

L'analisi dell'efficacia del vaccino anti-SARS-CoV-2 con studi sulla popolazione generale. Gli errori metodologici che provocano un aumento apparente dell'efficacia

Prof. Stefano Petti

Professore Ordinario, Dipartimento di Sanità Pubblica e Malattie Infettive, Università "La Sapienza", Roma

Come si valuta correttamente l'efficacia di un vaccino – i clinical trials

La corretta valutazione dell'efficacia di un vaccino si effettua esclusivamente attraverso i clinical trials, in cui due gruppi di persone simili tra loro per numero e caratteristiche sono selezionati attraverso un processo casuale detto randomizzazione. Ad un gruppo (test) viene somministrato il vaccino, ad un altro gruppo (controllo) viene somministrato il placebo (il placebo deve essere innocuo). I soggetti selezionati e i medici non sono a conoscenza di chi abbia ricevuto il vaccino e chi il placebo. Durante un periodo (follow-up) di osservazione si valuta la comparsa della malattia che si vuole prevenire. Il parametro si chiama incidenza (i):

$$i = (\text{Numero di casi di malattia successivi a vaccinazione o placebo}) / (\text{Numero di soggetti esposti}) \times 100$$

Ad esempio, se nel gruppo test di 10.000 soggetti si verificano 50 casi di infezione l'incidenza sarà:

$$i = (50/10.000) \times 100 = 0,5\%$$

L'efficacia del vaccino si valuta attraverso diverse formule, le più usate sono l'Efficacia Vaccinale (Vaccine Efficacy, EV). L'EV indica quanto sarebbe diminuita l'incidenza della malattia nel gruppo dei non vaccinati se fossero stati vaccinati. La formula è:

$$EV = [(incidenza nel gruppo controllo - incidenza nel gruppo test) / (incidenza nel gruppo placebo)] \times 100$$

Ad esempio, se l'incidenza nel gruppo test è stata 0,5% e nel gruppo controllo è stata 5%, l'EV sarà:

$$EV = [(5,0 - 0,5) / (5,0)] \times 100 = (4,5/5,0) \times 100 = 90\%$$

In pratica, l'incidenza nei non vaccinati sarebbe diminuita del 90% se si fossero vaccinati.

Per esempio, nel clinical trial del vaccino Moderna i risultati sono stati i seguenti:

- incidenza covid nel gruppo controllo = 5,3%

- incidenza covid nel gruppo test = 0,35%
- VE = 93%

Ci sono alcune REGOLE da rispettare, altrimenti i valori e i calcoli sono falsati (biased). Riporto nuovamente la formula dell'incidenza:

$$i = (\text{Numero di casi di malattia successivi a vaccinazione o placebo}) / (\text{Numero di soggetti esposti}) \times 100$$

- I "soggetti esposti" a cui fa riferimento la formula sono le persone che fanno parte del gruppo test oppure del gruppo placebo. Queste persone NON DEVONO MAI AVERE AVUTO LA MALATTIA IN PRECEDENZA. Questo esclude le persone che l'hanno già avuta o hanno avuto l'infezione asintomatica. Il motivo è chiaro: lo stato immunitario nei confronti del covid è diverso rispetto a chi non l'ha mai avuta (sono più protetti).
- Per lo stesso motivo, se una persona ha avuto il covid durante il follow-up e sviluppa di nuovo la malattia sempre durante il follow-up, il secondo episodio di malattia non si deve contare perché L'INCIDENZA RIGUARDA IL NUMERO DI PERSONE CHE SVILUPPA IL COVID E NON IL NUMERO TOTALE DI CASI DI COVID. Quindi NON È POSSIBILE CONTARE DUE VOLTE AL NUMERATORE DUE EPISODI DI MALATTIA NELLA STESSA PERSONA, ciò provocherebbe ad avere un valore FALSAMENTE più alto dell'incidenza in quel gruppo.
- Per calcolare l'incidenza nel gruppo di controllo (lo stesso discorso vale per il gruppo test) al numeratore va il numero di persone non vaccinate che sviluppano il covid durante il follow-up, al denominatore tutte le persone non vaccinate incluse nel clinical trial. Ne consegue che le PERSONE CHE SVILUPPANO IL COVID UTILIZZATE PER CALCOLARE IL NUMERATORE devono FAR PARTE DEL GRUPPO DI PERSONE INCLUSE NEL DENOMINATORE.

