



Valutazione casi-controlli dell'efficacia dello screening

Case-control evaluation of screening efficacy

Franco Berrino,¹ Gemma Gatta,¹ Paolo Crosignani²

¹Unità di epidemiologia eziologica e prevenzione, Dipartimento di medicina preventiva e predittiva, Istituto nazionale per lo studio e la cura dei tumori di Milano

²Unità di epidemiologia ambientale e Registro tumori, Dipartimento di medicina preventiva e predittiva, Istituto nazionale per lo studio e la cura dei tumori di Milano

Corrispondenza: Franco Berrino, Dipartimento di medicina preventiva e predittiva, Istituto nazionale per lo studio e la cura dei tumori di Milano, via Venezian 1, 20133 Milano; tel. 02 23903515; e-mail: berrino@istitutotumori.mi.it

L'efficacia dello screening è dimostrabile con studi sperimentali, che confrontino l'effetto sulla mortalità in popolazioni sottoposte e non sottoposte a screening. Lo stesso test di screening dimostratosi efficace in un dato periodo in una popolazione determinata può non essere efficace in un altro periodo o in un'altra popolazione. Lo screening mammografico, per esempio, si è rivelato efficace in studi controllati e randomizzati condotti a New York negli anni Sessanta e in Svezia in anni più recenti, ma non in uno studio condotto in Canada.¹ La differenza dei risultati può essere dovuta a numerosi fattori, che vanno dalla diversa sensibilità del test condotto da mani diverse, alla diversità delle popolazioni in esame, in particolare alla differente distribuzione degli stadi clinici della malattia in assenza di screening, a sua volta dipendente dall'attenzione culturale delle donne alla diagnosi precoce clinica.

I programmi organizzati di screening hanno sviluppato una serie di indicatori di *performance* e standard di adeguatezza, alcuni dei quali sono condizioni di efficacia, cioè se non sono raggiunti lo screening deve essere considerato poco o per niente efficace. Si tratta per esempio di indicatori che riflettono la sensibilità del test, quali il riscontro di carcinomi mammari al primo vaglio, e il tasso di incidenza nell'intervallo tra vagli successivi di screening. Una buona *performance* non è però prova di efficacia. Per provare l'efficacia occorre dimostrare che si riduce la mortalità, o l'incidenza di forme avanzate. La ripetizione di studi randomizzati in ogni popolazione non è proponibile per ragioni etiche, di costi e di fattibilità. Dove si inizia un nuovo programma in una popolazione non ancora molto 'contaminata' da attività di screening non organizzate si può procedere a un'introduzione programmata dello screening con tempi diversi in aree o generazioni differenti e valutare se dove si è iniziato prima si vedono prima gli effetti sull'incidenza e sulla mortalità.² Dove invece attività di screening sono in corso da anni con una copertura abbastanza alta, si può procedere a un monitoraggio dell'efficacia con la tecnica di *incidence-based mortality*³ o con la tecnica casi-controlli. Gli studi casi-controlli sull'efficacia degli screening presentano però grandi complessità metodologiche e potenzialità di distorsione. Lo scopo di questa nota è discutere questa complessità e proporre un disegno di studio e di analisi che eviti gravi distorsioni. Verranno trattate esclusivamente le distorsioni degli studi casi-controlli specifiche dello screening.

Disegno generale degli studi casi-controlli per la valutazione di efficacia degli screening

Ipotizziamo di voler condurre uno studio casi-controlli di valutazione di efficacia dello screening mammografico. I casi saranno tutte le donne che muoiono di carcinoma mammario in un dato periodo in una determinata popolazione. Per ogni caso si campioneranno, come controlli, una o più donne appartenenti alla stessa popolazione che ha generato i casi, della stessa classe di nascita, non affette da carcinoma mammario al momento della diagnosi del caso, perché solo chi non ha la (diagnosi della) malattia può essere sottoposto a screening. L'esposizione sarà definita in base alla storia di screening nel periodo in cui ciascun caso era teoricamente svelabile dal test di screening (la così detta *detectable pre-clinical phase* – DPCP – detta anche *sojourn time*, che per ora ipotizziamo di durata nota) e nello stesso periodo di tempo per i loro controlli appaiati. Eventuali test mammografici eseguiti prima della DPCP non dovrebbero essere considerati nella storia di esposizione perché privi di ogni potenzialità preventiva. La loro eventuale inclusione diluirebbe la stima di efficacia. La proporzione di donne che si sono sottoposte a screening prima della DPCP, condizionatamente ai fattori di confusione, dovrebbe risultare la stessa in casi e controlli. Una diversa proporzione farebbe sospettare che anche il confronto dell'esposizione nella DPCP, cioè nell'unica finestra temporale in cui lo screening può essere efficace, non sia valido (si veda il paragrafo «Esposizione definita come abitudine a sottoporsi ai controlli di screening»).

