Irish Journal of Medical Science (1971 -) (2019) 188:289–293
https://doi.org/10.1007/s11845-018-1828-8

ORIGINAL ARTICLE

Breast cancers missed by screening radiologists can be detected by reading mammograms at a distance

Ineke L. Schreutelkamp[1] · Robert M. Kwee[2] · Peter Veekmans[3] · Miraude E. A. P. M. Adriaensen[2]

Received: 23 November 2017 / Accepted: 25 April 2018 / Published online: 3 May 2018

Abstract

Background During locally organized quality assurance evaluation sessions for screening radiologists, we noticed that individual screening radiologists did miss tumours which in our opinion could be detected at a distance.

Aim To determine whether tumours missed by individual screening radiologists can be detected at a distance.

Methods Twenty-eight screening mammograms of 28 females (mean age 63 years, range 49–73) with a pathologically proven malignant tumour missed by individual screening radiologists were mixed with 56 normal screening mammograms of 56 females (mean age 63 years, range 53–74). This test set was independently assessed by a senior screening radiologist and by a radiology resident without prior training in screening mammography at 1.5 m distance from the screen display. Readers were unaware of the prevalence of pathologically proven malignant tumours in the test set. Primary outcome was whether the reader would recall the woman.

Results The senior screening radiologist recalled 28 of 28 women with a pathologically proven malignant tumour (sensitivity of 100%) and 16 of 56 women without pathology (specificity of 71%). The radiology resident recalled 25 of 28 women with a pathologically proven malignant tumour (sensitivity of 89%) and 10 of 56 women without pathology (specificity of 82%).

Conclusion Some malignant tumours missed by an individual screening radiologist can be detected from 1.5 m distance. Therefore, we recommend that screening radiologists consciously take a distant view before closely evaluating the mammogram in detail.

Keywords Breast cancer · Detection · Mammography · Screening · Test set

Introduction

Breast cancer is the top cancer in women worldwide, comprising 25.2% of all female cancers [1]. It is estimated that 521,817 women died in 2012 due to breast cancer [1].

Screening programmes have been developed to detect breast cancer at an early stage, in order to improve patient outcome and survival. All medical professional organizations in industrialized countries recommend screening mammography for women between 50 and 69 years of age [2]. The dutch breast cancer population-based screening programme, targeted at all women aged 50–74 years, provides biennial screening mammography [3]. All screening radiologists participating in the dutch breast cancer population-based screening programme must be registered at the National Expert and Training Centre for Breast Cancer Screening (NETCB) quality registry. In order to qualify, radiologists must successfully pass an initial screening training at a national level by the

Research was presented at the ECR 2016, Vienna, Austria, March 2016

✉ Miraude E. A. P. M. Adriaensen
miraude@gmail.com

Ineke L. Schreutelkamp
j.schreutelkamp@ziggo.nl

Robert M. Kwee
rmkwee@gmail.com

Peter Veekmans
peterveekmans@skynet.be

[1] Bevolkingsonderzoek Zuid, Australiëlaan 12, 6199 AA Maastricht-Airport, The Netherlands

[2] Department of Radiology, Zuyderland Medical Center, Henri Dunantstraat 5, 6419 PC Heerlen, The Netherlands

[3] Department of Medical Imaging, SJG Weert, Postbox 29, 6000 AA Weert, The Netherlands

NETCB, participate in Continuing Medical Education (CME), and read a minimum number of 3000 screening mammograms a year. Despite adequate training, the rate of missed breast cancers in screening is still estimated at 25% (range 13–41%) [4, 5]. During locally organized quality assurance evaluation sessions for screening radiologists participating in a Southern region of the dutch breast cancer population-based screening programme, we noticed that individual screening radiologists did miss tumours which in our opinion could be detected at a distance. Therefore, the objective of our study was to determine whether tumours missed by individual screening radiologists can indeed be detected at a distance.

Methods

Screening mammography

All women had undergone digital mammography in a Southern region of the dutch breast cancer population-based screening programme. Women were invited by mail and screening mammograms were performed in out-of-hospital mobile units with a Lorad Selenia FFDM system with a 70-μm pixel size and a 232 by 286 mm field of view (Hologic Inc., Danbury, USA). Screening mammograms were performed by dedicated radiographers trained at a national level by the NETCB. According to national guidelines, screening mammograms were performed in two views (craniocaudal (CC) and mediolateral oblique (MLO)) per breast. In accordance with the national Data Protection Act, women were informed at screening that their data could be used anonymously for scientific research. Women can opt out, if they do not agree. Since the test set only consisted of screening mammograms of women who agreed to opt in, no institutional review board approval was required. In daily practice, all screening mammograms are independently read by two individual screening radiologists in an out-of-hospital standardized quiet reading room. In cases with discrepant ratings, a final decision is made by a third screening radiologist who is aware of the scores given by the other two screening radiologists.