Nei clinical trials l'osservatore segue tutte le persone incluse nel gruppo controllo e quando si verifica un caso di malattia tra queste persone lo attribuisce correttamente al gruppo controllo. Non è dunque possibile che un malato, ad es del gruppo test venga inserito come caso di malattia nel gruppo controllo, perché ciò porterebbe ad avere un valore FALSAMENTE più alto dell'incidenza della malattia nel gruppo controllo.

Riassumendo, nella valutazione dell'incidenza sia nel gruppo test che nel gruppo controllo le persone non devono avere avuto la malattia in precedenza, quando si ammalano devono essere considerate nel gruppo originario di appartenenza, se si ammalano una seconda volta questo secondo episodio non si conta.

Studi osservazionali per la valutazione dell'efficacia di un vaccino

Gli studi osservazionali sono quelli che “osservano” ciò che succede nella vita di tutti i giorni, manca quindi l'intervento degli sperimentatori e manca il controllo su molti fattori che potrebbero differenziare le persone vaccinate da quelle non vaccinate introducendo quindi diverse forme di errore che porta ad avere risultati sbagliati (biased).

A livello di popolazione non ci sono più due gruppi, uno sottoposto a vaccinazione e uno di controllo a cui viene dato il placebo, ma c'è una singola popolazione a cui è stato chiesto di vaccinarsi. Per cui le persone che si vaccinano da quel momento in poi sono classificate come vaccinate (in realtà si aspettano, e senza alcun valido motivo scientifico, 21 giorni, e la sovrastima dell'efficacia vaccinale data da questa attesa è uno degli errori di cui parla Fenton), ma fino al momento precedente la vaccinazione erano, ovviamente, non vaccinate. Sembra un particolare senza apparente significato ma gli effetti sulla distorsione della stima dell'efficacia del vaccino sono notevoli.

Con una sola popolazione a disposizione quello che succede è quindi che non c'è più un gruppo test e un gruppo controllo nei quali calcolare e confrontare l'incidenza, ma ci sono persone che nel momento della vaccinazione passano automaticamente dal gruppo controllo non vaccinato al gruppo test vaccinato, mentre non è possibile il contrario cioè che persone vaccinate passino al gruppo dei non vaccinati per ovvie ragioni.

Perché l'uso dell'incidenza è sbagliato nello studio sulla popolazione

Prendiamo il caso del covid e consideriamo, solo per semplicità, esclusivamente la prima dose di vaccino. Abbiamo questi periodi:

- marzo-dicembre 2020 (10 mesi): a causa della pandemia tutta la popolazione è esposta al covid (ossia si può ammalare) e il vaccino non esiste, sono quindi tutti “controlli”;
- dicembre 2020: inizia la vaccinazione di massa con la prima dose; secondo l'osservatorio epidemiologico “Our World in Data”, a tutt'oggi il numero di Italiani che ha fatto la prima dose è pari all'81%, tale valore è stato raggiunto approssimativamente alla fine di febbraio 2022;

- gennaio 2021-febbraio 2022 (14 mesi): somministrazione della prima dose di vaccino (anche di dosi successive in realtà). Le persone che si sono vaccinate in questo periodo sono passate dal gruppo controllo al gruppo test, mentre il restante 19% che non si è vaccinato è rimasto nel gruppo controllo.

Distinguiamo tre casi possibili:

- Caso A, una persona che si vaccina appena è disponibile il vaccino: è stato “controllo” per 10 mesi e “test” per 14 mesi (controllo=10 mesi; test=14 mesi);
- Caso B, una persona che si vaccina durante il periodo di vaccinazioni, per esempio proprio a metà, ossia a fine luglio 2021: è stato “controllo” per 10 mesi, poi ancora “controllo” durante i primi 7 mesi nei quali si somministravano le vaccinazioni, infine “test” per i successivi 7 mesi (controllo=17 mesi; test=7 mesi),
- Caso C, una persona che non si vaccina: rimane “controllo” per tutti e 24 i mesi.

Come si fa a calcolare l’incidenza nei vaccinati (test) e nei non vaccinati (controllo)? Nel caso del calcolo dell’incidenza nei clinical trials è facile perché le persone restano rigorosamente all’interno del proprio gruppo di appartenenza (vaccinati, non vaccinati), ma qui c’è un flusso continuo di persone che passa dal gruppo controllo al gruppo test e l’incidenza non si può più calcolare seguendo il metodo dei clinical trials. Infatti, il Caso C non si sposta dal gruppo controllo per cui se si ammala sarà un caso di malattia nel gruppo controllo. Ma sia il Caso A che il Caso B sono stati nel gruppo controllo per un certo periodo e nel gruppo test da un certo momento in poi.