Facendo riferimento al disegno generale degli studi casi-controlli riportato in figura 1, dove i casi sono tutti i casi (decessi per carcinoma mammario) manifestatisi nella popolazione in studio e i controlli un campione rappresentativo della popolazione che ha generato i casi, l'efficacia preventiva dello screening sarà definita dalla proporzione di casi prevenuti sul totale dei casi attesi in assenza di screening, e l'efficacia potenziale, ovvero l'efficacia dello screening se tutta la popolazione vi si fosse sottoposta, sarà stimata dalla frazione di casi prevenuta negli esposti, cioè dal rapporto fra i casi prevenuti e il numero di casi esposti che ci si sarebbe potuti attendere se lo screening non fosse efficace:

■ *casi esposti osservati* = a

■ *casi esposti attesi in assenza di efficacia* = bc/d

- *casi prevenuti* = osservati - attesi = $bc/d - a$
- *efficacia potenziale* = casi prevenuti / casi attesi nella popolazione sottoposta a screening = $(bc/d - a)/(bc/d) = 1 - RR$
- *efficacia* = casi prevenuti / casi attesi nella popolazione in assenza di screening = $(bc/d - a)/(bc/d + b) = (1 - RR) c/(c+d)$, dove $c/(c+d)$ esprime la copertura della popolazione.

Dalla stessa tabulazione, conoscendo la frazione di casi esposti che erano risultati positivi (p) al test di screening, si può derivare la sensibilità del test nei confronti dei casi di carcinoma mammario potenzialmente letali:⁴

- *sensibilità* = veri positivi / (veri positivi + falsi negativi) = (casi prevenuti + positivi non prevenuti) / casi esposti attesi in assenza di efficacia = $(bc/d - a + ap)/(bc/d) = 1 - RR(1 - p)$.

Definizione dell'esposizione

Facendo riferimento allo schema della storia naturale della malattia riportata in figura 2, dove la DPCP, che continuiamo a supporre di durata nota, è indicata da una linea più spessa, dobbiamo considerare due casi:

- gli accertamenti diagnostici (con o senza mammografia) vennero iniziati su richiesta della paziente in seguito alla comparsa di sintomi;
- la presenza di un tumore del tutto asintomatico fu svelata da una mammografia di screening (cioè da una mammografia effettuata a scopo preventivo).

Supponiamo per il momento che questa distinzione sia semplice. Se la mammografia avesse una sensibilità del 100%, se cioè fosse capace di svelare tutti i casi in fase preclinica svelabile (la DPCP, la cui durata peraltro è funzione della sensibilità), i casi diagnosticati in seguito a sintomi non potrebbero essersi sottoposti a screening nella DPCP, altrimenti sarebbero stati diagnosticati prima (i casi a diventerebbero b).

Analogamente i casi diagnosticati in seguito a mammografia di screening non potrebbero avere avuto mammografie precedenti nella DPCP perché altrimenti sarebbero stati diagnosticati prima. I controlli, invece, non essendo affetti dalla malattia, potrebbero essersi sottoposti a un numero illimitato di mammografie nell'intervallo corrispondente alla DPCP del caso a cui sono appaiati. Se ne deduce che il confronto della frequenza di esposizione nei casi e nei controlli non può che essere di tipo esposto/mai esposto (cioè «tutto o nulla», *ever screened versus never screened*) nella DPCP senza alcuna possibilità di graduare l'esposizione in termini di numero di mammografie. I casi di tipo b (diagnosticati in seguito a screening) saranno classificati esposti e i casi di tipo a (diagnosticati in seguito a sintomi) non esposti.

Se invece la sensibilità della mammografia fosse inferiore al 100%, nella DPCP potrebbero essere state eseguite delle mammografie con risultato falso negativo. In questo caso anche un caso di tipo a con precedente mammografia negativa verrebbe considerato esposto. Nella pratica sono state proposte e utilizzate svariate altre definizioni di esposizione,⁵ ma

| Esposizione | Si | No |
|-------------|----|----|
| casi | a | b |
| controlli | c | d |

Casi esposti attesi = bc/d
 Casi prevenuti = $bc/d - a$
 Efficacia potenziale = $(bc/d - a)/(bc/d) = 1 - RR$
 Efficacia = $(bc/d - a)/(bc/d + b) = (1 - RR) c/(c+d)$

Figura 1. Schema generale del disegno e della logica di analisi degli studi casi-controlli per la valutazione di efficacia degli screening.

