

I carcinomi di intervallo quali indicatori di performance di un programma di screening. Modalità e standard per la valutazione

Stefano Ciatto, Carlo Naldoni, Antonio Ponti, Livia Giordano, Alfonso Frigerio, Marcello Vettorazzi,

Lauro Bucchi, Luigi Bisanti, Marco Petrella, Eugenio Paci, Gianni Saguatti, Dolores Santini, Marco Rosselli

Del Turco, Marco Zappa, Paolo Giorgi Rossi, Vittorio Corsetti, Luisella Milanese, Priscilla Sassoli de' Bianchi

Premessa

Il GISMa, Gruppo Italiano Screening Mammografico, in preparazione e occasione del suo convegno annuale svoltosi a Peschiera del Garda l'11 e 12 ottobre 2007, ha promosso e completato la stesura di un documento sui carcinomi di intervallo (CI) nell'ambito del programma di screening mammografico, per la definizione sistematica di questo evento, fornendo indicazioni precise relative alle modalità di identificazione e di revisione delle mammografie corrispondenti. Tali raccomandazioni sono rivolte agli operatori che lavorano nei programmi di screening mammografico in Italia, che devono includere la valutazione dei carcinomi di intervallo tra le procedure correnti di controllo di qualità degli stessi.

Il GISMa ringrazia sentitamente gli autori e tutti coloro che hanno contribuito per iscritto alla stesura del documento e delle raccomandazioni finali contenute nello stesso. Il documento completo è scaricabile anche dal sito GISMa (www.gisma.it). Nel contempo, cogliamo l'occasione per ringraziare l'apporto prezioso di tutti coloro che hanno partecipato al convegno GISMa e al dibattito sull'argomento.

Presentazione

Il cancro di intervallo (CI) è un'entità patologica che viene descritta solo nell'ambito dei programmi di screening organizzato. In genere, agli occhi dell'opinione pubblica, e in particolare dei mass

media, assume un significato molto negativo, di grave errore del programma, con riflessi che possono diventare problematici sull'andamento del programma stesso. Ma anche gli operatori, e in particolare i radiologi, temono questo pur inevitabile fenomeno (legato sia ai limiti del test, quindi alla possibilità di avere falsi negativi, sia all'errore umano, sempre in agguato nella pratica medica) per i suoi effetti sulla propria immagine professionale e per le implicazioni medico-legali e assicurative a cui esso può portare.

L'impostazione dei programmi di screening, che comporta per definizione il monitoraggio dell'intervento in tutti i suoi momenti, consente però di poter verificare la dimensione dello stesso e le sue caratteristiche e di poter quindi intervenire per migliorare gli aspetti organizzativi, tecnici e professionali per tenere più bassa possibile la percentuale di falsi negativi (quindi di CI), aumentando così la qualità del test e, di conseguenza, la sua sensibilità.

Tale caratteristica si deve connotare perciò, rovesciando l'atteggiamento comune critico nei confronti di questo dato, come un fatto decisamente vantaggioso per valutare la qualità del programma e del professionista che consente di ottimizzarne la qualità, di verificare costantemente le proprie prestazioni e, di conseguenza, di migliorarle (apprendimento dagli errori). Va quindi considerato come un valore aggiunto molto importante dello screening organizzato per veri-

ficare se esistono le premesse scientifiche e biologiche per arrivare a ottenere l'obiettivo di diminuire la mortalità organo-specifica nella popolazione interessata.

Da questo è scaturita l'esigenza da parte del GISMa di approfondire l'argomento per dar modo a tutti i professionisti e agli organizzatori dei programmi di poter disporre di un punto di riferimento scientifico-culturale che possa aiutare ad affrontare al meglio questo delicato problema nell'ottica del miglioramento continuo della qualità professionale e del programma stesso.