Si potrebbe erroneamente dire che se A e B si fossero ammalati prima della vaccinazione sarebbero stati consideranti come casi di covid nel gruppo dei non vaccinati, se invece si fossero ammalati dopo la vaccinazione sarebbero stati consideranti come nuovi casi di covid nel gruppo dei vaccinati. Non è così. Infatti, immaginiamo che B si ammali tre mesi dopo la vaccinazione. Essendo vaccinato, questo caso di covid va a finire nel calcolo dell’incidenza del gruppo dei vaccinati, però questa persona è rimasta non vaccinata per 17 mesi e non si può non tenere conto del fatto che in questo periodo questa persona era un controllo non vaccinato e per 17 mesi non si è ammalato. Questi 17 mesi di appartenenza al gruppo dei non vaccinati senza avere avuto il covid non vengono conteggiati nel calcolo erroneo dell’incidenza. Al contrario, se B si fosse ammalato prima di vaccinarsi sarebbe, senza rischio di errori, risultato come un caso di covid nel gruppo dei non vaccinati.

Quello che succede è quindi che il gruppo controllo si “svuota” progressivamente a favore del gruppo test che invece aumenta progressivamente e chi passa dal gruppo controllo al gruppo test è stato per un certo tempo un controllo non vaccinato. Bisogna quindi tenere conto di entrambi i periodi di esposizione al virus, precedente e successivo alla vaccinazione.

Per questo motivo, **negli studi di popolazione non si può utilizzare il parametro dell’incidenza ma si utilizza un parametro detto “densità di incidenza”** che si misura in “tempo-persona”. La formula è più complessa ma cambia solamente il denominatore. La formula per il gruppo controllo è:

$$\text{Densità di incidenza} = \left[\frac{\text{Numero di casi di malattia senza (pre/no) vaccinazione}}{\text{somma dei mesi senza (pre/no) vaccinazione da parte della popolazione}} \right] \times 100$$

La formula per il gruppo test è:

$$\text{Densità di incidenza} = \left[\frac{\text{Numero di casi di malattia dopo la vaccinazione}}{\text{somma dei mesi di appartenenza al gruppo dei vaccinati da parte della popolazione}} \right] \times 100$$

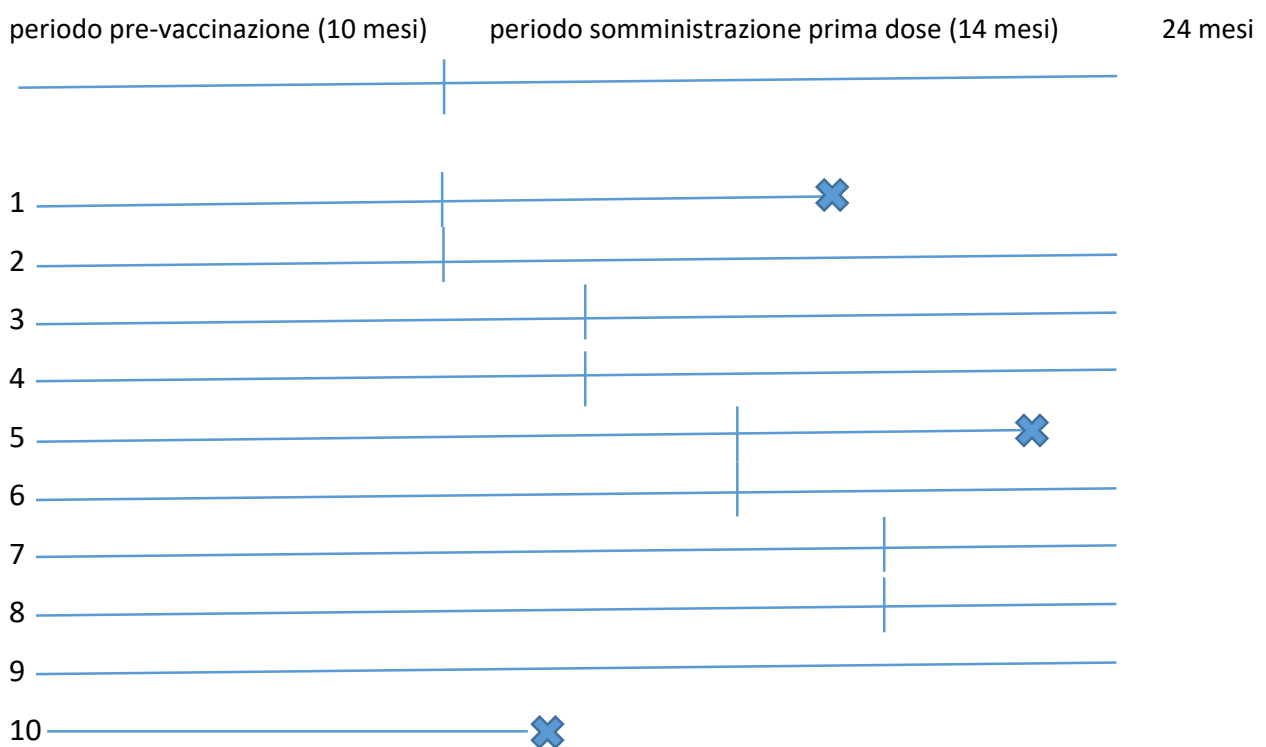
Ecco un esempio con una popolazione immaginaria di 10 persone (**Figura 1**):

