

RIASSUNTO

Ogni anno la sindrome influenzale colpisce una ampia fetta di popolazione. Le conseguenze si traducono in disagi e costi elevati per tutta la comunità.

I casi di influenza (ILI) in Italia vengono monitorati attraverso l'osservazione di medici e pediatri di base volontari distribuiti sul territorio. Per stimare l'incidenza delle sindromi influenzali è necessario coinvolgere un campione di medici sentinella sufficiente a garantire la sorveglianza di almeno il 2% della popolazione.

Il nostro studio si propone di rilevare l'efficacia del vaccino antinfluenzale stagionale utilizzando le risposte di un segmento di popolazione che afferisce a due farmacie. Il nostro obiettivo è di studiare l'effetto su tutta la popolazione ma ottenere un campione coerente è per vari motivi impraticabile. Abbiamo scelto una sede inusuale come la Farmacia considerando comunque che i dati ricavati sono sbilanciati verso le donne e anziani.

Lo scopo è quello di osservare la relazione tra assunzione del vaccino anti-influenzale e l'avvenuta sindrome influenzale nel campione osservato.

Attraverso la compilazione di un semplice questionario distribuito in due farmacie della provincia di Firenze abbiamo raccolto 642 schede nei mesi di Maggio e Giugno 2013.

Abbiamo osservato che i risultati raccolti nelle due farmacie sono sufficientemente omogenei e che esiste una correlazione positiva tra la vaccinazione antinfluenzale stagionale e la sindrome influenzale. I vaccinati si sono ammalati di influenza con il 10% in più nei confronti dei non vaccinati, a parità di sesso e di età.

Gianandrea Lanzara (1), Sergio Segantini (2)

PAROLE CHIAVE

ILI, vaccinazione, efficacia

Vaccino e influenza stagionale

Luglio 2013

Obiettivi della ricerca

1. Osservare la relazione tra l'assunzione del vaccino anti-influenzale e la probabilità di contrarre l'influenza stagionale in un campione di utenti di due farmacie della provincia di Firenze.
2. Stabilire se tra il vaccino anti-influenzale e la probabilità di ammalarsi esiste una relazione di tipo causale

3. Studiare la variazione di questa relazione, o del possibile effetto causale, a seconda dell'età, del sesso, e della provenienza geografica degli utenti

Contesto della ricerca

Un breve questionario è stato somministrato agli utenti di due farmacie. Il questionario, compilato su base volontaria, fornisce informazioni sulle seguenti variabili: (i) l'utente ha contratto l'influenza nella passata stagione; (ii) l'utente si è vaccinato nella passata stagione; (iii) l'età dell'utente; (iv) il sesso dell'utente. Il campione originario conteneva 642 osservazioni. Dopo aver eliminato le osservazioni che contenevano valori mancanti nelle variabili di interesse, rimane un campione di 621 osservazioni.

Considerazioni metodologiche

Per attribuire una relazione causale all'effetto di un trattamento (nel nostro caso il vaccino) su un particolare esito (in questo caso la probabilità di contrarre l'influenza stagionale) è necessario confrontare due campioni (statisticamente) *identici da tutti i punti di vista, tranne il fatto che uno ha ricevuto il trattamento e l'altro no*.

Il metodo più comune per ottenere questa condizione è selezionare su base casuale il campione che riceverà il trattamento: la legge dei grandi numeri assicura che i due campioni avranno le stesse caratteristiche osservabili e non osservabili, entro un intervallo di confidenza minimo. Siccome i due campioni sono *ex-ante* identici, qualunque differenza osserviamo *ex-post* deve essere attribuita all'effetto del trattamento.

Il contesto della ricerca differisce da questo *setting* ideale almeno per due principali ragioni:

1. Il segmento di popolazione che osserviamo tramite il questionario (in altre parole, chi si è recato in farmacia e si è preso la briga di compilare il questionario) potrebbe essere diverso dalla popolazione nel suo complesso. Come vedremo, il campione a nostra disposizione è sbilanciato verso le donne e verso gli anziani. Il nostro obiettivo è studiare l'effetto del vaccino su tutta la popolazione, non solo su chi si reca in farmacia.
2. L'appartenenza al campione che ha ricevuto il trattamento non dipende da una procedura di selezione casuale gestita dal ricercatore, bensì da una scelta individuale.

Un terzo problema, comune a tutte le ricerche ma potenzialmente rilevante in questa sede, è l'errore di misurazione, in particolare per quanto riguarda la variabile relativa all'influenza stagionale. Nei prossimi paragrafi vorrei discutere più in dettaglio ciascuno di questi problemi e le loro conseguenze sull'interpretazione dei risultati. In particolare, nel paragrafo *a* parlerò dello sbilanciamento del campione. I paragrafi dal *b*, *c* e *d* sono dedicati al punto 2 e a questioni ad esso connesse. Il terzo problema, l'errore di misurazione, è discusso nel paragrafo *e*.