Descrizione del fenomeno

Lo screening mammografico, dimostratosi efficace nel ridurre significativamente la mortalità per carcinoma mammario (CM), è correntemente raccomandato come procedura di sanità pubblica e, in Italia, è in via di implementazione la copertura nazionale della popolazione femminile dai 50 ai 69 anni.¹

L'efficacia dello screening, peraltro, è ben lungi dall'essere assoluta, e la riduzione della mortalità è solo del 20-30% nelle diverse metanalisi.² Questo dipende da molti motivi, anzitutto dal fatto che i trial randomizzati confrontano i soggetti invitati allo screening (solo parte dei quali si sottopongono alla mammografia) con quelli non invitati (parte dei quali si sottopongono a mammografia su base volontaria). La riduzione di mortalità conseguente non all'invito, ma all'esecuzione della mammografia, come suggerito da diversi studi caso controllo, è verosimilmente maggiore, dell'ordine del 40%.

Un limite certo dello screening è comunque la sua sensibilità non ottimale. Lo screening, che in base al protocollo di riferimento (linee guida UE) utilizza la mammografia biennale, non diagnostica tutti i CM. Da un terzo a un quinto di essi, infatti, non è diagnosticato dallo screening ma compare, per lo più per la presenza di sintomi soggettivi, nell'intervallo tra uno screening negativo e la successiva mammografia di screening, programmata dopo

due anni.³⁻⁵ Questi carcinomi sono per questo definiti "di intervallo" (CI).

Sul piano biologico i CI non sembrano essere sostanzialmente diversi dai carcinomi che compaiono nella popolazione in assenza di screening. Una teoria, in realtà, vorrebbe che i CI fossero più aggressivi, con una velocità di crescita elevata, che farebbe sì che da piccole dimensioni, al di sotto della soglia della visibilità mammografica, crescano rapidamente fino a dare sintomi soggettivi, prima dello screening successivo. Questa teoria ("length bias") che postula per i CI un tempo di soggiorno e una fase di diagnosticabilità preclinica più breve della media, non sembra però essere confermata da diverse osservazioni: anzitutto i CI, per stadio e grado di differenziazione, non differiscono dalla media dei carcinomi clinici; inoltre la sopravvivenza dalla diagnosi, a parità di stadio, non è diversa da quella dei carcinomi clinici. Queste evidenze tenderebbero a negare l'ipotesi di una maggiore aggressività e velocità di crescita per i CI, che sarebbero CM "normali", semplicemente non riconosciuti dallo screening per vari motivi.

Premesso che, dati i tempi di crescita del CM, un CI che compaia entro due anni da uno screening negativo era certamente "presente" in quel momento, esso può non essere stato riconosciuto perché non era effettivamente riconoscibile (dimensioni troppo piccole, mascheramento da parte del seno radiologicamente denso) o perché, seppur riconoscibile, non lo è stato per errore (più o meno evitabile) del radiologo o perché, pur essendo stato riconosciuto come sospetto, non è stato confermato dagli accertamenti diagnostici di secondo livello. Queste diverse possibilità hanno evidentemente importanti implicazioni sia sul piano delle possibilità di intervento (per esempio, miglioramento dell'accuratezza della mammografia, addestramento dei radiologi, ottimizzazione degli accertamenti di secondo livello) sia su quello medico-legale.

Qualsiasi sia la motivazione della loro genesi, i CI

rappresentano in ogni caso un insuccesso del programma di screening che, in teoria, mira a diagnosticare tutti i carcinomi in fase di screening e a non consentire la comparsa di CI. Non stupisce, quindi, che l'analisi dei CI (frequenza di comparsa nel tempo, caratteristiche morfologiche e cliniche, revisione della precedente mammografia di screening) sia stata proposta come un momento fondamentale della valutazione di performance dello screening.⁶

Definizione

Per CI si intende un carcinoma successivo a un processo di screening negativo e comparso prima del passaggio di screening successivo. In accordo alle linee guida UE, la dizione di CI spetta sia alle forme invasive sia a quelle *in situ*, per cui appare opportuno che, ove possibile, siano registrati anche i carcinomi *in situ*, se non altro per una revisione retrospettiva dell'esame di screening precedente ai fini della valutazione dell'eventuale errore interpretativo.