Figure 1. Basic design and analysis of case-control studies of screening efficacy.

l'applicazione simulata di un programma di screening a efficacia nota a popolazioni teoriche ha dimostrato che la definizione che proponiamo è l'unica formalmente corretta.⁶

Healthy screenee bias

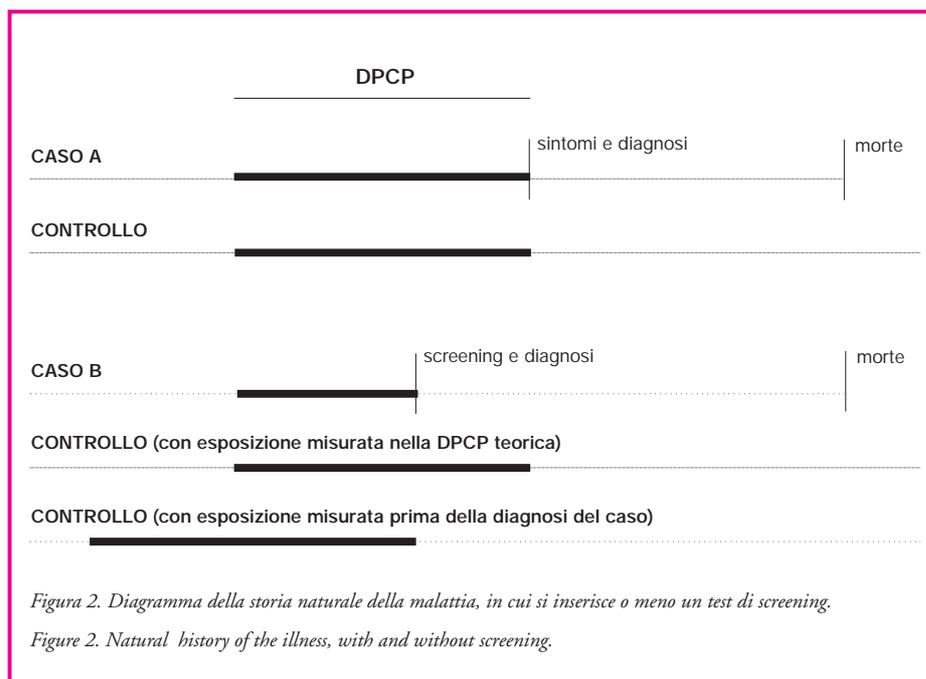
Qualora si volesse quantificare l'esposizione in termini di numero di mammografie effettuate nella DPCP, si configurerebbe l'errore denominato *healthy screenee bias*,⁷ dovuto al fatto che solo chi non ha la malattia può essere ripetutamente sottoposto a screening. La probabilità che un caso abbia una mammografia falsa negativa nella DPCP è funzione di $(1 - \text{sensibilità})$, cioè dell'ordine del 10%, la probabilità che ne abbia due è dell'ordine di $(1 - \text{sensibilità})^2$, la probabilità che ne abbia tre è praticamente nulla, mentre i controlli possono avere un qualunque numero di esami. Ne risulterebbe che una frequenza di mammografie dell'ordine di tre o più in un intervallo di tempo corrispondente alla durata della DPCP (dell'ordine di tre anni), cioè una frequenza di circa una mammografia all'anno, conferirebbe una protezione praticamente totale (nessun caso esposto contro un qualunque numero di controlli esposti), un risultato evidentemente non plausibile (vedi esempio di figura 3c). Qualunque sia la sensibilità (purché superiore a 0) c'è un periodo prima della diagnosi (la DPCP) in cui i casi non possono essersi sottoposti a screening con la stessa frequenza dei controlli, altrimenti presto o tardi il loro tumore sarebbe stato svelato, prevenendo così ogni ulteriore screening.