***a* Sbilanciamento del campione**

Tutto sommato, è lecito pensare che le farmacie siano frequentate da settori della popolazione ampi abbastanza da minimizzare la preoccupazione per lo sbilanciamento. Inoltre, l'età e il sesso sono

presumibilmente due discriminanti chiave tra chi va in farmacia e chi no, ed essendo queste due variabili osservate, la discrepanza tra campione e popolazione può essere in certa misura rimediata.

Il procedimento che permette di allineare il campione alla popolazione consiste nel *ripesare* ogni osservazione, attribuendole un peso maggiore o minore a seconda che sia sotto-rappresentata o sovra-rappresentata nel campione. Così facendo si assume implicitamente che sesso ed età siano le uniche determinanti della probabilità di recarsi in farmacia.

Per calcolare il peso è necessario conoscere la percentuale di donne e la distribuzione dell'età nella popolazione di interesse. Bisogna dunque chiedersi: qual è la popolazione di interesse? O meglio, a quale popolazione è lecito allineare il campione osservato? Volendo essere rigorosi, la "popolazione" include solo il quartiere che si serve abitualmente in quella farmacia. Cercando un compromesso tra il rigore statistico e la convenienza pratica, si può estendere la definizione di popolazione agli abitanti del comune o della provincia di Firenze (ma gli abitanti di un particolare quartiere saranno certamente diversi dagli abitanti del comune nel loro complesso: si pensi ad esempio all'età, al censo, alla nazionalità; e alcune di queste caratteristiche avranno anche a che fare con lo stato di salute: per questo problema si rimanda alla sezione successiva).

b Variabili omesse

Perché alcune persone hanno deciso di vaccinarsi e altre no? Chi sceglie di vaccinarsi avrà più fiducia nell'efficacia del vaccino, crederà di trarne benefici maggiori, oppure disporrà di più tempo per recarsi dal medico, magari perché non lavora, o abita più vicino, e così via. In generale, sorge la possibilità che i due campioni confrontati non differiscano solamente rispetto al trattamento, ma anche rispetto ad altre caratteristiche. Si badi che una differenza rispetto a una caratteristica non è un problema in tutti i casi, ma *solo se quella caratteristica è correlata con la probabilità di contrarre l'influenza*. Perché in quel caso non sapremo se la differenza osservata nella probabilità di contrarre l'influenza è da attribuire a quella caratteristica o al trattamento. Differenze nelle caratteristiche *osservate*, inoltre, possono essere eliminate mediante l'analisi di regressione, nel nostro caso il sesso e l'età; rimane dunque il problema delle caratteristiche *non osservate*.

Facciamo un esempio. Può darsi che le persone che decidono di vaccinarsi siano (in media) più attente alla propria salute e tendano perciò a mantenere uno stile di vita sano. Di conseguenza queste persone avranno in partenza una probabilità minore di ammalarsi, indipendentemente dall'efficacia del vaccino. In generale, tutte le caratteristiche correlate positivamente con la decisione di vaccinarsi e negativamente con la probabilità di ammalarsi, portano a sovrastimare l'efficacia del vaccino. Viceversa, tutte le caratteristiche correlate positivamente con entrambe le variabili porteranno a sottostimare l'efficacia del vaccino.

c Causalità inversa

Se oltre all'effetto di una variabile x su una variabile y , esiste un effetto della variabile y sulla variabile x , la relazione osservata tra quelle due variabili rifletterà la composizione di due effetti diversi. Nel nostro caso, può darsi che la probabilità stessa di ammalarsi, valutata *a priori* dal diretto interessato o dal medico curante, sia una determinante della decisione di vaccinarsi. Poniamo che il vaccino non abbia nessun effetto sulla probabilità di ammalarsi: osserveremo comunque una correlazione positiva tra le due variabili che sarebbe scorretto interpretare in senso causale. In generale, il problema della causalità inversa ci porterà a sottostimare l'efficacia del vaccino.

d Effetti eterogenei

Un terzo e ultimo problema sorge nel caso in cui l'effetto del trattamento non sia omogeneo nella popolazione. Potrebbe essere, ad esempio, negativo per alcune fasce di età e positivo per altre; potrebbe essere di maggiore o minore entità a seconda del singolo individuo. Si presume che le persone che decidono di vaccinarsi siano quelle che credono di trarre un beneficio maggiore dal vaccino. Calcolando l'efficacia del vaccino solo sul campione vaccinato otterremo quindi una stima al rialzo dell'efficacia nel complesso della popolazione: la stima ottenuta sarà corretta (al netto degli altri problemi) solo per coloro che decidono di vaccinarsi.

Tuttavia bisogna sottolineare che, per molti aspetti, quest'ultimo è l'effetto più rilevante per chi prende decisioni politiche, perché sarà quello che si realizzerà concretamente nella società; è bene però tenere distinte le interpretazioni.

e Errore di misurazione

E' ciascun individuo a dirci se ha avuto l'influenza stagionale o no, non la diagnosi di un medico. In alcuni casi le persone risponderanno in base alla diagnosi di un medico che le ha visitate, ma non è detto che sia così. In generale, possiamo fidarci che le persone sappiamo riconoscere correttamente l'influenza stagionale, e non la scambino con altre alterazioni dello stato di salute?