Dal momento che i Registri tumori non includono i carcinomi *in situ* nel calcolo dell'incidenza, i medesimi saranno esclusi dall'indicatore sull'incidenza proporzionale (vedi dopo). Pur essendo i carcinomi *in situ* in parte lesioni non evolutive (sovradiagnosi) o a lenta evoluzione in invasive (quindi di probabile diagnosi precoce anche allo screening successivo), una minoranza di essi (forme poco differenziate) hanno un rischio elevato di evoluzione rapida in forme invasive altrettanto aggressive, e l'efficacia dello screening dipende anche dalla diagnosi di queste lesioni.

Nell'accezione più classica il CI ha una mammografia di screening negativa e viene diagnosticato entro due anni da questa. Esistono peraltro limitate estensioni del concetto che vanno definite perché devono essere considerate al fine di confrontare diverse casistiche.

- **CI oltre i due anni:** è possibile (in genere per problemi organizzativi o per scelta a priori, come nel Regno Unito) che l'intervallo successivo a un test di screening sia superiore a due anni. Si configura così la possibilità che esistano CI del terzo anno. Generalmente l'intervallo temporale di due anni è usato per i confronti fra casistiche. Ove l'intervallo triennale sia sistematico per scelta, è evidente che il tasso proporzionale dei CI nel

terzo anno va valutato quale indicatore della bontà della scelta dell'intervallo triennale stesso. Altrimenti i casi intervallo del terzo anno sono da misurare come indicatore del mancato mantenimento dell'intervallo di screening stabilito e da utilizzare per la revisione radiologica.

- **Errori di assessment:** può capitare che il test di screening sia positivo (cioè abbia "percepito correttamente" il carcinoma), che l'assessment non riesca a identificarlo correttamente e che il carcinoma sia poi diagnosticato nell'intervallo. Trattandosi di un caso sfuggito al "processo" di screening esso viene considerato un CI a tutti gli effetti anche se, nella revisione delle cause dei CI, esso sarà imputato a una carenza della fase di assessment e non a un possibile errore della fase di screening.

- **Early recall:** può capitare che un assessment non sia conclusivo e rimandi a un controllo ravvicinato (early recall) e che questo porti a una diagnosi effettivamente tardiva. Convenzionalmente questi casi, in accordo alle linee guida europee, sono da considerare screen-detected e non intervallo. Qualora la diagnosi avvenga (per sintomi o altro) prima dell'appuntamento per il controllo anticipato, il caso deve essere contato tra gli intervallo. Infine, i carcinomi insorti oltre la data dell'early recall in donne che non vi hanno aderito non sono considerati CI, anche se si raccomanda che vengano attentamente monitorati e siano soggetti a revisione radiologica.

- **Early recall ripetuti:** qualora l'early recall sia reiterato, con diagnosi oltre l'anno dalla mammografia di screening, rimane difficile assimilare un simile caso a uno screen-detected. Nella valutazione dei CI si può consigliare di valutare separatamente i casi di carcinoma diagnosticati a seguito di controlli ripetuti che abbiano sfiorato l'anno dalla data del primo esame di screening: in questi casi il processo diagnostico è stato effettivamente iniziato da una positività del test di screening, ma a seguito di una procedura diagnostica (reiterazione dell'early recall senza adire alla conferma istologica) sostanzialmente scorretta l'anticipazione diagnostica si è fortemente ridotta. Se non CI, andrebbero almeno misurati come indicatore di un iter diagnostico non adeguato.

- **Follow up non-attenders:** soggetti con test di screening positivo che si sottraggono all'appro-

fondimento consigliato presso il programma (o che lo sospendono prima della sua conclusione) e che vengono diagnosticati nell'intervallo perché o provvedono all'approfondimento fuori del programma, o diventano sintomatici. Nel primo caso una diagnosi comunque precoce di fatto è stata "innescata" della positività allo screening, mentre nel secondo la responsabilità diretta non è del programma di screening (anche se questi casi possono essere spunto di riflessione sulla adeguatezza della modalità con cui viene offerto l'approfondimento). In entrambe le evenienze questi casi dovrebbero essere assunti come screen-detected.