Durata della DPCP e misclassificazione dell'esposizione

Naturalmente la durata della DPCP dei singoli casi non è nota, per cui è necessario ricorrere a stime di durata media del tipo

[durata = (prevalenza al primo screening)/(incidenza x sensibilità del test)] oppure derivanti dalla modellizzazione dell'incidenza dei casi di intervallo rispetto all'incidenza in assenza di screening,^{8,9} L'uso di una stima di durata media comporta però inevitabilmente una certa misclassificazione dei casi e quindi una sottostima della protezione:

- vi saranno casi diagnosticati in seguito a sintomi dopo



una DPCP breve che risulteranno erroneamente classificati esposti perché sottoposti a mammografia prima della loro DPCP (ma entro la durata media delle DPCP);

- vi saranno casi diagnosticati in seguito a sintomi dopo una DPCP lunga classificati non esposti anche se effettivamente esposti (con mammografia falsa negativa) nel primo periodo della loro DPCP (prima dell'inizio della DPCP media);

- i casi diagnosticati in seguito a sintomi, per effetto del *length bias*, cioè della maggiore probabilità di riscontrare allo screening i casi con DPCP più lunga,¹⁰ hanno DPCP più breve della media; l'utilizzo di una durata media di DPCP porta quindi a una sovrastima dell'esposizione potenzialmente protettiva dei casi (perché casi sintomatici a DPCP breve vengono erroneamente classificati esposti) e quindi a una sottostima della protezione.

In una situazione di stabilità di offerta di screening l'esposizione media dei controlli non risulterà distorta dall'utilizzo di una DPCP media. In caso di offerta crescente, al contrario, l'esposizione dei controlli appaiati ai casi diagnosticati in seguito a screening potrebbe risultare sottostimata se, come viene abitualmente fatto, si considerano esposti i controlli sottoposti a mammografia in un tempo pari alla DPCP media prima della diagnosi del caso. Occorrerebbe infatti considerare l'esposizione in un intervallo di tempo corrispondente a quella che sarebbe stata la DPCP se la diagnosi non fosse stata anticipata dallo screening (vedi controlli del caso b in figura 2). Supponiamo, per esempio, che la durata media della DPCP sia di due anni, che il caso sia stato diagnosticato a metà della sua DPCP il 31 dicembre del 2000, e che l'offerta di screening sia in aumento. La probabilità che

il controllo sia stato esposto sarà sottostimata se si considera la sua storia di esposizione prima della diagnosi del caso (1999 e 2000) piuttosto che nel periodo effettivamente corrispondente alla DPCP (2000 e 2001). Se la sensibilità dello screening fosse costante nel corso della DPCP, mediamente l'identificazione del caso cadrebbe alla metà della DPCP, e il periodo di esposizione pertinente del controllo sarebbe metà prima e metà dopo la diagnosi. Ma se l'offerta è crescente questa procedura sovrastimerebbe l'esposizione dei controlli perché la sensibilità dello screening aumenta nel corso della DPCP (con la crescita del-

le dimensioni del tumore), per cui i casi sono diagnosticati più facilmente verso la fine che all'inizio della DPCP. In caso di offerta crescente nel periodo di studio sarà comunque opportuno stimare un intervallo di validità dei risultati procedendo a due definizioni di periodo di esposizione dei controlli appaiati ai casi diagnosticati in seguito a screening, ipotizzando che tutta la DPCP sia trascorsa prima della diagnosi del caso (sottostima dell'esposizione) oppure metà prima e metà dopo (sovrastima).

Analisi di sensibilità della finestra di esposizione e DPCP *shrinkage bias*

Teoricamente se la durata della DPCP è sovrastimata, l'efficacia dello screening sarà sottostimata, perché un certo numero di donne (casi e controlli) verrà classificato esposto in quanto sottoposto a mammografia solo prima della DPCP, quando l'esposizione non può essere efficace. Riducendo progressivamente la durata della DPCP nella definizione dell'esposizione si dovrebbe veder aumentare asintoticamente la stima della protezione mano a mano che ci si avvicina alla durata reale. Se però si giunge a sottostimare la durata della DPCP si ridurrà la frequenza di esposizione dei controlli (in particolare se si giunge a durate inferiori all'intervallo di screening raccomandato) senza ridurre ulteriormente la frequenza di esposizione dei casi (si ricorda che in caso di alta sensibilità del test di screening la frequenza di esposizione dei casi è sostanzialmente definita dalla frequenza dei casi diagnosticati in seguito a screening), per cui ci si attende nuovamente una sottostima della protezione. Infatti, mentre i casi diagnosticati in seguito a screening saranno classificati esposti qualunque sia

In uno studio casi-controlli¹² di valutazione di efficacia della ricerca del sangue occulto nelle feci in una coorte di soggetti che avevano accettato l'invito a sottoporsi a un primo test (con la raccomandazione di ripeterlo annualmente), e che erano stati poi seguiti per molti anni, sono stati considerati esposti:

a. coloro che avevano effettuato un secondo test entro due anni dal primo:

| | ripetizione del test entro 2 anni | |
|---------------------------------|-----------------------------------|-----|
| | si | no |
| deceduti per cancro intestinale | 16 | 79 |
| controlli | 109 | 366 |

