Triage con intelligenza artificiale nello screening del tumore della mammella

Data di inserimento 25/03/2026

Anno 2025

Domanda di ricerca

Una strategia di triage basata sull'intelligenza artificiale per assegnare a singola o doppia lettura umana ciascuna mammografia invece che la doppia lettura umana per tutte le mammografie dovrebbe essere utilizzata nello screening mammografico basato su mammografia digitale (2DFFDM) o tomosintesi digitale del seno (DBT) per la diagnosi precoce del cancro al seno nei programmi di screening mammografico?

Quello che conta

La linea guida si basa su due studi condotti in contesti di screening organizzato di popolazione, entrambi di grandi dimensioni (oltre 200.000 donne in totale).
Studio MASAI (Svezia, Hernström 2025): trial randomizzato controllato in quattro centri di screening svedesi. Le donne sono state assegnate casualmente a ricevere screening con supporto dell'IA (*software Transpara*) oppure doppia lettura standard senza IA. Il sistema attribuiva a ciascuna mammografia uno score di rischio da 1 a 10: i casi con punteggio 1–9 venivano letti da un singolo radiologo supportato dall'IA, quelli con punteggio 10 ricevevano una doppia lettura radiologica.
Studio Lauritzen (Danimarca, 2024): studio osservazionale che ha confrontato le prestazioni del programma di screening della regione di Copenaghen prima e dopo l'introduzione dello stesso sistema di IA, su oltre 118.000 mammografie. I casi classificati dall'IA come probabilmente normali venivano assegnati alla lettura di un solo radiologo senior; gli altri ricevevano doppia lettura con supporto dell'IA.
I principali risultati sono sintetizzati nella tabella seguente:

Esito	Studio MASAI (Svezia, RCT)	Studio Lauritzen (Danimarca, pre/post)
Tumori identificati	+18% (da +4 a +25 casi ogni 10.000 donne)	+29% (da +3 a +24 casi ogni 10.000 donne)
Tasso di richiamo	+8% (non significativo)	–21% (significativo)
Falsi positivi	Equivalente alla doppia lettura	–32% (da –88 a –62 ogni 10.000 donne)
Letture risparmiate dai radiologi	–44% (9.129 letture ogni 10.000 donne)	–33,5% (6.700 letture ogni 10.000 donne)

Entrambi gli studi mostrano un aumento del tasso di identificazione dei tumori rispetto alla doppia lettura tradizionale. Il tasso di richiamo (donne richiamate per approfondimento) si riduce in uno studio e aumenta leggermente nell'altro, una differenza attribuibile anche ai diversi tassi di richiamo di partenza dei due programmi. In entrambi i casi l'aumento dei richiami, quando presente, è associato a un aumento parallelo della capacità di identificare tumori reali, per cui il valore predittivo positivo dei richiami migliora. Le letture aggiuntive evitate ai radiologi si collocano tra il 33% e il 44%.
Non sono ancora disponibili dati prospettici sui cosiddetti cancri intervallo, cioè i tumori che emergono tra un round di screening e il successivo. Dati retrospettivi su oltre un milione di mammografie indicano che i sistemi di IA identificano circa il 38% dei tumori non rilevati dalla doppia lettura tradizionale e il 93% dei tumori rilevati allo screening. Il panel ha giudicato improbabile un peggioramento di questo indicatore con l'introduzione del triage IA, pur sottolineando la necessità di monitorarlo prospetticamente.

Caveat	Entrambi gli studi hanno utilizzato lo stesso software (*Transpara, ScreenPoint Medical*): la generalizzabilità ad altri sistemi di IA è plausibile ma non dimostrata. Il panel ritiene che la raccomandazione possa applicarsi ad altri strumenti con accuratezza equivalente, previa validazione locale. L'analisi economica inclusa nella linea guida stima che il costo aggiuntivo di una lettura con IA (circa 2 euro per mammografia) sia inferiore al risparmio derivante dalla riduzione delle letture dei radiologi (circa 4,2 euro per lettura). L'intervento risulta cost-saving fino a un costo per lettura con IA di 3,83 euro, ma le stime dipendono dai volumi di attività e dalle infrastrutture già disponibili. Programmi di screening scarsamente informatizzati potrebbero affrontare costi iniziali più elevati. Il panel segnala il rischio di iniquità geografica in caso di implementazione disomogenea a livello nazionale, con centri che dispongono già di infrastrutture adeguate e altri che ne sono privi. È necessario un approccio graduale, con progetti pilota e monitoraggio sistematico, prima di una diffusione su larga scala. La raccomandazione è di tipo condizionale, non forte: ciò riflette la certezza moderata delle evidenze e la necessità di contestualizzare l'implementazione alle caratteristiche di ciascun programma. Non sostituisce la valutazione radiologica: il radiologo rimane parte integrante del processo diagnostico in tutti i casi.
Contesto	Il tumore della mammella è il più frequente nelle donne sia in Italia che nel resto d'Europa. In Italia si stimano circa 53.000 nuove diagnosi ogni anno (circa il 30% di tutti i tumori femminili). I programmi di screening mammografico organizzato, basati sulla mammografia biennale per le donne tra i 50 e i 69 anni, con estensione raccomandata fino ai 74 anni e con 'raccomandazione con riserva' per le fasce 45–49 anni, svolgono un ruolo centrale nella diagnosi precoce. Il modello attualmente raccomandato prevede la doppia lettura indipendente di ogni mammografia da parte di due radiologi, con arbitrato in caso di pareri discordanti. Questo approccio migliora la sensibilità diagnostica ma richiede un impegno significativo di personale specializzato, in un contesto in cui la carenza di radiologi formati nella lettura dei programmi di screening rappresenta un problema crescente in molti paesi, inclusa l'Italia. L'intelligenza artificiale applicata alla lettura delle mammografie è entrata nella pratica clinica in alcuni centri europei e sta per farlo in molti altri. La strategia di triage, che usa l'IA non per sostituire il radiologo, ma per decidere quanti radiologi devono leggere ogni mammografia, rappresenta un punto di equilibrio tra efficienza e supervisione umana, ed è quella per cui esistono le evidenze più solide al momento.
In evidenza	Un'analisi su oltre 1 milione di mammografie (Hovda et al., 2025) ha mostrato che il sistema di IA valutato è in grado di identificare il 38% dei tumori non rilevati dalla doppia lettura umana. Dati analoghi provengono da altri studi retrospettivi con diversi sistemi di IA, con una capacità di individuare tra il 20% e il 50% dei cancri intervallo a soglie di specificità pari o superiori al 90%. Dati preliminari presentati al Congresso Europeo di Radiologia 2025 dallo studio MASAI indicano una riduzione del 27% dei cancri intervallo non-luminali A nel gruppo sottoposto a screening con IA, suggerendo un potenziale impatto favorevole anche su questo esito critico, ancora in attesa di pubblicazione formale. Una survey condotta su 305 professionisti dei centri di screening italiani (Fattore et al., 2025) ha rilevato che il 75% è interessato al potenziale dell'IA e il 76% di quanti operano nello screening mammografico si dichiara disposto a sperimentare il modello di triage nel proprio centro. Le principali barriere identificate sono la mancanza di linee guida (ora parzialmente colmata), le risorse economiche e le infrastrutture informatiche.