- **Lapsed attenders:** i carcinomi insorti oltre due anni dopo un test negativo in donne che non hanno aderito all'invito successivo non sono considerati CI, anche se si raccomanda che siano monitorati separatamente.

- **Carcinomi prevalenti:** si intendono carcinomi metacroni ipsi o controlaterali ove i soggetti già diagnosticati con carcinoma mammario permangono nelle liste di invito allo screening. Di fatto molti programmi escludono questi soggetti dallo screening, affidandoli ai curanti per le procedure di sorveglianza tipiche del follow up, i cui ritmi e modalità differiscono da quelli dello screening. Ai fini valutativi non appare quindi consigliabile l'inclusione di questi soggetti nella valutazione dei CI, non essendo essi formalmente eleggibili per lo screening stesso, ma per altro tipo di sorveglianza. Sono invece da includere nella revisione radiologica.

Le donne che escono dallo screening (per raggiunti limiti di età o per emigrazione) andrebbero seguite per un periodo di tempo pari all'intervallo di screening tra due round e registrati gli eventuali CI che compaiono.

In merito alle note sopra, si può fare riferimento anche al documento del programma inglese (NHSBCSP publication n. 62, aprile 2006, pag 7) dove si indica che i CI possono essere ulteriormente suddivisi in sottocategorie:

- "screen interval cancer" (CI successivi a un test/processo di screening negativo);
- "assessment interval cancer" (CI successivi a un test di screening positivo e a un approfondimento falso negativo);
- "interval cancers in follow up non attenders" (CI

successivi a un test di screening positivo con mancata presentazione all'approfondimento).

Rilevazione

La durata effettiva dell'intervallo viene misurata dalla data della mammografia negativa di screening e la data della diagnosi istologica (o, in mancanza di questa, quella della registrazione al Registro tumori).

La diagnosi citologica (C5) potrebbe essere presa in considerazione dato il suo elevatissimo valore predittivo positivo se non fosse per il fatto che un C5 può benissimo derivare da un carcinoma *in situ*, come tale da non considerare CI. Di fatto, i Registri assumono come riferimento la data della prima istologia di cancro invasivo. Assumere l'istologia come riferimento è più universale (i Registri spesso ignorano se e quando è stato formulato il primo sospetto diagnostico che ha portato alla diagnosi e ci sono casi di CI in cui l'unico dato è di fatto il dato di Registro).

Considerando i possibili ritardi tra diagnosi e biopsia/intervento, la regola sopra indicata può costare lo shift di una minoranza di casi dal primo a secondo e dal secondo al terzo anno.

I CI vengono diagnosticati per lo più perché divengono sintomatici e la paziente si presenta alla consultazione senologica e, in un numero limitato di casi, per la ripetizione anticipata della mammografia su base spontanea (tipicamente alla scadenza dell'anno dall'esame di screening). Quando la diagnosi avviene al di fuori del programma di screening, da parte del centro diagnostico o della paziente non si provvede sistematicamente alla segnalazione del CI al programma di screening. Questo comporta che non tutti i CI siano noti al centro di screening al momento della diagnosi.

Per poter rilevare tutti i CI è necessario cercarli attivamente in vario modo:

- innanzitutto vanno considerati i casi diagnosticati presso il centro di screening (che spesso opera anche come centro senologico per pazienti a presentazione spontanea); la proporzione di casi intervallo diagnosticati dal centro di screening stesso è un buon indicatore della capacità del programma di comunicare correttamente con le assistite e di mantenere un ruolo di riferimento anche a seguito di test negativo, oltre che del

fatto di poter disporre sollecitamente della possibilità di rivedere casi utili ai fini della formazione permanente;