Test set

A senior screening radiologist with more than 25 years of experience (JS) created a test set of 84 screening mammograms. The test set consisted of 28 screening mammograms of 28 females (mean age 63 years, range 49–73) with pathologically proven malignancies which were missed by one but detected by one other screening radiologist and 56 normal screening mammograms of 56 females (mean age 63 years, range 53–74). Note that the "missed tumours" were missed by different radiologists with experience in reading screening mammograms ranging from 5 years to more than 25 years. The target recall rate at the national level is set at less than 2%. The average individual recall rate of the different radiologists who missed the tumours was 2.1% (ranging from 1.3 to 3.4%). The 28 screening mammograms with pathologically proven malignancies were chosen by the senior screening radiologist (JS) from the lists reporting the individual performances of all screening radiologists in our region. These lists are also used for locally organized quality assurance evaluation. According to the senior screening radiologist (JS), all malignancies were visible on the 28 mammograms. Interval cancers were not included in our study. Thirteen pathologically proven malignant tumours were located in the right breast and 15 pathologically proven malignant tumours were located in the left breast. For all included negative cases, a subsequent negative screening examination (i.e. final BI-RADS [6] score 1 or 2) was available. The screening mammograms with pathologically proven malignancies and the normal screening mammograms were randomly mixed. No prior mammograms were available for comparison in this test set.

Test set evaluation

The test set was only read from 1.5 m distance under otherwise normal viewing room circumstances on a dedicated diagnostic work station (Hologic SecurView DX; Hologic Inc., Danbury, USA) by a senior screening radiologist (PV) with 10 years of experience as a screening radiologist and awarded the European Diploma in Breast Imaging in 2013, and by a radiology resident (MW) who received no prior training in screening mammography. The original screening radiologists who missed the malignancy and the senior screening radiologist (PV) were not matched for experience. Mammograms were presented at the diagnostic working station according to our standard hanging protocol. That is to say first a hanging overview with four empty slots of the prior images of the previous screening round at the top and the four images (from left to right: right MLO view, left MLO view, right CC view, left CC view) of the mammogram included in the test set at the bottom, followed by a hanging with the two MLO views on the left and the two CC views on the right (from left to right: right MLO, left MLO, right CC, left CC), followed by the hanging of the two MLO views (from left to right: right MLO view, left MLO view) and followed by the hanging of the two CC views (from left to right: right CC-view, left CC-view). Readers were allowed to go back and forward between these hangings as often as they liked. No other tools were allowed. Both readers independently completed the test set at their own speed. Readers were unaware of the number of screening mammograms with pathologically proven malignancies in the test set. Before finishing analyzing the test set, the readers were also unaware of the total number of mammograms included. The readers indicated whether they would

recall the woman on the basis of the screening mammogram. In case of a recall, the readers indicated laterality (left breast, right breast or both breasts) and were asked to mention the quadrant in which they noticed the abnormality if possible (i.e. right upper lateral, right lower lateral, right upper medial, right lower medial, left upper lateral, left lower lateral, left upper medial, left lower medial, one projection only MLO cranial, one projection only MLO caudal, one projection only CC lateral, one projection only CC medial). Furthermore, the readers assigned a lesion type (well-defined mass, ill-defined mass, spiculated mass, architectural distortion, focal asymmetry, cluster of microcalcifications) to the most suspicious lesion. In addition, the readers assigned a mammographic breast density category to all mammograms in the test set according to the description in the American College of Radiology (ACR) Breast Imaging Reporting and Data System Atlas (BI-RADS) 5th Edition [7].

Statistical analysis

Sensitivity was determined as the ratio of the number of correctly identified true positive cases divided by the total number of true positive cases. Specificity was determined as the ratio of the number of correctly identified true negative cases divided by the total number of true negative cases. The inter-observer variability with regard to recall was determined by calculating Cohen's kappa (κ). The Landis and Koch classification of the agreement based on κ value was used to interpret the κ values [8].

Results

Mammographic breast density categories assigned by the senior screening radiologist and radiology resident are displayed in Table 1.

From 1.5 m distance from the screen display, the senior screening radiologist recalled 28 of 28 women with a pathologically proven tumour (sensitivity of 100%) and 16 of 56 women without pathology (specificity of 71%). From 1.5 m distance from the screen display, the radiology resident recalled 25 of 28 women with a pathologically proven tumour (sensitivity of 89%) and 10 of 56 women without pathology (specificity of 82%). Lesion types assigned by the senior screening radiologist and radiology resident are displayed in Table 2.

