Computer-Aided Detection (CAD) in Mammography Screening

SBU ALERT REPORT NO 2011-05 • 2011-05-18 • WWW.SBU.SE/ALERT

Summary and conclusions

In Sweden, all women aged 50 through 69 years are offered mammography at regular intervals. Most county councils even offer this type of examination to women aged 40 through 49 years and 70 through 74 years. Swedish and European guidelines recommend that two specially trained radiologists (breast radiologists) review the breast images. Computer-aided detection (CAD) is a computerised method for analysing images from e.g. mammography screening. Although the method has existed for approximately 10 years, its clinical use in Sweden is limited.

SBU's appraisal of the evidence

- ❑ The scientific evidence is insufficient to determine whether CAD plus single reading by one breast radiologist would yield results that are at least equivalent to those obtained in standard practice, i.e. double reading where two breast radiologists independently read the x-ray images.
- ❑ Since the medical consequences are uncertain, it is not possible to determine the cost-effectiveness or the socioeconomic consequences of replacing one of the readings with CAD in the context of mammography screening.
- ❑ Since this literature review, CAD technology has advanced further by virtue of improved features in computer software and digitalisation of images. Additional studies are essential to understand the specific benefits, risks, and costs of the method.

Technology and target group

The method under investigation is computer-aided detection (CAD) in breast cancer screening. The target group includes women aged 40 through 74 years who receive mammography within the framework of population-based screening.

The program used in CAD identifies and marks areas that the software identifies as abnormal breast tissue. The CAD program is not intended as the only method to be used when analysing mammography images. Rather, it is designed to alert the radiologist about possibly suspicious areas. Hence, interpretation of the image by a breast radiologist must accompany CAD.

It has been suggested that CAD in conjunction with mammography screening could replace one of the two independent readings that are done in accordance with European and Swedish guidelines. A prerequisite would be that the diagnostic accuracy and patient benefit are as good when the images are read by one breast radiologist plus CAD as when they are read by two breast radiologists. Another important prerequisite is that not too many women need to be called back for further diagnostic work-up (recall). In Europe, the highest recommended recall rate is 5 percent.

High average age among practicing breast radiologists, and poor replacement rate in this group of specialists, has increased the interest for computerised analysis of mammography images.

In this report we have studied whether diagnostic accuracy is at least as good, while recall rates are not higher, when CAD plus single reading by one breast radiologist is used in conjunction with mammography screening instead of independent readings by two breast radiologists.

Primary questions

Is the reading of mammographic images by a single breast radiologist plus CAD at least as accurate as readings by two breast radiologists (current practice) in terms of:

- sensitivity (probability that a person with the disease has a positive test result)?
- specificity (probability that a healthy person has a negative test result)?
- cancer detection rate (number of cancer cases detected per 1 000 women examined)?
- recall rate (women called back for further investigation)?
- cost-effectiveness?

Inclusion criteria

The report includes population-based screening studies only. The studies should include at least 5 000 women and the study settings should be comparable to Swedish conditions. Furthermore, the studies should compare mammography readings by one breast radiologist plus CAD against readings by two breast radiologists. Since prospective studies based on digital mammography could not be identified, scanned analogue images were accepted.

Patient benefit

Recall increases short-term anxiety among the women affected and also increases cost. Therefore, the requirement for using single reading plus CAD is that the method must detect at least as many cancers as double reading, without increasing the recall rate, i.e. the method's specificity must be at least as high as that in double reading.

Only one study of sufficient quality met the inclusion criteria. It compared single reading plus CAD with double reading in conjunction with mammography screening, and reported no difference in the percentage of cancer cases detected. The recall rate, however, was statistically significantly higher for single reading plus CAD (3.9 percent compared to 3.4 percent for double reading). The generalisability of the study is reduced since all of the breast radiologists participating in the study had extensive experience in mammography screening. Therefore, this single study, having deficiencies in study quality and generalisability, cannot be used to draw conclusions (insufficient scientific evidence ⊕○○○).