OR = 0,64, con intervallo di confidenza al 95% 0,36-1,15 (aggiustato per storia familiare di cancro e storia personale di adenomi). Questo OR fu interpretato come un indicatore che la propensione a ripetere il test, come raccomandato dal programma, conferisce una protezione.

b. coloro che avevano effettuato uno o più test nella DPCP (analisi tutto o nulla) con diverse ipotesi di durata della DPCP:

| durata ipotetica della DPCP | OR | IC 95% |
|-----------------------------|------|-----------|
| ≤ 1 anno | 2,32 | 0,85-6,35 |
| ≤ 2 anni | 1,09 | 0,48-2,44 |
| ≤ 3 anni | 0,88 | 0,42-1,85 |
| ≤ 4 anni | 0,80 | 0,39-1,61 |
| ≤ 5 anni | 0,78 | 0,40-1,52 |
| ≤ 6 anni | 0,81 | 0,43-1,52 |

suggerendo una durata media della DPCP dell'ordine di 5 anni, e una protezione soltanto del 22%, inferiore al risultato dello studio della «propensione», a causa della misclassificazione dell'esposizione dentro e fuori della DPCP.

c. l'analisi basata sul numero di test effettuati prima della diagnosi, invece, suggeriva un forte *bias* di *healthy screenee*:

| numero di esami | OR | IC 95% |
|-----------------|------|-----------|
| solo 1* | 1 | |
| 2-3 | 0,71 | 0,44-1,14 |
| 4 o più | 0,34 | 0,16-0,75 |

*criterio di inclusione nello studio

Figura 3. Esempio di analisi.

Figure 3. Analysis sample.

a) by screening proneness. b) by different hypothetical duration of the detectable pre-clinical phase. c) by number of screening examinations (showing the effects of the healthy screenee bias).

il momento della DPCP in cui sono stati colti dallo screening, per i loro controlli appaiati si finirebbe per riconoscere valido per la definizione di esposizione un intervallo di tempo più breve (DPCP *shrinkage bias*). In un'analisi di sensibilità, la durata di DPCP che massimizza la protezione è verosimilmente la scelta migliore (vedi esempio di figura 3b).

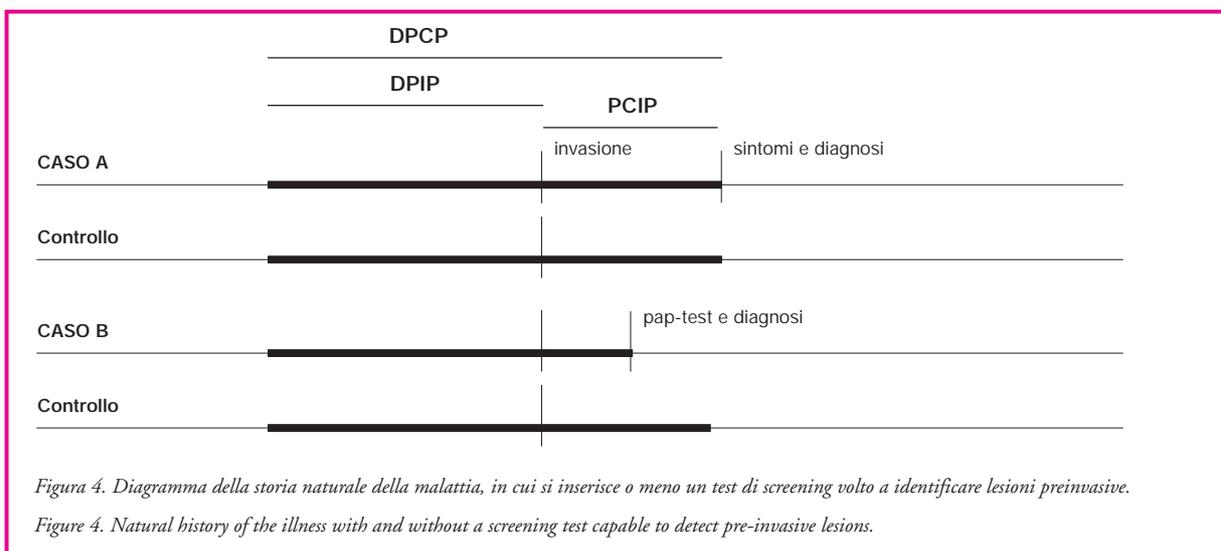
Misclassificazione dei test eseguiti in seguito a sintomi (*bias* di informazione)