Quando l'errore di misurazione è solo “rumore”, cioè non è correlato con le variabili indipendenti, esso tende a comprimere le stime verso lo zero, un effetto noto come “attenuazione”. Se invece ha una direzione sistematica, può ingrandire o rimpicciolire le stime.

Faccio due esempi.

1. Chi si è vaccinato è fermamente convinto di non poter prendere l'influenza, quindi, anche se compaiono i sintomi, la scambia per un malessere passeggero. In questo caso si sovrastimerà l'efficacia del vaccino.
2. Chi si è vaccinato ha la pretesa di non ammalarsi, non solo di influenza, ma anche di altre frequenti patologie invernali. E scambia per sintomi influenzali anche semplici mal di testa o colpi di tosse, frequenti durante l'inverno. In questo caso si sottostimerà l'efficacia del vaccino.

Naturalmente sono esempi narrativi, utili a illustrare cosa si può immaginare che accada, piuttosto che ciò che accade realmente.

Alla luce di queste considerazioni, possiamo adesso all'analisi statistica dei dati: per via dei problemi appena discussi, non parlerò di effetto casuale, bensì di correlazione.



Analisi descrittiva

f Caratteristiche generali del campione

Nella tabella 1 sono riportate la media e l'errore standard (tra parentesi) per le quattro variabili disponibili nella base dati. Il 62% del campione è composto da donne, l'età media è pari a 57 anni circa. Poco meno del 40% del campione si è vaccinato mentre circa il 20% ha contratto l'influenza stagionale.

Tabella 1. Descrizione delle variabili				
	% donne	età media	% vaccino	% influenza
Tutto il campione				
	0.62	57.11	0.38	0.19
	0.02	0.72	0.02	0.02
Farmacia Le Mura				
	0.62	58.73	0.37	0.19
	0.03	1.07	0.03	0.02
Farmacia Ninci				
	0.61	55.78	0.39	0.19
	0.03	0.96	0.03	0.02

Se dividiamo il campione nelle due farmacie in cui è stato distribuito il questionario, non compaiono differenze degne di nota tra le variabili. Le lievi discrepanze non sono significative dal punto di vista statistico.

E' utile dare un'occhiata non solo alla media ma all'intera distribuzione della variabile età. Per rendere più chiara l'analisi ho diviso il campione in cinque fasce di età. Come si vede nel grafico 1, i due terzi circa del campione ha più di cinquant'anni, mentre appena il 2% del campione ha meno di diciotto anni.

La composizione del campione per sesso e per età può riflettere il fatto che le persone più anziane fanno un uso più frequente di medicinali, e che il compito di fare acquisti in farmacia, nelle famiglie, è molto spesso delegato alla donna. Come si è detto in precedenza, la ponderazione del campione permette di allineare queste percentuali con quelle della popolazione.

Grafico 1



g Differenze tra vaccinati e non vaccinati

Differenze tra vaccinati e non vaccinati			
	% donne	età media	influenza
NON si è vaccinato	0.69	50.67	0.17
	0.02	0.87	0.02
Si è vaccinato	0.54	67.81	0.22
	0.03	0.96	0.03

Diamo adesso uno sguardo preliminare alle caratteristiche demografiche e alle percentuali di ammalati nei due principali sotto-campioni di interesse. La tabella 2 mostra che il 69% dei non vaccinati è di sesso femminile contro il 54% dei vaccinati. Questo indica che le donne, in proporzione, si vaccinano meno degli uomini. L'età media, inoltre, è molto più alta tra i vaccinati che tra i non vaccinati: quasi 68 anni tra i primi contro 50 tra i secondi. Infine, tra i vaccinati la percentuale di ammalati di influenza è leggermente superiore: 22% contro il 17% di ammalati tra i non vaccinati. Quest'ultima differenza, tuttavia, non è statisticamente significativa.

h Vaccino ed influenza per fascia d'età

Prima di passare all'analisi di regressione, nella quale calcoleremo la correlazione tra il vaccino e l'influenza a parità di sesso e di età, vediamo come la decisione di vaccinarsi e la probabilità di contrarre l'influenza variano per fascia d'età.

La percentuale di vaccinati, ci dice il grafico 2, aumenta in maniera vistosa con l'età. E' pari al 14% tra i 35-49enni, al 35% tra i 50-69enni, quasi il doppio, e raddoppia ulteriormente nella fascia di età successiva, raggiungendo il 70% tra gli ultrasessantenni. La percentuale di vaccinati in realtà diminuisce in un primo momento, tra i minori di diciotto anni e i 19-34enni, come sarebbe naturale aspettarsi se i bambini, ad esempio, vengono vaccinati più di frequente a scopo precauzionale. Tuttavia lo scarso numero di osservazioni in questa parte del campione non permette di trarre conclusioni stabili: gli intervalli di confidenza, rappresentati nel grafico per mezzo di barrette rosse, sono estremamente ampi, e si sovrappongono a quelli di tutte le fasce di età successive, tranne quella degli ultra-settantenni.