Due persone (1 e 2) si vaccinano non appena il vaccino si è reso disponibile, due persone (3 e 4) si vaccinano 3 mesi dopo l’inizio delle vaccinazioni, due persone (5 e 6) si vaccinano 6 mesi dopo l’inizio delle vaccinazioni, due persone (7 e 8) si vaccinano 9 mesi dopo l’inizio delle vaccinazioni, due persone (9 e 10) pari al 20% del totale, non si vaccinano.

Si ammalano tre persone, la numero 1, 8 mesi dopo la vaccinazione, la numero 5 sei mesi dopo la vaccinazione, la numero 9, 12 mesi dopo lo scoppio della pandemia.

Figura 1

Popolazione immaginaria di 10 persone. La barra verticale indica la somministrazione della prima dose, la croce indica lo sviluppo del covid, la lunghezza della linea corrisponde al periodo di esposizione al covid. Quando una persona sviluppa la malattia non è più oggetto del calcolo dell'incidenza. È possibile sviluppare il covid più volte, ma se si vuole calcolare l'incidenza si deve stabilire esattamente l'incidenza di che cosa, in questo caso è il primo episodio di covid, ma può anche essere il secondo episodio di covid, oppure il terzo, oppure può essere l'ospedalizzazione per covid, o il decesso per covid, ecc.



Calcolo erroneo (biased) dell'incidenza dopo 24 mesi dall'inizio della pandemia:

Vaccinati, sono 8 con 2 casi di covid. **Incidenza nei vaccinati**=(2/8)x100=25%

Non vaccinati, sono 2 con 1 caso di covid. **Incidenza nei non vaccinati**=(1/2)x100=50%

Il vaccino è apparentemente protettivo.

Come abbiamo detto, in questo modo non si tiene conto del tempo in cui le persone vaccinate non lo erano ancora, anche in quel periodo erano a rischio di covid.

Calcolo corretto, della densità di incidenza dopo 24 mesi dall'inizio della pandemia:

Vaccinati, il NUMERATORE è lo stesso dell'incidenza, sono 2 casi di covid. La differenza sta nel denominatore che non considera il numero di persone ma corrisponde alla somma dei mesi passati da vaccinati in coloro che hanno ricevuto la prima dose.

Calcolo del DENOMINATORE:

Persona 1: 8 mesi e poi ha avuto il covid

Persona 2: 14 mesi

Persona 3: 11 mesi

Persona 4: 11 mesi

Persona 5: 6 mesi e poi ha avuto il covid

Persona 6: 8 mesi

Persona 7: 5 mesi

Persona 8: 5 mesi

La somma dei mesi senza covid dopo la vaccinazione è pari a:

$$8+14+11+11+6+8+5+5=68$$

Densità di incidenza nei vaccinati=(2/68)x100=2,9 per 100 persone per ogni mese

Non vaccinati, il NUMERATORE è lo stesso dell'incidenza, 1 caso di covid.