- per l'identificazione dei CI non diagnosticati al centro di screening l'ideale è disporre di un Registro tumori, che è lo strumento deputato a rilevare l'incidenza, e tende a ignorare solo i casi (molto pochi, in genere) diagnosticati al di fuori del SSN; il linkage ai file del Registro tumori non deve limitarsi ai soli casi incidenti ma anche ai tumori in donne che avevano già avuto un cancro della mammella e ai carcinomi *in situ*; questi due ultimi casi dovrebbero essere registrati separatamente e non utilizzati nel calcolo dell'incidenza proporzionale, in quanto entrambi vengono esclusi dal denominatore (incidenza di riferimento, vedi oltre) e dunque non devono essere inclusi nemmeno al denominatore: i primi perché non incidenti per definizione; i secondi, carcinomi *in situ*, data la probabilità non assoluta di progressione e la possibilità di sovradignosi;

- nelle aree (ancora buona parte del Paese) non coperte da un Registro tumori, un surrogato può essere la creazione di un Registro di patologia (monitorando i servizi di anatomia patologica dell'area) in considerazione che è molto raro che un CM non abbia almeno una diagnosi istologica; un simile strumento, però, tende a ignorare i casi che sono diagnosticati e trattati al di fuori dell'area di riferimento;

- molto utile è la consultazione delle schede di dimissione ospedaliera (SDO) che coprono il territorio regionale e, in via differita, registrano anche i ricoveri dei propri residenti in altre Regioni; un simile strumento tende a ignorare solo i casi che siano stati trattati completamente al di fuori del SSN e/o non ospedalizzati, il che è poco probabile dato che il CM, nella stragrande maggioranza dei casi, viene trattato chirurgicamente in regime di ricovero.^{7,8}

La rilevazione quanto più completa dei CI secondo le modalità sopra esposte è compito del programma di screening, nella sua funzione deputata alla verifica di qualità della propria performance. Di tutte le misure comunemente messe in atto per monitorare la performance di un programma, la valutazione dei CI, probabilmente anche per le sue difficoltà, è stata finora la più negletta, nonostante sia uno degli indicatori più affidabili.

Indicatori

La frequenza di comparsa nel tempo dei CI è comunemente misurata come la frazione (incidenza proporzionale) di CI osservati rispetto ai CM "attesi" in assenza di screening (incidenza di base, "underlying incidence"). Questi ultimi sono noti direttamente solo in aree servite da un Registro tumori di popolazione o di patologia, ma stime di incidenza, ricavate in base alla mortalità e all'incidenza di Registri tumori geograficamente adiacenti, sono disponibili pressoché ovunque in Italia. I CM attesi sono calcolati attribuendo agli anni/persona per classi di età dei soggetti con screening negativo i tassi di incidenza età specifici.

Idealmente, uno screening perfetto non ha CI e quindi l'incidenza proporzionale è pari a zero. Viceversa, uno screening pessimo (o la sua assenza) prevede tanti CI quanti sono i CM attesi e una incidenza proporzionale del 100%. La realtà sta nel mezzo. Tanto più sensibile è lo screening, tanto meno CI compariranno, tanto più bassa sarà l'incidenza proporzionale. Il tasso di incidenza proporzionale costituisce il complemento a 1 (o al 100%) della sensibilità: per esempio, un tasso di incidenza proporzionale (osservati/attesi) del 30% (30 CI osservati/100 attesi) nel primo anno dell'intervallo corrisponde a una sensibilità (a 1 anno) del 70% (100 - 30).