Grafico 2

La percentuale di persone che hanno contratto l'influenza ha un andamento in parte opposto. Stabile attorno al 23% fino alla fascia dei 35-49enni, essa scende al 20% tra i 35-49enni, fino a toccare il 13,5% tra gli ultrasessantenni. Tuttavia, come risulta dalla sovrapposizione degli intervalli di confidenza, nessuna di queste percentuali è significativamente diversa dalle altre. Almeno a partire da un rozzo confronto tra fasce di età, quindi, non emerge nessuna correlazione evidente tra l'assunzione del vaccino e la probabilità di contrarre l'influenza stagionale.

Analisi di regressione

i Risultati di base

L'analisi di regressione permette di calcolare la correlazione tra due variabili (nel nostro caso la probabilità di contrarre l'influenza e la vaccinazione) *tenendo costanti* gli altri fattori (nel nostro caso il sesso e l'età). In altre parole, consente di misurare quanto varia la probabilità di ammalarsi, in seguito all'assunzione del vaccino anti-influenzale, per due individui dello stesso sesso e di età uguale.

Prima di analizzare i risultati, è necessaria una precisazione. La variabile indipendente è una variabile binaria, che assume valore 1 nel caso che l'individuo abbia avuto l'influenza, e 0 nel caso contrario. Parlare di "probabilità di ammalarsi" significa perciò definire la probabilità, come è comune fare, come la percentuale di casi di influenza nel campione, così come la probabilità di ottenere testa con il lancio di una moneta equivale al numero di realizzazioni favorevoli diviso il totale delle realizzazioni. Ho scelto inoltre di utilizzare *il modello di probabilità lineare*, ovvero di considerare la probabilità come una combinazione lineare delle variabili esplicative. Questo modello ha lo svantaggio che la probabilità stimata in certi casi può cadere al di fuori dell'intervallo 0-1, il che è impossibile; tuttavia esso è di gran lunga più immediato da interpretare rispetto a modelli più rigorosi come il *Probit* e il *Logit*, perché permette di interpretare i coefficienti direttamente come effetti sulla probabilità, e questo senza che venga alterata la sostanza delle conclusioni (la significatività e la magnitudine dei coefficienti).

Scorriamo adesso la tabella dei risultati. La prima sezione contiene le stime per l'intero campione, mentre la seconda e la terza ripetono l'analisi per ciascuna farmacia separatamente. La prima colonna riporta la correlazione incondizionata (cioè senza controllare per altri fattori) tra il vaccino e l'influenza: essa è pari a 0,048, positiva ma non statisticamente significativa. Come ho spiegato nel paragrafo precedente, questo coefficiente può essere interpretato direttamente come una probabilità: la probabilità di ammalarsi è del 5% *più alta* tra chi si è vaccinato. Controllare per il sesso dell'individuo, come si legge nella seconda colonna, non altera i risultati. Al contrario, quando si controlla per l'età, la correlazione tra vaccino ed influenza si rafforza notevolmente e diventa statisticamente significativa (livello di confidenza superiore al 99%): la probabilità di contrarre l'influenza è del 10% *più alta* tra chi si è vaccinato, a parità di sesso e di età (terza colonna). Il coefficiente dell'età (-0.003) è negativo e statisticamente significativo: questo era già stato illustrato con il grafico 2. Dieci anni in più diminuiscono la probabilità di ammalarsi di influenza del 3%.