Calcolo del DENOMINATORE:

Persona 1: 10 mesi prima di vaccinarsi

Persona 2: 10 mesi prima di vaccinarsi

Persona 3: 13 mesi prima di vaccinarsi

Persona 4: 13 mesi prima di vaccinarsi

Persona 5: 16 mesi prima di vaccinarsi

Persona 6: 16 mesi prima di vaccinarsi

Persona 7: 19 mesi prima di vaccinarsi

Persona 8: 19 mesi prima di vaccinarsi

Persona 9: 24 mesi

Persona 10: 12 mesi prima di sviluppare il covid

La somma dei mesi prima del covid nei non vaccinati e nei vaccinati prima di vaccinarsi è pari a:

$$10+10+13+13+16+16+19+19+24+12=152$$

Densità di incidenza nei non vaccinati= $(1/152) \times 100 = 0,7$ per 100 persone per ogni mese

Il vaccino non è protettivo

Bisogna ribadire che questo è un **esempio fittizio, inteso a dimostrare che l'incidenza non è un indicatore valido per studiare l'efficacia della vaccinazione negli studi di popolazione**, e che in questi casi si deve utilizzare la densità di incidenza. I numeri, come detto, sono di pura fantasia.

Calcolo della densità di incidenza nella popolazione italiana

Per calcolare la densità di incidenza serve un sistema di sorveglianza che segua ogni singola persona dall'inizio della pandemia e riporti per quella persona:

(1) la comparsa del covid (o del secondo o terzo caso di covid, o l'ospedalizzazione, o il decesso, secondo qual è l'obiettivo dell'indagine) in rapporto allo stato vaccinale; con questo parametro si calcola il NUMERATORE della densità di incidenza nei vaccinati e nei non vaccinati;

(2) la data di vaccinazione (o la seconda dose, o il primo, secondo, terzo booster, ecc.) e quindi i tempi di ognuno passati senza essere vaccinati, oppure vaccinati con una dose, o due, ecc. con questo parametro si calcola il DENOMINATORE della densità di incidenza nei vaccinati e nei non vaccinati.

Il problema dei sistemi di sorveglianza, incluso quello italiano è che **non c'è una banca dati sufficientemente accurata per ogni singola persona** che consenta di dire per quanto tempo è stato senza vaccinazione e per quanto tempo è stato con vaccinazione (il denominatore del parametro da utilizzare). Ci sono due sistemi, uno che riporta quanti hanno ricevuto la vaccinazione (incluse le varie dosi), mentre un secondo sistema avvisa quando una persona sviluppa il covid e se quella persona è o non è vaccinata al momento dell'insorgenza della malattia. Con **due sistemi di sorveglianza disconnessi tra loro, non si può calcolare il denominatore della densità di incidenza**,

perché manca l'informazione relativa al tempo passato da non vaccinati di coloro che poi si sono vaccinati. Per contro, si può conoscere il periodo passato da vaccinati da parte delle persone che si sono vaccinate.

Questo è l'errore (bias) notato da Fenton, ossia che non c'è connessione tra il sistema che comunica quelli che si vaccinano e il sistema che comunica quelli che si infettano. L'unico elemento noto è se chi si infetta è o non è vaccinato in quel momento.

Con questo sistema si può soltanto fare un calcolo errone (biased) dell'incidenza, infatti per un determinato intervallo di tempo è possibile sapere quante persone sono e non sono vaccinate, quante hanno sviluppato il covid e quali tra queste ultime era o non era vaccinato.

L'esempio precedente dimostra la presenza del bias nel calcolo dell'incidenza e soprattutto dimostra che questo bias "favorisce" la vaccinazione, perché manca tutto il tempo passato da non vaccinato da parte del soggetto vaccinato. In pratica, quindi, non è il numeratore delle formule ad essere sbagliato ma il denominatore, che in realtà è molto più grande nel gruppo dei non vaccinati e molto più piccolo nel gruppo dei vaccinati, per cui l'incidenza "appare" (ma non è) più alta nei non vaccinati e più bassa nei vaccinati.

Ecco un esempio con dati del tutto inventati che dimostra come l'errore sia nel DENOMINATORE delle formule.

- Immaginiamo che in una popolazione di 20.000 persone ad un certo punto della pandemia e della campagna vaccinale ci siano 10.000 persone non vaccinate e 10.000 persone vaccinate con la prima dose e che in entrambi i gruppi ci siano stati 50 casi di covid.
- Immaginiamo che invece di una vaccinazione reale sia somministrata soluzione fisiologica (acqua e sale) che non ha alcun potere immunogeno.
- Immaginiamo inoltre che si vaccinino con soluzione fisiologica 500 persone al giorno.
- Immaginiamo per assurdo che non ci siano più casi di covid da questo momento in poi. In questo modo il numeratore non cambia più e si può capire l'effetto sul calcolo dell'incidenza che ha l'errore del calcolo del denominatore.

