



European Commission Initiative on Breast Cancer (ECIBC): European guidelines on breast cancer screening and diagnosis Evidence profile

Healthcare question	Should an optimal number of readings vs. no specific number be used for allowing mammography readers to work in mammography screening programmes?
Date	October 2018
Authors	Guidelines Development Group (GDG): Mariangela Autelitano, Bettina Borisch, Mireille Broeders, Xavier Castells, Edoardo Colzani, Jan Daneš, Stephen Duffy, Patricia Fitzpatrick, Markus Follmann, Livia Giordano, Paolo Giorgi Rossi, Axel Gräwingholt, Solveig Hofvind, Lydia Ioannidou-Mouzaka, Susan Knox, Miranda Langendam, Annette Lebeau, Helen McGarrigle, Lennarth Nyström, Elsa Pérez Gómez, Cecily Quinn, Holger Schünemann, Alberto Torresin, Ruben Van Engen, Cary Van Landsveld-Verhoeven, Sue Warman, Kenneth Young. Systematic Review team: David Rigau Comas, Pablo Alonso Coello, Iván Sola Arnau, Margarita Cynthia Posso Rivera. JRC Healthcare Quality team: Marta Garcia-Escribano, Sazan Pakalin, Zuleika Saz-Parkinson, Asli Uluturk
Abbreviations	CI: Confidence interval

N° of studies	Study design	Certainty assessment					N° of patients		Effect		Certainty	Importance
		Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	An optimal number of readings	No specific number	Relative (95% CI)	Absolute (95% CI)		
False positive rate (assessed with: defined as the proportion of positive screening examinations among all screening exams without a breast cancer diagnosis within the follow-up period.)												
7 1,2,3,4,5,6,7,ab	observational studies	serious ^c	serious ^d	not serious	serious ^e	none	The increasing number of readings is related with an overall decreasing of false positive rate considering all data from the 7 studies together. (linear R ² =0.4951; p<0.001). A cut-off ≤ 3500 annual				⊕○○○ VERY LOW	CRITICAL

Certainty assessment							N° of patients		Effect		Certainty	Importance
N° of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	An optimal number of readings	No specific number	Relative (95% CI)	Absolute (95% CI)		
							readings is associated with a statistically significant reduction of an absolute 1.35% false positives every 1000 readings increase of experience (p=0.01) whereas a cut-off ≥ 3500 annual readings is associated with a non-significant decreasing of an absolute 0.13% false positives every 1000 readings increase of experience (p=0.126)					
Sensitivity (assessed with: Proportion of screening examinations interpreted as positive (BIRADS 0, 4, 5 or referred for assessment) among all screening exams who were given a diagnosis of BC within the follow-up period)												
6 2,3,4,6,7,8,ab	observational studies	serious ^c	serious ^f	not serious	serious ^g	none	The visual inspection of the scatter plot reveals that the increasing number of readings is related with an overall increasing of sensitivity up to approximately 7000 annual readings (between 0.893 and 0.939). Over this number of readings, it appears to be a decrease of sensitivity.		⊕○○○ VERY LOW			CRITICAL
Breast cancer detection (assessed with: defined as the number of cancers detected per 1000 screening exams. Cancer included all breast cancer (ductal carcinoma in situ or invasive))												
4 2,6,7,8,ab	observational studies	serious ^c	serious ^f	not serious	serious ^g	none	The visual inspection of the scatter plot reveals that the increasing number of readings is related with an overall increasing of breast cancer detection up to approximately 7000 annual readings (between 5.05 and 8.3 per 1000 screening exams). Over this number of readings it appears to be a decrease of the breast cancer detection.		⊕○○○ VERY LOW			CRITICAL
Recall rate - not measured												
-	-	-	-	-	-	-	-	-	-	-	-	-

Explanations

- The research evidence considered refers to organised screening programs.
- Hoff et al. 2019 has used data from the Cancer Registry of Norway. The interpretation and reporting of these data are the sole responsibility of the authors, and no endorsement by the Cancer Registry of Norway is intended nor should be inferred.
- All studies were retrospective and these outcomes were not adjusted for relevant confounding variables. Number of readings was mostly self-reported or not accounting for professionals' experience in other settings.

- d. Two out of 7 studies did not show a statistically significant relationship (trend) between false positive rate and number of readings according to their analysis. According to our slope analysis, 4 out of 7 studies do not show a statistically significant relationship (trend).
- e. The available data is non-weighted and with no estimation of the 95%CI. The overall slope for false positives and number of readings is statistically significant. A cut-off was selected when the slope above this point was no longer significant (flat slope or relationship).
- f. According to visual inspection of the scatter plot, only 2 studies were compatible with the estimation of a maximum number of readings (Cornford 2011 and Hoff 2019).
- g. The summary of findings represent a visual interpretation of the scatter plot and no statistical inference was taken. The available data is non-weighted and with no estimation of the 95%CI.

References

1. Alberdi RZ, Llanes AB Ortega RA Expósito RR Collado JM Verdes TQ Ramos CN Sanz ME Trejo DS Oliveres XC, group., CFPR (Cumulative False Positive Risk). Effect of radiologist experience on the risk of false-positive results in breast cancer screening programs. *Eur Radiol*; 2011.
2. Hoff SR, Myklebust TÅ, Lee CI, Hofvind S.. Influence of Mammography Volume on Radiologists' Performance: Results from BreastScreen Norway.. *Radiology*; 2019.
3. Elmore JG, Jackson SL, Abraham L, Miglioretti DL, Carney PA, Geller BM, Yankaskas BC, Kerlikowske K, Onega T, Rosenberg RD, Sickles EA, Buist DS.. Variability in interpretive performance at screening mammography and radiologists' characteristics associated with accuracy.. *Radiology*; 2009.
4. Duncan KA, Scott NW. Is film-reading performance related to the number of films read? The Scottish experience. *Clinical Radiology*; 2011.
5. Théberge I, Hébert-Croteau N, Langlois A, Major D, Brisson J. Volume of screening mammography and performance in the Quebec population-based Breast Cancer Screening Program. *CMAJ*; 2005.
6. Théberge I, Chang SL, Vandal N, Daigle JM, Guertin MH, Pelletier E, Brisson J.. Radiologist interpretive volume and breast cancer screening accuracy in a Canadian organized screening program. *J Natl Cancer Inst*; 2014.
7. Buist DS, Anderson ML, Haneuse SJ, Sickles EA, Smith RA, Carney PA, Taplin SH, Rosenberg RD, Geller BM, Onega TL, Monsees BS, Bassett LW, Yankaskas BC, Elmore JG, Kerlikowske K, Miglioretti DL.. Influence of annual interpretive volume on screening mammography performance in the United States. *Radiology*; 2011.
8. Cornford, E. J., Turnbull, A. E., James, J. J., Tsang, R., Akram, T., Burrell, H. C., Hamilton, L. J., Tennant, S. L., Bagnall, M. J., Puri, S., Ball, G. R., Chen, Y., Jones, V.. Accuracy of GE digital breast tomosynthesis vs supplementary mammographic views for diagnosis of screen-detected soft-tissue breast lesions. *Br J Radiol*; 2016.