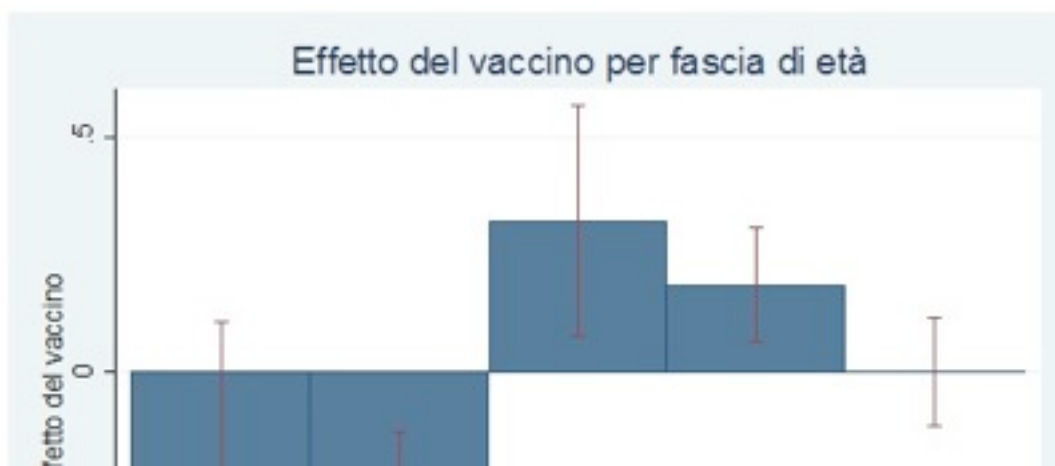


Tabella2. Variabile Dipendente: Probabilità di contrarre l'influenza stagionale									
	Tutto il campione			Farmacia Ninci			Farmacia Le Mura		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
vaccinato	0.048	0.054	0,101***	0.011	0.019	0.076	0.091*	0.095*	0.133**
	(0.033)	(0.034)	(0.038)	(0.044)	(0.045)	(0.051)	(0.050)	(0.051)	(0.058)
donna		0.041	0.039		0.038	0.036		0.040	0.038
		(0.033)	(0.032)		(0.045)	(0.044)		(0.048)	(0.048)
età			-0.003***			-0.004***			-0.002
			(0.001)			(0.001)			(0.002)
costante	0.173***	0.145***	0.288***	0.184***	0.157***	0.340***	0.161***	0.134***	0.238**
	(0.019)	(0.029)	(0.063)	(0.027)	(0.041)	(0.081)	(0.027)	(0.041)	(0.100)
Osservazioni	621	621	621	334	334	334	287	287	287
Errori standard robusti tra parentesi									
*** p<0.01, ** p<0.05, * p<0.1									

Ripetendo l'analisi per ogni farmacia, emergono due fatti degni di nota: primo, la correlazione tra vaccino e influenza si concentra nella farmacia le Mura, nella farmacia Ninci non è mai statisticamente significativa; in secondo luogo, e a parziale correzione del primo, bisogna notare che l'andamento dei coefficienti e la loro grandezza (almeno nella terza colonna, dove sono incluse tutte le variabili di controllo) è simile nelle due farmacie. Di fatto, i coefficienti delle colonne (6) e (9) non sono significativamente diversi tra di loro. La struttura dei risultati insomma sembra essere più o meno coerente nelle due farmacie: le dimensioni ridotte dei campioni non permettono in questo caso di formulare conclusioni più solide.

***j* Correlazione tra vaccino e influenza per fascia di età**

Come varia la correlazione tra vaccino ed influenza con il variare dell'età? Ho stimato la regressione corrispondente alla colonna (3) della tabella precedente, separatamente per ogni fascia di età. Per rendere i risultati più leggibili, ho riportato i coefficienti stimati in un grafico, insieme ai rispettivi intervalli di confidenza.

Grafico 3

Il grafico 3 rivela un andamento dei risultati molto particolare. La correlazione tra vaccino e influenza è negativa per le prime due fasce di età, anche se dato il numero esiguo di minorenni presenti nel campione, è significativa solo per i 18-34enni. Al contrario, è significativa e positiva per i 34-49enni e i 50-69enni. Per gli ultra-settantenni, invece, è indistinguibile da zero. Si veda il grafico 6 per un'illustrazione ancora più evidente di questo risultato.

Di seguito, presento anche gli stessi grafici prodotti separatamente per ogni farmacia. I campioni si assottigliano a tal punto che i risultati non sono molto informativi, è utile però vedere se emergono delle incongruenze palesi con il grafico 3. A prima vista, nonostante l'ampiezza degli intervalli di confidenza, i risultati sono coerenti sia tra le due farmacie che con il grafico 3.