Fino a questo punto la situazione è la seguente, ovviamente l'incidenza è la stessa perché non si tratta di un vaccino ma di acqua e sale:

incidenza nei vaccinati con soluzione fisiologica= $(50/10.000) \times 100 = 0,5\%$

incidenza nei non vaccinati= $(50/10.000) \times 100 = 0,5\%$

Torniamo alla formula dell'Efficacia Vaccinale:

$$EV = [(incidenza\ nel\ gruppo\ controllo - incidenza\ nel\ gruppo\ test) / (incidenza\ nel\ gruppo\ placebo)] \times 100$$

Nel presente esempio

$$EV = [(0,5 - 0,5) / 0,5] \times 100 = 0\%$$

Il giorno successivo, come detto, essendo il covid assente ci saranno sempre i 50 casi iniziali nei due gruppi, ma si sono vaccinate altre 500 persone. Il sistema di sorveglianza tiene conto e calcola le due incidenze:

incidenza nei vaccinati con soluzione fisiologica= $[50 / (10.000 + 500)] \times 100 = 0,48\%$

incidenza non vaccinati= $[50 / (10.000 - 500)] \times 100 = 0,53\%$

L'incidenza nei vaccinati diminuisce e quella nei non vaccinati aumenta senza che ci siano stati nuovi casi di covid. L'efficacia vaccinale sarà:

$$EV = [(0,53 - 0,48) / 0,53] \times 100 = 9,4\%$$

Compare un'efficacia della vaccinazione senza che la ci siano stati casi di covid e senza che la vaccinazione sia in realtà una vaccinazione, solo per lo spostamento di 500 persone da un denominatore all'altro.

Dopo altri nove giorni (10 giorni totali), si sono vaccinate altre 4.500 persone, ed essendo il covid assente ci sono sempre i 50 casi iniziali nei due gruppi. Si calcola l'efficacia vaccinale della soluzione fisiologica al giorno 10:

incidenza nei vaccinati con soluzione fisiologica= $[50 / (10.000 + 5.000)] \times 100 = 0,33\%$

incidenza non vaccinati= $[50 / (10.000 - 5.000)] \times 100 = 1\%$

L'incidenza nei non vaccinati è ora tripla rispetto a quella nei vaccinati. L'efficacia vaccinale sarà:

$$EV = [(1,00 - 0,33) / 1,00] \times 100 = 67\%$$

Dopo altri dieci giorni (19 giorni totali), si sono vaccinate altre 5.000 persone, i casi sono sempre 50 in entrambi i gruppi, ma il flusso di persone non vaccinate che passano nel gruppo dei vaccinati cambia ulteriormente:

incidenza nei vaccinati con soluzione fisiologica= $[50/(10.000+9.500)] \times 100=0,26\%$

incidenza non vaccinati= $[50/(10.000-9.500)] \times 100=10\%$

L'incidenza nei non vaccinati è ora tripla rispetto a quella nei vaccinati. L'efficacia vaccinale sarà prossima al 100%:

$EV=[(10,00-0,33)/10,00] \times 100=96,7\%$

Conclusioni

Questa spiegazione aiuta a capire meglio il discorso di Fenton, che dovrebbe essere il punto di vista di TUTTI, in pratica l'uso di due flussi SEPARATI di dati, ossia il numero di vaccinazioni somministrate quotidianamente e il numero di infezioni in base allo stato vaccinale implica un GRAVE BIAS perché cambia il denominatore che aumenta da una parte e diminuisce dall'altra e quindi anche lasciando invariati i numeratori (ossia anche senza nessun nuovo caso di infezione) compare magicamente un'efficacia.

Ma c'è di più, perché Fenton fa notare che i casi di infezione nei vaccinati da meno di 21 giorni vengono ARBITRARIAMENTE assegnati al gruppo dei non vaccinati per cui, all'errore suddetto se ne aggiunge un secondo, ossia nei non vaccinati non solo diminuisce il denominatore che provoca un aumento arbitrario dell'incidenza anche in assenza di nuovi casi, ma c'è anche un aumento del numeratore MAGGIORE DI QUELLO CHE DOVREBBE ESSERE in realtà perché le infezioni nel gruppo di transizione non hanno ragione di essere attribuite ai non vaccinati.

Quindi, nella review di Fenton gli esempi con il miscategorization bias riguardano quanto appena scritto, mentre gli esempi con l'exclusion bias (il gruppo di transizione è escluso sia dal gruppo dei vaccinati che da quello dei non vaccinati) è a metà tra il miscategorization bias e il discorso dell'ultimo esempio.