

# Riproducibilità inter-lettore del giudizio relativo alla densità radiologica secondo la classificazione quantitativa BI-RADS

Daniela Bernardi<sup>1</sup>, Marco Pellegrini<sup>1</sup>, Stefano Di Michele<sup>1</sup>, Paola Tuttobene<sup>1</sup>, Carmine Fantò<sup>1</sup>, Marvi Valentini<sup>1</sup>, Maria Adalgisa Gentilini<sup>2</sup>, Stefano Ciatto<sup>3</sup>

<sup>1</sup> U.O. Senologia Clinica e Screening Mammografico, Dipartimento di Radiodiagnostica, APSS, Trento Italy  
<sup>2</sup> Servizio Osservatorio Epidemiologico, Direzione per la Promozione e l'Educazione alla Salute, APSS, Trento Italy  
<sup>3</sup> Centro di Prevenzione Screening, ULSS 16, Padova Italy

## Obiettivi

valutare la **riproducibilità interosservatore** nella attribuzione della categoria di densità percentuale **BI-RADS** al fine di definire se la valutazione visuale sia affidabile nella pratica clinica o debba essere abbandonata in favore a metodi alternativi più affidabili quale quello computer-assistito.

**Tabella 1** Frequenza percentuale di attribuzione delle categorie di densità BI-RADS (P complessiva: 0,15)

Letttore	D1-2	D3-4	D3-4%	P (vs altri)
A	63	37	37	0,50
B	55	45	45	0,37
C	49	51	51	0,02
D	61	39	39	0,82
E	65	35	35	0,26
F	64	36	36	0,37

Sei lettori esperti di mammografia sono stati testati su un set di 100 mammografie. La **concordanza interosservatore** è stata valutata mediante la statistica kappa, che aggiusta per la concordanza casuale, rispetto a quattro (D1 vs. D2 vs. D3 vs. D4) o due categorie (D1-2 vs. D3-4). La **Tabella 1** mostra la **distribuzione percentuale delle categorie** di densità attribuite dai sei lettori. La distribuzione delle categorie varia tra i lettori ma non raggiunge la significatività statistica ( $p=0.15$ ).

**Tabella 2** Concordanza (Kappa pesato) dei sei radiologi tra loro nella attribuzione delle quattro categorie di densità (D1 vs D2 vs D3 vs D4) BI-RADS (set di 100 osservazioni): I limiti di confidenza al 95% sono indicati in parentesi

Letttore	A	B	C	D	E	F
A	-	0,72 (0,63-0,82)	0,62 (0,53-0,71)	0,69 (0,59-0,79)	0,67 (0,58-0,76)	0,73 (0,63-0,82)
B	0,72 (0,63-0,82)	-	0,67 (0,58-0,76)	0,73 (0,64-0,83)	0,65 (0,56-0,74)	0,71 (0,61-0,80)
C	0,62 (0,53-0,71)	0,67 (0,58-0,76)	-	0,61 (0,52-0,71)	0,61 (0,52-0,70)	0,63 (0,54-0,72)
D	0,69 (0,59-0,79)	0,73 (0,64-0,83)	0,61 (0,52-0,71)	-	0,72 (0,63-0,80)	0,87 (0,80-0,93)
E	0,67 (0,58-0,76)	0,65 (0,56-0,74)	0,61 (0,52-0,70)	0,72 (0,63-0,80)	-	0,80 (0,72-0,88)
F	0,73 (0,63-0,82)	0,71 (0,61-0,80)	0,63 (0,54-0,72)	0,87 (0,80-0,93)	0,80 (0,72-0,88)	-

La **Tabella 2** riporta i risultati della **valutazione di concordanza** (Kappa semplice e pesato) dei sei lettori nella attribuzione delle quattro categorie BI-RADS **rispetto al panel di riferimento**. La concordanza è moderata per 5 su 6 lettori e sostanziale per uno mentre, in base al Kappa pesato, la concordanza è sostanziale per 5 su 6 lettori, e quasi perfetta per uno.

**Tabella 3** Concordanza (Kappa semplice) dei sei radiologi rispetto al panel di riferimento nella attribuzione di due categorie di densità (D1-2 vs 3-4) BI-RADS (set di 50 osservazioni) con indicazione dei limiti di confidenza al 95%

Letttore	kappa semplice	Kappa semplice: intervallo di confidenza al 95%
A	0,72	0,52 - 0,91
B	0,76	0,58 - 0,94
C	0,72	0,52 - 0,91
D	0,76	0,58 - 0,94
E	0,76	0,59 - 0,94
F	0,72	0,54 - 0,91

Successivamente è stata verificata la **concordanza con un panel di altri 12 lettori** che, in uno studio precedente, avevano valutato lo stesso set di mammografie.

La **Tabella 3** riporta i risultati della valutazione di concordanza (Kappa semplice) dei sei lettori nella attribuzione di due categorie BI-RADS di densità (D1-2 vs D3-4) rispetto al panel di riferimento [21]. La concordanza risulta sostanziale per tutti i lettori.

## Risultati

I sei lettori hanno mostrato una **buona concordanza** quando valutati in tutte le possibili combinazioni di coppia. La concordanza su quattro categorie, sostanziale (kappa = 0,60-0,80) per 13 coppie, è stata quasi perfetta (kappa >0,80) per due coppie. La concordanza su due categorie è stata sostanziale per 12 coppie e quasi perfetta per tre. Rispetto al panel di 12 radiologi la concordanza su quattro categorie è stata sostanziale per 5 su 6 lettori e quasi perfetta per uno, mentre quella su due categorie è stata sostanziale per tutti i lettori.

## Conclusioni

Il lavoro, in accordo con studi precedenti, conferma come la **classificazione visuale della densità** secondo BI-RADS sia **altamente riproducibile**; l'attribuzione alla categoria "denso" (BI-RADS 3-4) varia tuttavia discretamente tra i lettori (range 36-51%) (v. tabella 1). Di qui la necessità di studi controllati che confrontino la classificazione visuale con **quella computerizzata** potendo quest'ultima rappresentare una **valida alternativa** alla prima in virtù della sua riproducibilità pressoché assoluta.