Implicazioni per la pratica	L'introduzione di un sistema di triage basato sull'IA nello screening mammografico è raccomandabile, ma richiede attenzione a diversi aspetti organizzativi e regolativi: - Ogni centro deve validare localmente il software scelto sulle proprie immagini prima della messa a regime, verificando compatibilità tecnica e performance attese. - La soglia di score per il triage (che distingue le mammografie da inviare a singola o doppia lettura) deve essere definita in base alle caratteristiche del programma locale, non importata acriticamente dalla letteratura. - Il risparmio di tempo radiologo si realizza effettivamente solo se i flussi di lavoro sono riorganizzati di conseguenza e le risorse liberate vengono reindirizzate (ad esempio verso il secondo livello diagnostico). - L'implementazione deve avvenire nel rispetto del GDPR e del Regolamento europeo sull'IA (AI Act, 2024/1689), che classifica questi sistemi come ad alto rischio e prevede obblighi specifici di trasparenza, sorveglianza e responsabilità. - Le donne devono essere informate in modo chiaro e trasparente dell'uso dell'IA nel percorso diagnostico, con materiali elaborati con il coinvolgimento dei portatori di interesse. - È opportuno prevedere un monitoraggio prospettico degli indicatori di programma (tasso di identificazione, tasso di richiamo, cancri intervallo) separato per le unità che adottano il triage IA, per alimentare la base di evidenze e consentire aggiornamenti della raccomandazione.
Giudizio di qualità revisione	ALTO - linea guida di alta qualità metodologica, sviluppata con processo trasparente e multidisciplinare, con certezza delle evidenze moderata per gli esiti clinici chiave.
Riferimento bibliografico revisione	Sistema Nazionale Linee Guida – ISS. Screening e diagnosi del tumore della mammella (Adolopment LG Europee) – Capitolo sull'uso dell'Intelligenza Artificiale. Roma, 23 dicembre 2025.
Altri riferimenti	Hernström V et al. Screening performance and characteristics of breast cancer detected in the Mammography Screening with Artificial Intelligence trial (MASAI). Lancet Digit Health. 2025;7(3):e175-e183. Lauritzen AD et al. Early Indicators of the Impact of Using AI in Mammography Screening for Breast Cancer. Radiology. 2024;311(3):e232479. Hovda T et al. Retrospective evaluation of a CE-marked AI system, including 1,017,208 mammography screening examinations. Eur Radiol. 2025;35(9):5685-5694.
Parole chiave	- Intelligenza artificiale - screening - servizio sanitario/ospedale - patologie oncologiche - interventi di comunità -
Aree di intervento	- screening -
Setting	- comunitario -
Outcome	Di seguito i principali outcome clinici e organizzativi: - Tumori identificati (Cancer Detection Rate): misura la capacità del sistema (IA + radiologo) di individuare carcinomi mammari rispetto alla doppia lettura tradizionale. - Tasso di richiamo (Recall Rate): indica la percentuale di donne che, a seguito dello screening, vengono invitate a effettuare ulteriori accertamenti diagnostici. - Falsi positivi: si riferisce ai casi in cui lo screening sospetta un tumore che viene poi smentito dagli approfondimenti successivi. - Carico di lavoro (Letture evitate): valuta l'impatto sulla sostenibilità organizzativa misurando la riduzione del numero di letture richieste ai radiologi grazie al triage dell'IA. - Valore Predittivo Positivo (VPP) dei richiami: esprime la probabilità che una donna richiamata per approfondimenti sia effettivamente affetta da un tumore. - Cancri intervallo: tumori che vengono diagnosticati nel periodo che intercorre tra un test di screening negativo e l'appuntamento successivo. - Sostenibilità economica (Cost-saving): analisi del rapporto tra i costi di implementazione del software e il risparmio derivante dal minor tempo impiegato dal personale medico

Sintesi e traduzione	Sintesi a cura di Andrea Guida Azienda USL Toscana Centro. Revisione a cura di Cristiano Piccinelli - CPO, Centro Prevenzione Oncologica, Regione Piemonte