SBU's assessment shows that the scientific evidence is insufficient to comment on single mammographic reading by one breast radiologist plus CAD in comparison to current practice of double reading involving two breast radiologists.

Economic aspects

Since the medical consequences are uncertain, it is not possible to determine the cost-effectiveness and/or the socioeconomic consequences of replacing one of the readings with CAD in the context of mammography screening.

Four levels are used in grading the strength of the scientific evidence on which conclusions are based:

Strong scientific evidence (⊕⊕⊕⊕). Based on high or medium quality studies with no factors that weaken the overall assessment.

Moderately strong scientific evidence (⊕⊕⊕○). Based on high or medium quality studies with isolated factors that weaken the overall assessment.

Limited scientific evidence (⊕⊕○○). Based on high or medium quality studies having factors that weaken the overall assessment.

Insufficient scientific evidence (⊕○○○). Scientific evidence is deemed insufficient when scientific findings are absent, the quality of available studies is low, or studies of similar quality present conflicting findings.

Project group

- **Edward Azavedo**, MD, PhD, Associate Professor, Karolinska University Hospital, Stockholm
- **Sophia Zackrisson**, MD, PhD, Skåne University Hospital, Malmö
- **Marianne Heibert Arnlind**, Project Manager, Health Economist, SBU, heibert.arnlind@sbu.se
- **Ingegerd Mejàre**, Professor Emerita, Assistant Project Manager, SBU
- SBU Staff: Lars-Åke Marké, Health Economist, Karin Rydin, Literature Searcher, Lena Wallgren, Project Assistant

Scientific reviewers

- **Olof Jarlman**, MD, PhD, Associate Professor, Sophiakliniken, Lund
- **Per Skaane**, MD, PhD, Professor, Oslo University Hospital, Oslo