Agreement on recall between the senior screening radiologist and the radiology resident was moderate ($\kappa = 0.50$; 95% CI 0.32–0.69). In the group of 25 pathologically proven tumours recalled by both readers, the readers agreed on the breast quadrant assigned in 23 cases. In the first discrepant case, the senior screening radiologist noticed the tumour only on one projection, whereas the radiology resident noticed the tumour on two projections. In the second discrepant case, the pathologically proven tumour was located right at the border between the left upper lateral quadrant and the left upper medial quadrant. The senior screening radiologist located this pathologically proven tumour in the left upper lateral quadrant, whereas the radiology resident located this tumour in the left upper medial quadrant.

Discussion

During locally organized quality assurance evaluation sessions for screening radiologists participating in a Southern region of the dutch breast cancer population-based screening programme, we noticed that individual screening radiologists did miss tumours which in our opinion could be detected at a distance. All screening radiologists have been trained to make an initial overview assessment and in daily practice all screening mammograms by default are initially presented at the diagnostic working station as an overview with the four prior images of the previous screening round if available at the top and the four images of the current screening round at the bottom. However, while observing

Table 1 Mammographic breast density categories assigned by the readers

	Mammograms with a pathologically proven malignancy (*n*)		Normal mammograms (*n*)	
Mammographic breast density category	Reader 1[a]	Reader 2[b]	Reader 1[a]	Reader 2[b]
Almost entirely fatty	13	4	28	9
Scattered areas of fibroglandular density	12	20	17	39
Heterogeneously dense	3	1	10	7
Extremely dense	0	3	1	1

[a] Senior screening radiologist

[b] Radiology resident

Table 2 Lesion types assigned by the readers

	True positives: mammograms with a pathologically proven malignancy (*n*)		False positives: normal mammograms (*n*)	
Lesion type	Reader 1[a]	Reader 2[b]	Reader 1[a]	Reader 2[b]
Well-defined mass	1	1	1	3
Ill-defined mass	16	15	11	4
Spiculated mass	5	5	0	0
Architectural distortion	2	1	0	0
Focal asymmetry	1	1	3	3
Cluster of microcalcifications	3	2	1	0

[a] Senior screening radiologist

[b] Radiology resident

screening radiologists participating in our locally organized quality assurance evaluation sessions, we noticed that screening radiologists did get absorbed by the screen of the diagnostic working station (i.e. a short distance between the eyes of the screening radiologist and the screen of the diagnostic working station) assumingly looking for subtle abnormalities and microcalcifications. When the screening radiologists were forced to put their seat backwards and look at the screening mammogram again at a distance, new suspicious findings were noted.

Our study confirms that screening mammograms with pathologically proven malignancies which were initially missed by an individual screening radiologist can indeed be identified by an experienced screening radiologist but also by a resident in radiology without prior training in screening mammography, when read from a distance. Distant reading provides a better overview of the mammogram compared to close reading, which may help to identify ill-defined masses, spiculated masses, architectural distortions and focal asymmetries. The importance of an initial glance and global impression when viewing chest radiographs has already been reported in the 1980s [9]. Our study shows that this also applies for screening mammography. We suggest that screening mammograms should be presented as an initial eight-view hanging (two prior MLO views and two prior CC views at the top and two current MLO views and two current CC views at the bottom) to allow a general assessment then followed by current and prior hangings.

It is an interesting point that the majority of proven malignancies in this study set were ill-defined masses and that reader 1 (the senior screening radiologist) was also the least specific regarding this imaging feature when it came to recalling patients. Ill-defined masses and asymmetries tend to be quite subjective findings and have been shown to be the most commonly missed imaging finding [10, 11]. Perhaps the relatively subjective nature of these findings is a contributing factor as to why these mammograms were not recalled by the original screening radiologists who "missed" the malignancy in the first instance.

Screening radiologists should be aware that malignant tumours which were initially missed can be detected from a distance. If screening radiologists would add this specific task of reading each screening mammogram from a distance, the tumour detection rate might improve. This would not necessarily imply that each screening mammogram should also be read from exactly 1.5 m distance, such as was done in our study. An active mindset of the screening radiologist, i.e. reinforcing the knowledge that it is important to take a global view of the mammogram, could suffice. Although we did not assess reading time in our study, we estimate that the extra time needed for a distant view does not have to exceed 15 s. In comparison, one previous study reported a mean interpretation time for digital screen mammography between 171 and 300 s [12]. In addition, technical adjustments to the reading software of the diagnostic working station, for example minifying options, could prove to be helpful in practice.