Nelle linee guida UE viene suggerito per il calcolo della sensibilità il rapporto:⁶

$$\frac{\text{carcinomi diagnosticati dallo screening}}{\text{carcinomi diagnosticati dallo screening} + \text{carcinomi di intervallo}}$$

Tale formula può essere criticata: confronta infatti i carcinomi screen-detected (che sono quelli destinati a comparire in un periodo ben superiore ai due anni, data la possibilità che la mammografia abbia un'anticipazione diagnostica maggiore) con i CI, che per definizione compaiono nel biennio. Questo indicatore consente di includere nel calcolo i tumori *in situ*, sia al numeratore sia al denominatore (sia quelli che compaiono come intervallo sia quelli che compaiono come casi screen-detected). Si raccomanda di indicare sempre se si sono compresi gli *in situ*, dato che per queste lesioni l'anticipazione diagnostica è mediamente maggiore. Tale definizione della sensibi-

lità risulta inevitabilmente sovrastimata rispetto alla sensibilità vera e premia inevitabilmente la sensibilità di programmi al round di prevalenza. D'altro canto, anche l'incidenza proporzionale ha i suoi limiti, soprattutto per la possibile difficoltà di stimare l'incidenza di base in assenza di screening ("underlying"):

- per la mancanza di un Registro;
- per la disponibilità di dati di incidenza pre-screening di molti anni prima, con incidenza non rappresentativa di quella odierna;
- per l'effetto sull'incidenza di base dello screening spontaneo;
- per la limitata affidabilità di tassi di incidenza calcolati per aree più vaste o più limitate che non l'area del programma di screening, quindi limitatamente rappresentative;
- per la necessità di escludere dal calcolo gli *in situ*.

Se è quindi possibile usare entrambe le formule per il calcolo della sensibilità, è essenziale che confronti tra diversi scenari siano fatti in base allo stesso metodo di calcolo.

Nelle linee guida UE viene suggerito come indicatore anche il tasso di CI/10.000 donne sottoposte a screening. Tale valore, alla stregua del tasso diagnostico di carcinomi allo screening, è un indicatore imperfetto, in quanto fortemente condizionato dalla incidenza di base. Ove lo si volesse usare per confronti tra diversi scenari, dovrebbe essere aggiustato in base all'incidenza attesa. Tale indicatore può essere utilizzato senza necessità di aggiustamento per l'incidenza attesa per confronti interni a un programma di screening (tra centri o tra lettori).

Inoltre, per valutare l'impatto dei CI sull'efficacia dello screening, potrebbe essere impiegato come indicatore anche il tasso delle sole forme avanzate (T2+ o stadio II+), che non risulta influenzato dalla diagnosi comunque precoce di CI conseguente alla pressione diagnostica spontanea (per esempio, mammografie intermedie annuali spontanee).

L'incidenza proporzionale è d'abitudine misurata separatamente per il primo e per il secondo anno dell'intervallo o complessivamente, per il biennio. Le linee guida UE indicano come standard di riferimento per l'incidenza proporzionale dei CI un tasso massimo del 30% al primo anno,

del 50% al secondo anno, del 40% nel biennio. E' poi opportuno che tutti gli indicatori sui casi intervallo siano stratificati o aggiustati per età (classi quinquennali o decennali). Non è invece normalmente considerata importante, per i casi intervallo, la distinzione tra primo e successivi test di screening.

Stadio alla diagnosi

Le linee guida UE raccomandano la raccolta dell'informazione relativa allo stadio alla diagnosi dei CI, confrontato a quello dei casi screen-detected e dei casi insorti nelle non rispondenti.⁶ Il dato è effettivamente interessante, in quanto consente di stimare indirettamente l'esistenza di un "length bias" e l'impatto dei CI sull'efficacia dello screening ma, per essere completo, non può prescindere da una fonte di informazione dettagliata, quale un Registro tumori, purtroppo non sempre disponibile.

Revisione radiologica

Connessa all'identificazione dei CI è la revisione dei radiogrammi di screening precedenti, uno degli strumenti più efficaci per il miglioramento della qualità della lettura mammografica.

