ABSTRACT

Objectives: We aim to contribute to the assessment of the screening performance in Flanders (Belgium) and to identify valuable mammograms for subsequent studies and training.

Materials and Methods: Initially negative prior screening mammograms (sMx) of 210 women with confirmed breast cancer detected by the Flemish screening programme between 2011–2013 were reviewed by a highly experienced radiologist. The review of the prior sMx was performed in three steps: 1) only prior mammograms available; 2) with index sMx (=subsequent positive sMx) present; 3) with index sMx and clinical information present.

Results: The radiological review yielded 94 (45%) mammograms ‘without suspicious lesions’, 77 (37%) ‘with minimal signs in at least one breast’, and 39 (19%) ‘with clearly visible tumours’. In univariate analyses, the reclassification of prior sMx was significantly associated with the date of the prior sMx, the need for a third reader for arbitration, image quality and the detector system used (computed radiography versus direct readout digital radiography), and it was not associated with the interval between screening rounds, age at prior sMx, breast density, or tumour characteristics (<T2 versus ≥T2, in situ versus invasive). In multivariate analyses, the date of the prior sMx (p = 0.001), need for arbitration (p = 0.001) and image quality (p < 0.001) remained significantly associated with the reclassification.

Conclusion: This retrospective review reclassified 19% of the sMx as clearly visible tumours. With this, the Flemish screening programme performs in accordance with similar studies. The sMx reviewed in this study, form a valuable set of mammograms for training and further research.
INTRODUCTION

Breast cancer screening programmes have substantially increased the number of early detected cancers [1]. However, studies have made clear that current screening programmes only capture about 70% of all breast cancers that occur in participating women [2–4]. To improve cancer detection by mammography screening the European guidelines advise quality control using predefined performance indicators and quality assurance including review and training. An important performance indicator is rating interval cancers (breast cancers arising after a negative screening episode and before the next scheduled screening round). Performing a radiological review of prior screening mammograms (sMx) of interval cancers is part of the quality assurance and also an important teaching tool [1]. Screen-detected cancers have different characteristics than interval cancers [3, 5], and it is therefore useful to also review the priors of screen-detected cancers in order to improve the programme’s quality [1, 6].

This study comprises a review of confirmed breast cancer cases detected by the Flemish screening programme. The aims were to quantify the proportion of visible tumours on the prior sMx and to gather insight into associated variables that may hinder cancer detection, such as breast density, age, image quality, imaging technique, tumour size, type of tumour, need of arbitration, screening interval, and date of prior sMx. The study also aimed to identify a valuable set of sMx for training and subsequent studies.

MATERIALS AND METHODS

In the breast cancer screening programme in Flanders, biennial two-view mammographic screening is offered free of charge to women aged 50–69 years. Two radiologists (first and second reader) independently evaluate the screening mammograms, with third reader arbitration if needed. Between 2009–2013, 254,350 women participated in the Flemish Breast Cancer Screening Programme. From this group, cases for review were selected based on the following inclusion criteria: 1) informed consent for use of data in scientific research, 2) participation in minimum two consecutive screening rounds, 3) a screening interval of 16–30 months, 4) the index sMx (latest sMx) in 2011, 2012, or 2013 resulted in a referral for further diagnostic workup confirming and correctly documenting breast cancer, 5) where the prior sMx (previous sMx) was considered negative, 6) where the index and prior sMx were digital and available in the PACS (Picture Archiving and Communication System) at the Centre for Prevention and Early Detection of Cancer. In total, 292 cases met these inclusion criteria. From those a predefined sample size of 210 was selected by standard SPSS algorithms for random selection.

The 210 prior sMx were thoroughly reviewed by a single, highly experienced radiologist (reading > 10,000 sMx/year since 2006). The review followed a stepwise procedure: 1) review of prior sMx, in the absence of other images or information, 2) review of prior sMx with index sMx (subsequent positive screening mammogram) present, and 3) review of prior sMx, where index sMx and clinical information on tumour localization and characteristics (size, type, and stage) from diagnostic follow up were present. All steps were performed per case in succession. The expert radiologist reviewed all prior sMx for the presence of malignancy, the image quality, and breast density. The reviewing radiologist was not informed of the purpose of the study.

