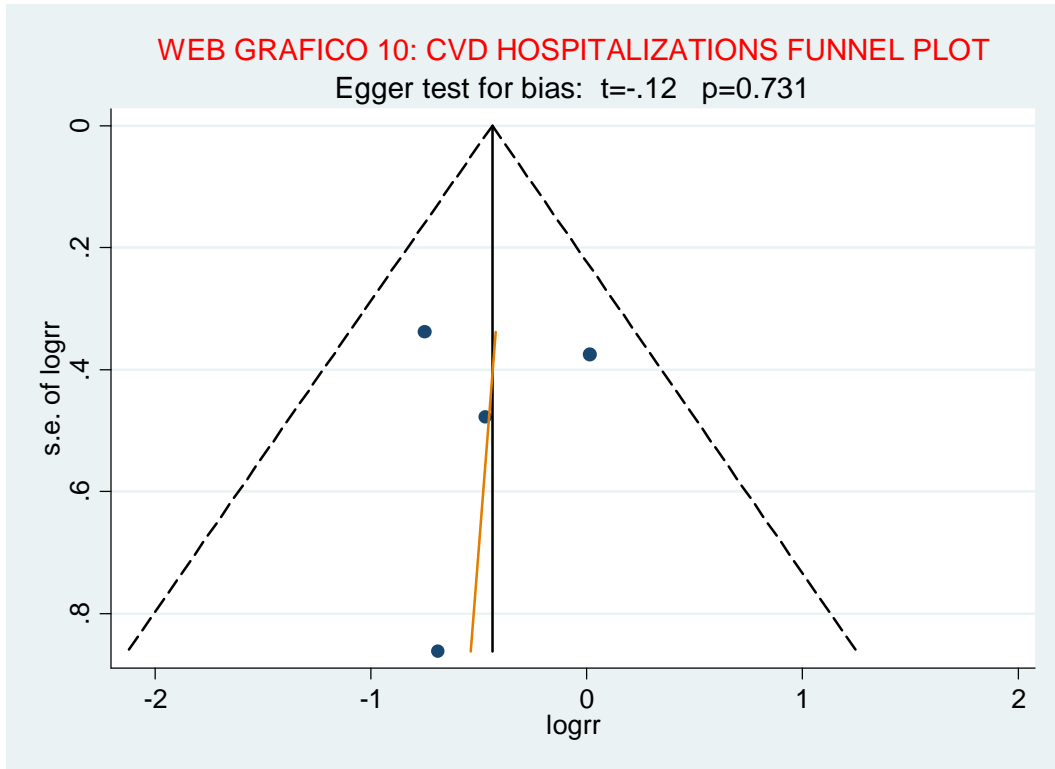
Il Funnel plot e l' Egger test non confermano esistenza di publication bias



Il Funnel plot e l' Egger test suggeriscono esistenza di publication bias



Il Funnel plot e l' Egger test non confermano esistenza di publication bias



Il Funnel plot e l' Egger test non confermano esistenza di publication bias