Il vantaggio di questa procedura è che è applicabile anche in caso di identificazione incompleta dei CI per carenza di fonti informative, o di mancata disponibilità dei radiogrammi di screening precedenti, in quanto queste condizioni in genere non configurano un bias di selezione quanto a tipologia di errore diagnostico. Necessita però dei radiogrammi sui quali è stato poi diagnosticato il CI, spesso non disponibili nel centro di screening. La revisione dei radiogrammi dovrebbe essere eseguita più correttamente con modalità cieca, cioè mescolando i radiogrammi precedenti i CI con controlli negativi, non riconoscibili come tali, in rapporto 1:4-1:5. In tal modo la revisione si avvicina di più alla condizione in cui potrebbe essere stato commesso l'errore diagnostico ed è certamente più corretta (e garantista nei confronti del radiologo) che non la revisione parzialmente (revisione dei soli CI) o completamente informata (revisione di CI con disponibilità dei radiogrammi diagnostici per confronto).

Queste due modalità (l'ultima in particolare) purtroppo sono le più seguite, anche in ambito di

perizia medico-legale, e tendono a sopravvalutare l'eventuale errore diagnostico per la facilità di riconoscere, con il "senno del poi", alterazioni minime alle quali di fatto nella pratica corrente non si dà importanza.

In ordine di affidabilità e preferenza la revisione cieca è certamente la migliore, seguita dalla parzialmente informata e dalla completamente informata. La prima, peraltro, è più complessa e potrebbe non essere facile da applicare per revisioni a livello regionale e nazionale per l'elevato numero di casi da rivedere. Resta il fatto che la modalità di revisione deve essere specificata e i confronti tra diversi programmi sono possibili solo quando sia stata impiegata la stessa modalità di revisione.

Alla revisione i radiogrammi devono essere classificati come negativi ("occuliti", "true interval", eventualmente radiologicamente occulti anche alla diagnosi, escludono una responsabilità del radiologo), o devono essere indicate con precisione sul radiogramma una o più lesioni focali, che in seguito un arbitro verifica coincidere o meno con il CI. Quando il revisore indica una lesione deve definire:

- se si tratta di lesione certamente meritevole di approfondimento ("errore di screening" o "falso negativo", implica una responsabilità del radiologo);
- se si tratta di alterazione minore ("minimal sign", non configura una responsabilità del radiologo), che non comporta necessariamente l'indicazione all'approfondimento diagnostico e che il revisore segnala con l'attenuante che, essendo comunque conscio di eseguire una revisione, è verosimile che la propria soglia del sospetto diagnostico risulti involontariamente abbassata.

I "minimal sign" sono soprattutto utili sul piano dell'apprendimento, in quanto consentono al radiologo di verificare come il carcinoma possa avere manifestazioni subdole e minimali e di affinare le proprie categorie diagnostiche di riferimento.

Dopo la revisione, un arbitro indipendente verifica quali delle lesioni segnalate corrispondano di fatto a CI e quali no, e definisce un valore di sensibilità (tasso dei CI riconosciuti come "errori di screening") e di specificità (tasso di controlli negativi classificati come "occuliti" o "minimal sign"). È evidente che il valore di sensibilità assume significati diversi se associato a una specificità alta o bassa.

Studi di riproducibilità delle revisione dimostrano che questa è in qualche misura soggettiva. Per tale motivo, anche se la procedura può risultare assai più complessa, potrebbe essere opportuno che la revisione fosse condotta non da uno ma da un numero dispari di revisori, con definizione delle controversie a maggioranza.

La procedura della revisione dei radiogrammi, oltre all'obiettivo del miglioramento continuo della qualità delle letture attraverso il confronto dei radiogrammi di screening e diagnostici, può avere in alcuni casi anche lo scopo di accertare se vi sia stata imperizia o meno da parte del radiologo nella lettura di screening. Appare evidente da quanto detto che per questo secondo scopo l'unica modalità di revisione che possa dare una risposta valida è quella in cieco.

Le linee guida UE indicano come standard di riferimento un tasso di CI classificati come "errore di screening" non superiore al 20%; relativamente alla revisione dei CI, si limitano a fornire le categorie diagnostiche per la revisione e lo standard di riferimento ma non entrano nei dettagli di come dovrebbe essere fatta la revisione (cieca, informata, singola, multipla con consenso o a maggioranza).⁶ Senza queste specifiche la variabilità dei risultati può essere tale che i confronti intercentro possono essere largamente inaffidabili.