Possible associations between relevant variables and the intermediate (step 2) or final classification (step 3) were studied in univariate (chi-square) and multivariate analyses (logistic regression). 1) Breast density (≤25%, 26–50%, >50%), 2) age (50–54, 55–59, 60–64, 65–69 years), 3) tumour size (<T2 versus ≥T2), 6) type of tumour (in situ versus invasive), 7) the need of a third reader for arbitration during the original reading process of the prior sMx (arbitration, no arbitration), 8) the interval between prior and index screening (17–20, 21–24, 25–28 months), and 9) the date of screening of prior sMx (earliest, intermediate and latest tertile) were tested. Tertiles were used instead of screening years due to an imbalanced distribution of cases across calendar years (see Table 1).

Because of the limited number of clearly visible tumours in the intermediate and final classification, bootstrap validation with bias correction and accelerated bootstrap interval was performed. Statistical significance was set at p < 0.05.

In the multivariate analysis, the group of clearly visible tumours was first compared with the compound group of minimal and no signs, subsequently the group of clearly visible tumours was compared with the no signs group only.

RESULTS

DESCRIPTIVE CHARACTERISTICS OF SMX

Table 1 lists data from prior and index sMx and diagnostic follow up.

The sMx dataset contained images of 102 left, 103 right, and 5 bilateral breast cancers.
EXPERT REVIEW OF PRIOR SMX

The results of the expert review are summarized in Table 2.

By reviewing prior sMx alone (step 1), 24 of the sMx (11.4%) were labelled ‘probably malignant’ and might have been referred. The intermediate classification (step 2), prior sMx with index sMx present, identified 25 cases (11.9%) with ‘clearly visible tumours’. The final classification of prior sMx (step 3), including the use of index images and clinical information, revealed 39 ‘clearly visible tumours’ (18.6%).

UNIVARIATE ANALYSES

The intermediate classification was significantly associated with the date of prior sMx (p <= 0.001) and the need of arbitration on the prior sMx (p = 0.002). The final classification was significantly associated with the date of the prior sMx (p <= 0.001); the need of arbitration (p = 0.004), also with the image quality (p = 0.004) and the detector system used (CR versus DR) (p = 0.036). See Table 3. More ‘clearly visible tumours’ were detected in older sMx, sMx that required arbitration, in sMx of inferior quality, and in those using CR-technique.
When clearly visible tumours were compared to the compound group of minimal and no signs, the need of arbitration on the prior sMx (p = 0.005) and the date of the prior images (p = 0.044) were independently significantly associated with false negative clearly visible tumours in step 2 (i.e., only using prior and index images). When clearly visible tumours were compared only to the group of no signs, the significance level for the need of arbitration (p = 0.001) and date of priors (p = 0.004) appeared even higher.

In step 3, the final classification (i.e., with prior and index images and clinical information available), the need of arbitration (p = 0.001) and the date of the prior images (p = 0.006) were still independently significantly associated with false negative clearly visible tumours. Furthermore, the image quality was statistically significant (p < 0.001). These conclusions held, whether comparing to the compound group of minimal and no signs or only to the no signs group. See Table 4.

All statistically significant associations were confirmed by bootstrap validation.
DISCUSSION

This review of a substantial set of ‘initially negative’ prior sMx resulted in 39 (19%) being labelled as ‘clearly visible tumours’. This result is in accordance with similar studies [6, 7]. It concerns tumours missed twice during the normal screening procedure (by the first and second reader, or if arbitration was necessary, by the third reader and one of first two readers) and are therefore very valuable for training.