A limitation of the use of a test set in our study is that it does not resemble daily screening practice [13]. Total recall rates of 44/84 (52.3%) and 35/84 (41.7%) and false positive recall rates of 16/56 (28.6%) and 10/56 (17.9%) (for the senior screening radiologist and radiology resident, respectively) were artificially high in this study and not applicable in a real screening situation. For comparison, recall rates of 10.0% for first and 6.7% for subsequent mammograms are recommended targets [14]. The readers may have been more focused to detect abnormalities compared to screening radiologists in daily practice. However, the readers were unaware of the number of screening mammograms with pathologically proven malignancies in the test set, and before finishing analyzing the test set, the readers were also unaware of the total number of screening mammograms included. It should also be noted that besides perceptual errors, there are many other reasons in the daily practice of reading

screening mammograms that could explain missing a tumour by an individual screening radiologist [15], for example, fatigue, time of the day, interruption in the reading room, recently too many false positives, the number of patients recalled in the same reporting session, and recent experiences at assessment clinics. Other limitations of our proof-of-concept study were the relatively small sample size, the use of only two readers, and non-matching of readers for experience and recall rates. In order to determine whether adding the specific task of reading each screening mammogram from a distance might indeed improve the tumour detection rate and improve the individual performances of screening radiologists, a much larger study is needed including a greater mix of cases and including different reading strategies.

Conclusion

To conclude, some malignant tumours missed by an individual screening radiologist can be detected from 1.5 m distance. Therefore, we recommend that screening radiologists consciously take a distant (i.e. global) view before closely evaluating the mammogram in detail.

Acknowledgements The authors thank M.S.O. van Wissen, MD for reading the test set during her residency in radiology.

Compliance with ethical standards

Conflict of interest Authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Informed consent In accordance with the national Data Protection Act, women were informed at screening that their data could be used anonymously for scientific research. Women can opt out, if they do not agree. The test set only consisted of screening mammograms of women who agreed to opt in.

References

1. International Agency for Research on Cancer. World Health Organization. GLOBOCAN 2012. Available via http://globocan.iarc.fr/Pages/fact_sheets_population.aspx. Accessed 7 May 2016
2. Warner E (2011) Clinical practice. Breast-cancer screening. N Engl J Med 365:1025–1032
3. Fracheboud J, van Luijt PA, Sankatsing VD et al (2014) National evaluation of breast cancer screening in the Netherlands. 1990–2011/2012 (XIII). Thirteenth evaluation report. National Evaluation Team for Breast cancer screening (NETB), Rotterdam
4. Boyer B, Hauret L, Bellaiche R, Gräf C, Bourcier B, Fichet G (2004) Retrospectively detectable carcinomas: review of the literature. J Radiol 85:2071–2078
5. Saarenmaa I, Salminen T, Geiger U, Holli K, Isola J, Kärkkäinen A, Pakkanen J, Piironen A, Salo A, Hakama M (1999) The visibility of cancer on earlier mammograms in a population-based screening programme. Eur J Cancer 35:1118–1122
6. Timmers JM, van Doorne-Nagtegaal HJ, Zonderland HM et al (2012) The breast imaging reporting and data system (BI-RADS) in the Dutch breast cancer screening programme: its role as an assessment and stratification tool. Eur Radiol 22:5
7. American College of Radiology (2013) American College of Radiology (ACR) Breast Imaging Reporting and Data System Atlas (BI-RADS), 5th edn. Reston (Va)
8. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33:159–174
9. Nodine CF, Kundel HL (1987) Using eye movements to study visual search and to improve tumor detection. Radiographics 7:1241–1250
10. Daly CA, Apthorp L, Field S (1998) Second round cancers: how many were visible on the first round of the UK National Breast Screening Programme, three years earlier? Clin Radiol 53:25–28
11. Goergen SK, Evans J, Cohen GP, MacMillan JH (1997) Characteristics of breast carcinomas missed by screening radiologists. Radiology 204:131–135
12. Haygood TM, Wang J, Atkinson EN, Lane D, Stephens TW, Patel P, Whitman GJ (2009) Timed efficiency of interpretation of digital and film-screen screening mammograms. AJR Am J Roentgenol 192:216–220
13. Timmers JM, Verbeek AL, Pijnappel RM, Broeders MJ, den Heeten GJ (2014) Experiences with a self-test for Dutch breast screening radiologists: lessons learnt. Eur Radiol 24:294–304
14. Schell MJ, Yankaskas BC, Ballard-Barbash R, Qaqish BF, Barlow WE, Rosenberg RD, Smith-Bindman R (2007) Evidence-based target recall rates for screening mammography. Radiology 243:681–689
15. Pow RE, Mello-Thoms C, Brennan P (2016) Evaluation of the effect of double reporting on test accuracy in screening and diagnostic imaging studies: a review of the evidence. J Med Imaging Radiat Oncol 60(3):306–314