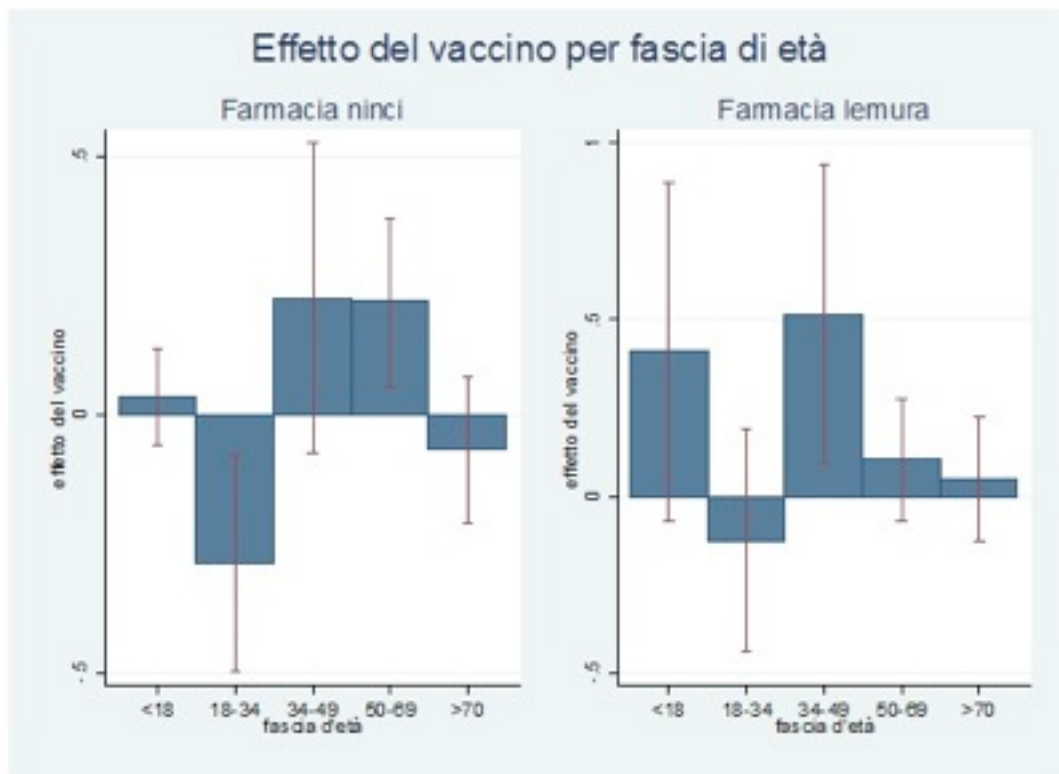


Grafico 4

***k* Correlazione tra vaccino e influenza per gli uomini e per le donne**

La correlazione tra vaccino e influenza pressoché identica per gli uomini e per le donne (grafico 5), positiva e per i primi statisticamente significativa al livello di confidenza del 5% (ma per le donne è comunque significativa al 10%). La leggera perdita di significatività in questo caso è da attribuirsi al dimezzamento del campione.

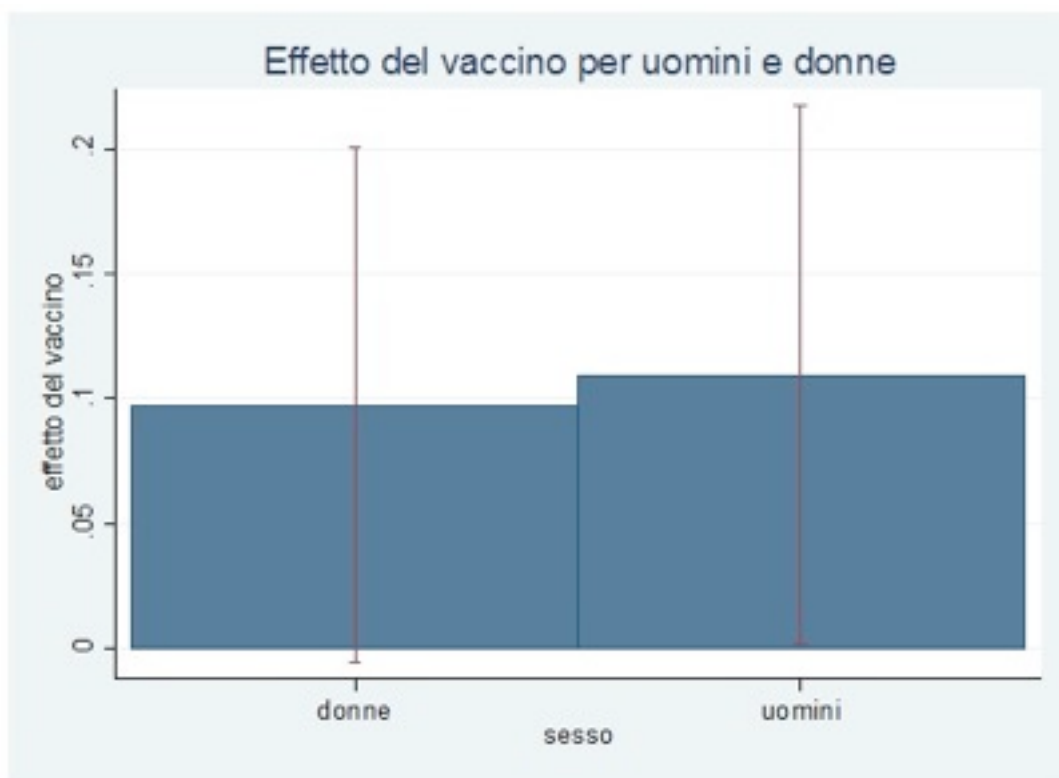


Grafico 5

1 Definizione alternativa delle fasce di età

L'analisi svolta fino a qui poggia su delle fasce di età definite in maniera arbitraria. Per assicurarci che i risultati ottenuti non dipendano esclusivamente da questa definizione, ho prodotto i grafici 2 e 3 con fasce di età più fini, di dieci anni ciascuna, tranne la prima e l'ultima che comprendono rispettivamente i minori di quindici anni e gli ultra-settantacinquenni.

Senza discutere in dettaglio questi nuovi grafici, mi limito a notare come la struttura dei risultati appaia sostanzialmente confermata. Come prima, i numeri relativi alle prime due fasce di età vanno presi con beneficio di inventario, dato il numero esiguo di osservazioni disponibili. Tuttavia l'andamento a "U rovesciata" dell'effetto del vaccino appare ancora più evidente nel grafico 7.