Per definizione lo screening è rivolto a persone senza sintomi della malattia oggetto di screening. Una delle maggiori difficoltà nella definizione dell'esposizione è la distinzione tra i casi diagnosticati in fase del tutto asintomatica grazie a una mammografia di screening (casi «esposti» per definizione) e i casi in cui il ricorso alla mammografia era motivato dalla presenza di sintomi e di un sospetto diagnostico del curante. Se questi ultimi casi venissero per errore classificati esposti l'efficacia dello screening risulterebbe sottostimata.

La distinzione è abbastanza semplice in caso di screening organizzato: tutte le persone che hanno risposto all'invito partecipando a uno o più vagli di screening nel periodo corrispondente alla DPCP verrebbero considerate esposte, mentre le persone che non hanno risposto verrebbero classificate non esposte. Verrebbero però classificati esposti anche alcuni casi la cui partecipazione all'invito era motivata dalla percezione di un sintomo, e non esposti alcuni casi di tumore diagnosticati a seguito di una mammografia effettuata in fase asintomatica al di fuori del programma di screening, con conseguente sottostima dell'efficacia, tanto mag-

giore quanto maggiore fosse la pratica di mammografie preventive al di fuori del programma organizzato.

In assenza di programma di screening organizzato la distinzione dovrebbe basarsi su un'anamnesi accurata, impossibile nel caso di pazienti decedute. Un'alternativa è definire come casi le pazienti con metastasi accertate di carcinoma mammario invece che le pazienti decedute per carcinoma mammario. Intervistando le pazienti sarebbe teoricamente possibile ricostruire la storia delle manifestazioni della malattia, delle mammografie precedenti, e le motivazioni della prima mammografia positiva. È il disegno di studio proposto dal Centro studi e ricerche in medicina generale (CSeR-MEG) e dalla Società italiana dei medici di medicina generale (SIMMG)¹¹ per valutare se è utile prescrivere mammografie alle loro assistite al di fuori dei programmi organizzati. Rimarrebbero comunque casi in cui la paziente non è in grado di ricostruire la sequenza dei sintomi e delle manovre diagnostiche, nonché casi di difficile attribuzione, per esempio donne che segnalano un sintomo che induce la prescrizione di una mammografia che evidenzia una neoplasia ma in sede diversa rispetto al sintomo percepito dalla paziente, oppure la mammografia risulta negativa ma il tumore viene riscontrato l'anno successivo nella sede sospettata, e così via. Entrambi questi casi vanno considerati esposti. È possibile che anche tra i controlli vi siano donne che hanno effettuato una mammografia motivata da sintomi sospetti. Queste mammografie dei controlli contribuiscono all'esposizione perché non si trattava di sintomi di tumore oppure si trattava di falsi negativi per tumori destinati a manifestarsi dopo la diagnosi del caso.



Esposizione definita come attitudine a sottoporsi ai controlli di screening

Per evitare l'*healthy screenee bias*, la distorsione introdotta dalla misclassificazione di casi sintomatici, o della durata della DPCP, e il DPCP *shrinkage bias*, si possono classificare le persone in studio non in termini di esposizione nella DPCP, bensì in termini di persone che partecipano o meno al programma, seguendo o meno le raccomandazioni di periodicità. A parte i casi identificati al primo vaglio di screening, infatti, è l'aderenza al programma di controlli successivi che determina l'efficacia dello screening. Chi vi si sottopone regolarmente sarà più protetto di chi non vi si sottopone o di chi ripete il test saltuariamente. Un indicatore di accettazione, di attitudine a sottoporsi allo screening, può essere la densità di esami di screening in un certo periodo della vita, anche se deve trattarsi di un periodo precedente alla DPCP, in cui i casi possano essere ripetutamente sottoposti a screening come i controlli. Si tratta di un periodo in cui i test di screening non hanno di per se potenzialità preventiva; ciò nondimeno la frequenza di screening in questo periodo costituisce un indicatore della propensione a sottoporsi ai test, propensione che ha potenzialità preventiva. Si potrebbero per esempio stratificare le persone in studio a seconda che si siano sottoposte a una o più mammografie in un tempo certamente precedente alla DPCP (per esempio 5 anni prima dell'inizio della DPCP media) e classificare esposte («propense») quelle che hanno ripetuto la mammografia nei due o tre anni successivi (vedi figura 3a).