References

1. Socialstyrelsen. www.socialstyrelsen.se
2. Sveriges Kommuner och Landsting, Socialstyrelsen. Öppna jämförelser av hälso- och sjukvårdens kvalitet och effektivitet: jämförelser mellan landsting 2009. Stockholm: Sveriges Kommuner och Landsting; 2009.
3. Nyström L, Andersson I, Bjurstam N, Frisell J, Nordenskjöld B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. Lancet 2002;359(9310):909-19.
4. 1177. Mammografi. www.1177.se
5. Olsson S, Andersson I, Bjurstam N, Frodis E, Håkansson S, Lithander E, et al. [600,000 women are examined by mammography per year. Every fifth of them refuses screening]. Läkartidningen 1995;92(6):552-6.
6. Swedish Organised Service Screening Evaluation Group. Reduction in breast cancer mortality from organized service screening with mammography: 1. Further confirmation with extended data. Cancer Epidemiol Biomarkers Prev 2006; 15(1):45-51.
7. Vainio H, Bianchini F, editors. IARC Handbooks of Cancer Prevention – Breast Cancer Screening. Lyon: IARCPress; 2002.
8. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition – summary document. Ann Oncol 2008;19(4):614-22.
9. Helvie M. Improving mammographic interpretation: double reading and computer-aided diagnosis. Radiol Clin North Am 2007;45(5):801-11, vi.
10. Azavedo E, Svane G. Radiologic aspects of breast cancers detected through a breast cancer screening program. Eur J Radiol 1991;13(2):88-90.
11. Azavedo E, Svane G. Radial scars detected mammographically in a breast cancer screening programme. Eur J Radiol 1992; 15(1):18-21.
12. Brett J, Bankhead C, Henderson B, Watson E, Austoker J. The psychological impact of mammographic screening. A systematic review. Psychooncology 2005;14(11):917-38.
13. Brett J, Austoker J. Women who are recalled for further investigation for breast screening: psychological consequences 3 years after recall and factors affecting re-attendance. J Public Health Med 2001;23(4):292-300.
14. Brewer NT, Salz T, Lillie SE. Systematic review: the long-term effects of false-positive mammograms. Ann Intern Med 2007;146(7):502-10.
15. Brodersen J, Thorsen H, Cockburn J. The adequacy of measurement of short and long-term consequences of false-positive screening mammography. J Med Screen 2004;11(1):39-44.
16. Lidbrink E, Elfving J, Frisell J, Jonsson E. Neglected aspects of false positive findings of mammography in breast cancer screening: analysis of false positive cases from the Stockholm trial. BMJ 1996;312(7026): 273-6.
17. Boudin F, Nie JY, Bartlett JC, Grad R, Pluye P, Dawes M. Combining classifiers for robust PICO element detection. BMC Med Inform Decis Mak 2010;10:29.
18. Khoo LA, Taylor P, Given-Wilson RM. Computer-aided detection in the United Kingdom National Breast Screening Programme: prospective study. Radiology 2005;237(2):444-9.
19. Houssami N, Given-Wilson R, Ciatto S. Early detection of breast cancer: overview of the evidence on computer-aided detection in mammography screening. J Med Imaging Radiat Oncol 2009;53(2):171-6.
20. Skaane P, Kshirsagar A, Stapleton S, Young K, Castellino RA. Effect of computer-aided detection on independent double reading of paired screen-film and full-field digital screening mammograms. AJR Am J Roentgenol 2007;188(2):377-84.
21. Gilbert FJ, Astley SM, Gillan MG, Agbaje OF, Wallis MG, James J, et al. Single reading with computer-aided detection for screening mammography. N Engl J Med 2008;359(16):1675-84.
22. Georgian-Smith D, Moore RH, Halpern E, Yeh ED, Rafferty EA, D'Alessandro HA, et al. Blinded comparison of computer-aided detection with human second reading in screening mammography. AJR Am J Roentgenol 2007;189(5):1135-41.
23. Gromet M. Comparison of computer-aided detection to double reading of screening mammograms: review of 231,221 mammograms. AJR Am J Roentgenol 2008;190(4):854-9.
24. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008;336(7650):924-6.
25. Taylor P, Champness J, Given-Wilson R, Johnston K, Potts H. Impact of computer-aided detection prompts on the sensitivity and specificity of screening mammography. Health Technol Assess 2005;9(6):iii, 1-58.
26. Hellquist BN, Duffy SW, Abdsaleh S, Björneld L, Bordás P, Tabár L, et al. Effectiveness of population-based service screening with mammography for women ages 40 to 49 years: evaluation of the Swedish Mammography Screening in Young Women (SCRY) cohort. Cancer 2011;117(4):714-22.
27. Olsson S, Andersson I, Karlberg I, Bjurstam N, Frodis E, Håkansson S. Implementation of service screening with mammography in Sweden: from pilot study to nationwide programme. J Med Screen 2000;7(1):14-8.
28. Moberg K, Bjurstam N, Wilczek B, Rostgård L, Egge E, Muren C. Computed assisted detection of interval breast cancers. Eur J Radiol 2001;39(2):104-10.
29. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol 2003;3:25.

SBU evaluates healthcare technology
The Swedish Council on Health Technology Assessment (SBU) is a national governmental agency that assesses healthcare technologies. SBU analyses the benefits, risks, and costs of different methods and compares the scientific facts to prevailing practices in Sweden. SBU's goal is to provide stronger evidence for everyone engaged in shaping the delivery of health services.

The SBU Alert reports are produced in collaboration with experts from the respective subject areas, the National Board of Health and Welfare, the Medical Products Agency, the Swedish Association of Local Authorities and Regions, and a special advisory panel (the Alert Advisory Board).

This assessment was published in 2011. Findings based on strong scientific evidence usually continue to apply well into the future. However, findings based on insufficient, limited, or contradictory evidence might have already been replaced by more recent findings.

The complete report is available in Swedish.

Publisher: Måns Rosén, Director, SBU
Program Manager: Sofia Tranæus, SBU
Graphic Production: Anna Edling, SBU