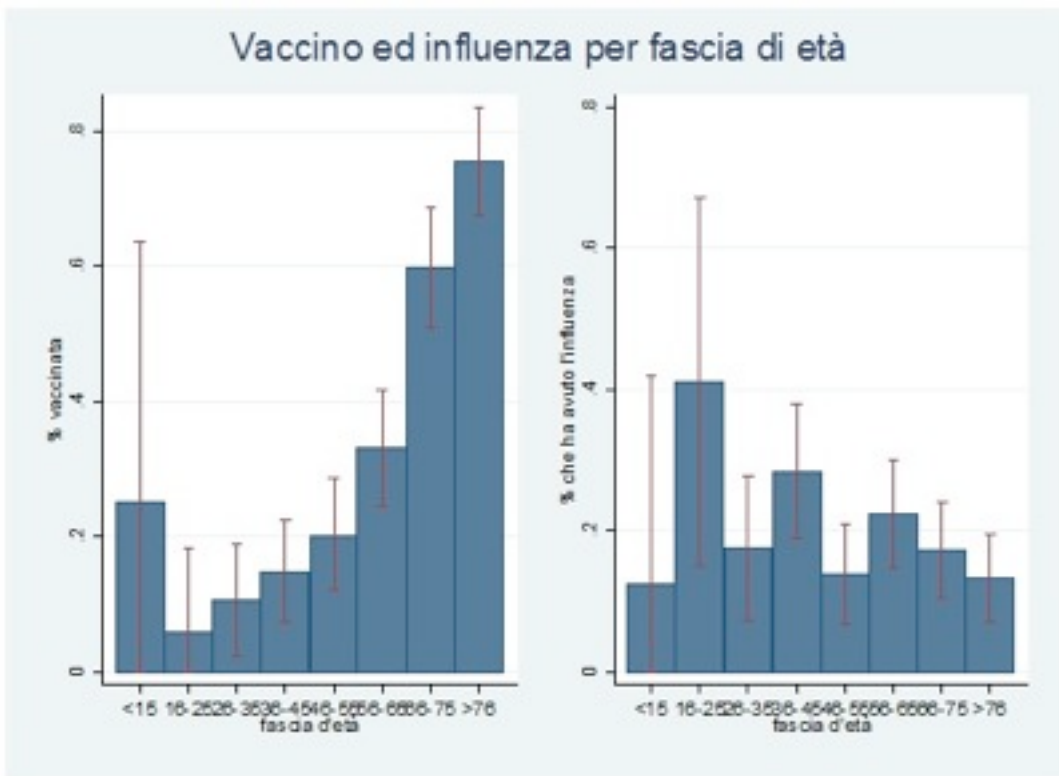


Grafico 6

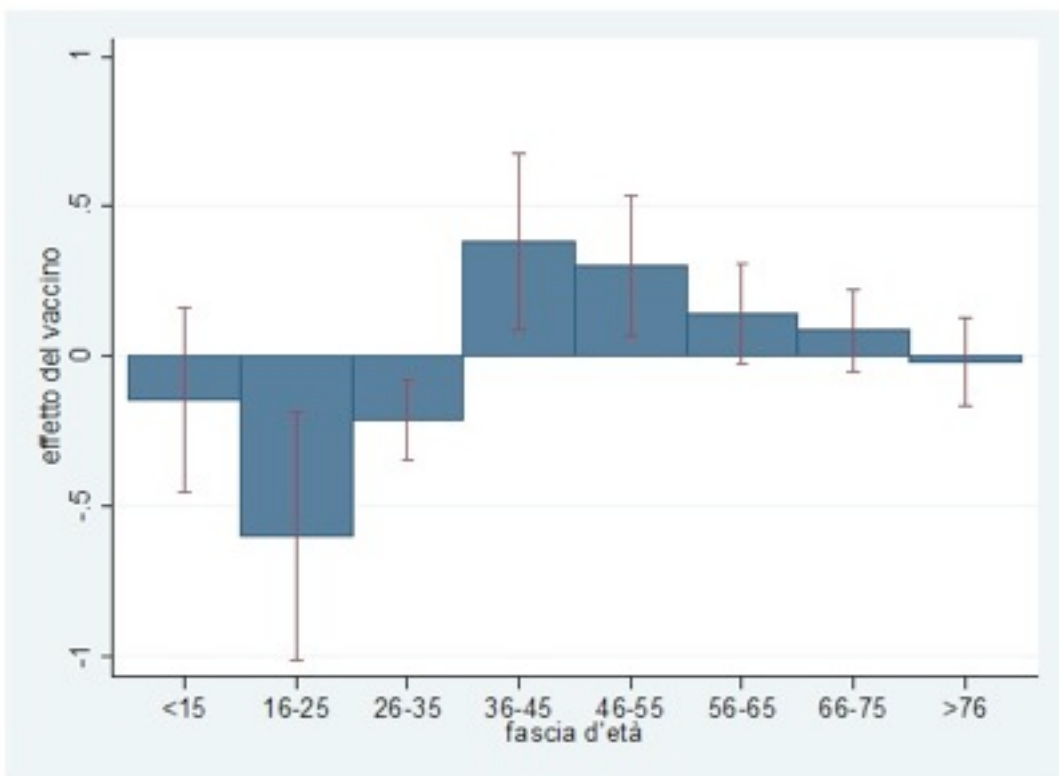


Grafico 7

Conclusioni

Per riassumere, mi pare che da questa ricerca si possono trarre tre conclusioni principali:

1. Esiste una netta correlazione positiva tra l'essersi vaccinati e l'aver contratto l'influenza stagionale. I vaccinati si sono ammalati di influenza con 10% di frequenza in più rispetto ai non vaccinati, a parità di sesso e di età.
2. Questa correlazione varia in maniera significativa col variare dell'età: in particolare, il coefficiente mostra un chiaro andamento a "U rovesciata". Al contrario non si ravvisano differenze tra i due sessi.
3. I limiti imposti all'analisi dal contesto della ricerca e dai dati non consentono di interpretare questa correlazione in senso causale. Sarebbe facile criticare dal punto di vista statistico una tale interpretazione, visto che non è possibile adottare contromisure per nessuno dei comuni problemi (variabili omesse, causalità inversa, etc.) che ostacolano questo tipo di analisi.

In merito al punto 3, aggiungo che in certi casi è possibile stabilire se la correlazione misurata è maggiore o minore dell'effetto causale. Bisogna capire se si può attribuire un segno univoco ai vari effetti distorsivi che agiscono sui risultati. Ho già abbozzato questo tipo di ragionamento nella sezione relativa alle considerazioni metodologiche. Ritorniamoci brevemente sopra:

- Effetti eterogenei: se si presume che chi decide di vaccinarsi si attende di trarne un beneficio *maggiore* rispetto a chi non si vaccina, si deve anche concludere che per questi ultimi l'efficacia del vaccino dev'essere ancora *minore* di quella che abbiamo misurato;
- Variabili omesse: inoltre, è più facile pensare a variabili omesse, come la cura per la propria salute, che da un lato aumentano la propensione a vaccinarsi e dall'altro rendono meno vulnerabili all'influenza, piuttosto che variabili che esercitano l'effetto inverso. Se queste variabili fossero incluse nell'analisi, allora, l'efficacia del vaccino sarebbe ulteriormente ridotta;
- Causalità inversa: tuttavia, il problema della causalità inversa va nella direzione opposta: le persone che hanno assunto il vaccino, si può immaginare, avevano già in partenza una maggiore probabilità di ammalarsi.
- Errore di misurazione: l'errore di misurazione può giocare sia a favore che a sfavore dell'efficacia del vaccino. A favore, se le persone vaccinate dichiarano, sbagliandosi, di non aver avuto l'influenza anche quando l'hanno avuta; a sfavore in caso contrario.

Poiché le forze distorsive hanno segni differenti, e non avendo strumenti per stabilire *a priori* quale

forza prevale, è difficile anche fare ipotesi sul rapporto tra la correlazione osservata e l'effetto causale.

Come affinare la metodologia di ricerca e renderne più solide le conclusioni? Come ultima nota, vorrei passare brevemente in rassegna le principali opzioni:

1. Randomizzazione: dal punto di vista statistico è sicuramente l'opzione più sicura. Quando il trattamento è assegnato su base casuale, non possono sussistere differenze significative tra il campione trattato e il campione di controllo, e ogni discrepanza nella variabile dipendente è attribuibile con certezza al trattamento. Purtroppo è anche il metodo più difficile da implementare.
2. Variabili strumentali: si tratta di una variabile correlata con la propensione a vaccinarsi ma **non** con la vulnerabilità all'influenza. Ad esempio, la presenza di un'altra specifica patologia che aumenta i rischi per la salute *in seguito* all'influenza stagionale. Simili variabili permettono di indurre una variazione nella propensione a vaccinarsi che non ha *nulla a che vedere* con la probabilità di ammalarsi, riproducendo così (con certi distinguo, nei quali non mi dilungo) le condizioni di randomizzazione. In questo caso, bisognerebbe aggiungere al questionario una domanda che richieda le informazioni relative alla variabile strumentale prescelta.
3. Discontinuità: si possono sfruttare le discontinuità nella propensione a vaccinarsi per studiare se ad essere corrispondono altrettante discontinuità nella probabilità di ammalarsi, e dall'entità del "salto" risalire a una stima dell'efficacia del vaccino. Per esempio, una discontinuità potrebbe essere indotta dalla gratuità del vaccino in seguito al compimento del 65esimo anno di età. Questo metodo consente stime molto precise, ma ha alcuni difetti: in primo luogo, richiede un campione numeroso intorno alla discontinuità; in altre parole: alcune centinaia di persone comprese tra i 50 e i 70 anni, indicativamente. In secondo luogo, i risultati ottenuti sono validi solo nel punto di discontinuità, in questo caso per i 65enni; visto che l'efficacia del vaccino, quale essa sia, sembra variare notevolmente con l'età, non è chiaro quale sarebbe il contenuto informativo di una tale stima.