Restrignendo lo studio alle sole persone con uno o più test prima di una certa data si ridurrebbero anche eventuali distorsioni da autoselezione, per esempio, nel caso del carcinoma mammario, dovute alla minore propensione a sottoporsi a screening da parte delle persone a rischio più basso. Questo tipo di analisi consentirebbe anche di valutare l'effetto di diverse periodicità di screening.

Il caso dello screening per identificare lesioni preinvasive allo scopo di prevenire il cancro invasivo della cervice uterina

Ipotezziamo di pianificare uno studio casi-controlli per valutare l'efficacia del pap-test nella prevenzione del carcinoma invasivo della cervice. In questo caso lo screening è sia un fattore potenzialmente preventivo, se il test avviene prima dell'inizio della fase invasiva (nella *detectable pre-invasive phase* - DPIP), sia un determinante dell'evento stesso che si intende prevenire, quando porta alla diagnosi della malattia dopo l'inizio della fase invasiva (nella *pre-clinical invasive phase* - PCIP). Contrariamente al caso del carcinoma mammario (dove l'*outcome* dello studio è la morte) l'eventuale test di screening che porti alla diagnosi di carcinoma invasivo non è da includere nell'esposizione potenzialmente preventiva (non si può prevenire qualcosa che c'è già).

L'esposizione non sarà definita dalla presenza di (uno o più) test di screening in un dato intervallo di tempo prima della diagnosi del caso. C'è infatti un periodo, la PCIP, in cui non solo il pap-test non può prevenire, ma in cui ci si aspetta che i casi non siano stati sottoposti a screening, altrimenti la diagnosi sarebbe stata fatta prima.

Lo screening in fase invasiva preclinica (PCIP) anzi è da considerare un fattore di confusione (un determinante della diagnosi di carcinoma invasivo potenzialmente associato all'esposizione in studio, cioè all'esposizione nel tratto preinvasivo della DPCP, la *detectable pre-invasive phase* o DPIP) (figura 4). Per la definizione dell'esposizione potenzialmente preventiva e dei confonditori si rende quindi necessaria sia la stima della durata media della PIDP sia quella della PCIP.¹³ Chiunque abbia uno o più pap-test nella PIDP sarà da considerare esposto, sia che si tratti di pap-test negativi sia di pap-test che abbiano identificato lesioni preinvasive. Includere nel computo dell'esposizione solo i pap-test negativi, come cu-

riosamente hanno fatto vari studi, porterebbe a una sottostima dell'esposizione dei casi. Questi studi non hanno considerato che per essere preventivi i test di screening devono essere positivi (cioè devono individuare lesioni preinvasive curabili). L'analisi dovrà poi essere stratificata per presenza di pap-test nella PCIP.

Per sincronizzare i tempi in cui si valuta l'esposizione di casi e controlli, se il caso è stato diagnosticato in seguito a screening si può ipotizzare che in media sia stato diagnosticato a metà della PCIP (figura 4, caso b). Diversamente dal caso del carcinoma mammario, infatti, la sensibilità dello screening in fase invasiva non cresce necessariamente con il tempo, anzi un piccolo tumore inizialmente invasivo potrebbe essere più facile da riconoscere che non un tumore più avanzato in preda a necrosi. Per stabilire quale durata di PIDP e PCIP utilizzare si può procedere ad analisi di sensibilità e scegliere le durate associate alla massima protezione.

Si noti che non si possono inserire le neoplasie intraepiteliali tra i casi di uno studio casi controlli per la valutazione dell'efficacia dello screening cervicale. Le neoplasie intraepiteliali, infatti, non possono essere prevenute ma solo svelate dallo screening. Gli studi che le hanno inserite tra i casi hanno evidenziato una forte protezione apparente, perché ovviamente una donna con neoplasia intraepiteliale ha una bassa probabilità di essere stata sottoposta al test in un passato recente, altrimenti la neoplasia sarebbe stata riconosciuta prima (*healthy screenee bias*).¹⁴

Alcune prescrizioni

I punti seguenti sintetizzano le caratteristiche che deve avere uno studio casi controlli per evitare distorsioni, tra cui quelle discusse sopra:

1. definizione dei casi:

- solo manifestazioni cliniche (decesso, metastasi, o tumore invasivo se lo screening è volto a identificare lesioni preinvasive)
- non fattori di rischio o lesioni preinvasive riscontrabili solo allo screening
- periodo di studio definito dall'evento da valutare (decessi o metastasi occorsi dal... al...), non dal periodo in cui i casi sono stati diagnosticati, perché un certo numero di casi che sarebbero stati diagnosticati in quell'intervallo verrebbero esclusi se sottoposti a screening in precedenza, lasciando invece nello studio tutti i casi non sottoposti a screening (eventuali casi la cui diagnosi è stata anticipata nell'intervallo in studio potrebbero essere troppo recenti e non ancora deceduti);

2. definizione dei controlli:

- devono descrivere la frequenza di esposizione nella popolazione sottoposta a screening fino alla data di diagnosi dei casi, o meglio alla data in cui i casi sarebbero stati diagnosticati in assenza di screening;
- i controlli devono aver avuto la stessa opportunità dei casi di essere sottoposti a screening (nella valutazione di uno screening organizzato, per esempio, se si escludono i casi diagno-

sticati prima dell'invito a sottoporsi a screening devono essere esclusi anche i controlli non ancora invitati al tempo della diagnosi del caso a cui sono appaiati);

3. definizione dell'esposizione:

- in assenza di efficacia l'esposizione dovrebbe essere uguale per casi e controlli;
- negli studi che valutano la prevenzione della morte o della malattia avanzata (ma non negli studi che valutano la prevenzione del tumore invasivo) l'esposizione deve includere il test di screening che porta alla diagnosi del caso;
- deve escludere i test eseguiti dopo la comparsa dei sintomi (in fase clinica);
- deve includere sia risultati negativi sia positivi;
- la finestra di esposizione potenzialmente preventiva dovrebbe corrispondere alla DPCP (alla DPIP nel caso di screening volti alla prevenzione del tumore invasivo) e non deve essere né troppo lunga (per non diluire l'effetto con screening prima della DPCP) né troppo breve (per dare il tempo ai controlli di essere esposti). La durata ottimale è quella che massimizza la stima di efficacia;
- alternativamente all'effettiva esposizione nella DPCP si può definire l'attitudine, la proclività, allo screening, in base alla densità di screening in un periodo precedente alla DPCP.

Bibliografia

1. Vainio H, Bianchini F. *Breast Cancer Screening*. IARC Handbooks of Cancer Prevention vol 7. Lyon, IARC Press, 2002.
2. Hakama M, Pukkala E, Soderman B, Day N. Implementation of screening as a public health policy: issues in design and evaluation. *J Med Screen* 1999; 6: 209-16.
3. Paci E, Duffy SW, Giorgi D, et al. Quantification of the effect of mammographic screening on fatal breast cancer: the Florence Programme 1990-96. *Br J Cancer* 2002; 87: 65-69.
4. Berrino F, Bouchardy C, Sasco AJ. Screening for breast cancer in pre-menopausal women. Who should decide? *Int J Cancer* 1995; 60: 153-5.
5. Cronin KA, Weed DL, Connor RJ, Prorok PC. Case-control studies of cancer screening: theory and practice. *J Natl Cancer Inst* 1998; 90: 498-504.
6. Connor RJ, Boer R, Prorok PC, Weed DL. Investigation of design and bias issues in case-control studies of cancer screening using microsimulation. *Am J Epidemiol* 2000; 151: 991-8.
7. Morrison AS. *Screening in chronic disease*. Oxford, Oxford University Press, 1992.
8. Walter SD, Day NE. Estimation of the duration of a pre-clinical disease state using screening data. *Am J Epidemiol* 1983; 118: 865-86.
9. Duffy SW, Chen HH, Tabar L, Day NE. Estimation of mean sojourn time in breast cancer screening using a Markov chain model of both entry to and exit from the preclinical detectable phase. *Stat Med* 1995; 14: 1531-43.
10. Feinleib M, Zelen M. Some pitfalls in the evaluation of screening programs. *Arch Environ Health* 1969; 19: 412-15.
11. <http://www.csermeg.it/ricerca/index.html> ; <http://www.simg.it/step-mammografia>
12. Bertario L, Russo A, Crosignani P, Sala P, Spinelli P, Pizzetti P, Andreola S, Berrino F. Reducing colorectal cancer mortality by repeated faecal occult blood test: a nested case-control study. *Eur J Cancer* 1999; 35: 973-7.
13. Berrino F, Gatta G, d'Alto M, Crosignani P, Riboli E. Efficacy of screening in preventing invasive cervical cancer: a case-control study in Milan, Italy. *LARC Sci Publ* 1986; 76: 111-23.
14. Berrino F. Papanicolaou smears and risk of cervical neoplasia. *Lancet* 1984; ii: 1099-1100.