Ove possibile la revisione, sempre con l'intento di identificare errori sistematici o più comuni di interpretazione (e quindi di innescare i relativi correttivi), potrebbe essere estesa alle mammografie negative seguite da carcinomi screen-detected al round successivo, almeno limitatamente alle forme diagnosticate in stadio avanzato.

Raccomandazione

Ogni programma di screening dovrebbe includere tra gli indicatori di performance da definire periodicamente:

- il tasso assoluto e proporzionale di CI, fornendo l'informazione relativa alle modalità adottate per la definizione sia degli attesi (incidenza di base), sia degli osservati (censimento dei CI); ove non disponibile un Registro tumori, l'identificazione può essere fatta tramite SDO, descrivendo le modalità adottate, in attesa che ne venga proposto a livello nazionale un protocollo di utilizzo omogeneo;

- il tasso di CI rivalutati come errori di screening, fornendo l'informazione relativa alle modalità adottate (tipo di revisione, cieca o informata, numero dei revisori, definizione delle controversie);
- la distribuzione per stadio alla diagnosi dei CI, confrontata con i carcinomi screen-detected e insorti nei non rispondenti.

In considerazione della complessità della definizione di tali indicatori, la loro determinazione può essere eseguita con periodicità meno frequente che annuale, anche se la revisione radiologica dei casi (non necessariamente svolta sulla totalità dei casi intervallo) dovrebbe essere un'attività routinaria a cadenza almeno annuale.

Il GISMa individua inoltre due bisogni conosciuti urgenti per l'implementazione su tutto il territorio nazionale di queste necessarie procedure di monitoraggio e miglioramento della qualità dei programmi di screening mammografico e raccomanda:

- che venga costituito un gruppo di lavoro che operi con l'Associazione Italiana Registri Tumori al fine di produrre stime di incidenza attesa per le diverse aree geografiche italiane, utilizzabili per il calcolo dell'incidenza proporzionale dei casi intervallo;
- che venga costituito un gruppo di lavoro per la definizione di metodi standardizzati per l'uso delle SDO nell'identificazione dei CI nelle zone non coperte da un Registro tumori o di patologia.

Bibliografia

1. Ministero della Sanità. Linee guida elaborate dalla Commissione oncologica nazionale, in applicazione di quanto previsto dal Piano Sanitario Nazionale per il triennio 1994-1996, relativo all'azione programmatica "Prevenzione e cura delle malattie oncologiche", concernenti l'organizzazione della prevenzione e dell'assistenza oncologica. Gazzetta Ufficiale 1996; 127: 25-38.
2. Kerlikowske K, Grady D, Rubin SM et al. Efficacy of screening mammography. A meta-analysis. *JAMA*. 1995; 273: 149-54.
3. Paci E, Ciatto S, Buiatti E et al. Early indicators of efficacy of breast screening programmes. Results of the Florence District Programme. *Int J Cancer* 1990; 46: 198-202.
4. Ciatto S, Rosselli Del Turco M, Zappa M. The detectability of breast cancer by screening mammography. *Br J Cancer* 1995; 71: 337-39.
5. Peeters PHM, Verbeek ALM, Hendricks JHCL et al. The occurrence of interval cancer in the Nijmegen screening programme. *Br J Cancer* 1989; 59: 929-32.
6. Perry NM, Broeders M, de Wolf C et al. *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition*. European commission, Luxembourg, 2006.
7. Vettorazzi M, Stocco C, Chirico A et al. Quality control of mammography screening in the Veneto Region. Evaluation of four programmes at a Local Health Unit level - Analysis of the frequency and diagnostic pattern of interval cancers. *Tumori* 2006; 92: 1-5.
8. Burani R, Caimi F, Maggioni C et al. Quality assessment of the mammographic screening programme in the Azienda Sanitaria locale Provincia Milano 1 - Analysis of interval cancers and discussion of possible causes of diagnostic error. *Radiol Med (Torino)* 2005; 109: 260-67.