The 19% missed tumours cannot automatically be considered ‘screening errors’, for several reasons:

1. the proportion of cases with ‘clearly visible tumours’ based on image review alone was 1/3 lower, at 12%. The availability of clinical information is known to alter the reading outcome [8, 9].

2. Even if we tried to reproduce the conditions of routinely assessing sMx in the screening programme, the radiologist’s attention was presumably triggered by the clustering of challenging image sets, the slightly different protocol form, the specific categorization, and the stepwise assessment for the review [3, 9].

3. The normal response of the human mind to low probability events (i.e., the low prevalence of cancer in the sMx) can be a substantial contributor to false negative errors in breast cancer screening [4].

Therefore, the clustering of challenging sMx in this study may have affected the reader’s awareness and the results of the review.

The image quality was significantly associated with the final classification of clearly visible tumours, as well as with the interval classification (Table 4).

Table 4 Multivariate analyses: Variables associated with the interim or final classification after reviewing prior mammograms.

| VARIABLE & CLASSES       | CLEARLY VISIBLE TUMOURS COMPARED TO NO OR MINIMAL SIGNS | CLEARLY VISIBLE TUMOURS COMPARED TO NO SIGNS |
|--------------------------|--------------------------------------------------------|---------------------------------------------|
|                          | ODDS RATIO 95% CONFIDENCE INTERVAL p-VALUE            | ODDS RATIO 95% CONFIDENCE INTERVAL p-VALUE |
| Need of arbitration on prior images | 0.005 0.001                                         |                                            |
| No arbitration            | 1 1                                                   |                                            |
| Arbitration              | 4.85 (1.61–14.61) 0.005 16.65 (2.98–93.00) 0.001     |                                            |
| Date of prior imaging    | 0.044 0.004                                           |                                            |
| Earliest tertile         | 11.13 (1.39–88.93) 0.024 39.71 (3.43–459.09) 0.003   |                                            |
| Intermediate tertile     | 5.75 (0.68–48.72) 0.109 12.30 (1.06–142.17) 0.045   |                                            |
| Latest tertile           | 1 1                                                   |                                            |
| Image quality at the tumour side | 0.510 0.220                                      |                                            |
| Good                     | 1 1                                                   |                                            |
| Not good                 | 1.60 (0.52–3.78) 0.510 2.02 (0.66–6.20) 0.220        |                                            |

A. MULTIVARIATE ANALYSES: VARIABLES ASSOCIATED WITH THE INTERIM CLASSIFICATION (STEP 2) AFTER REVIEWING PRIORS WITH INDEX IMAGES PRESENT.

B. MULTIVARIATE ANALYSES: VARIABLES ASSOCIATED WITH THE FINAL CLASSIFICATION (STEP 3) AFTER REVIEWING PRIORS WITH INDEX IMAGES AND CLINICAL INFORMATION PRESENT.
performance. Often higher sensitivity is found with higher cancer detection rates and less interval cancers, especially in dense breasts [2, 10, 11].

Since this review was performed by a single – albeit highly experienced – radiologist, the results of this retrospective review could not be corrected for inter-observer variability. This is a major limitation of this study.

CONCLUSION

The radiological review yielded 94 (45%) mammograms ‘without suspicious lesions’, 77 (37%) ‘with minimal signs in at least one breast’, and 39 (19%) ‘with clearly visible tumours’. These results are in line with similar studies.

The screening mammograms assessed in this review are valuable for training and subsequent studies.

DATA ACCESSIBILITY STATEMENT

All relevant documentation or data in order to verify the validity of the results presented is available, but not openly. Due to the nature of this research, participants of this study did not agree for their data to be shared publicly.

ABBREVIATIONS

sMx: screening mammogram
CR: computed radiography
DR: direct readout digital radiography
PACS: Picture Archiving and Communication System
BI-RADS: Breast Imaging Reporting & Data System

ETHICS AND CONSENT

All participants gave their written informed consent for the Breast Cancer screening programme in Flanders, including its quality assessment. This research project was approved by the Ethics Committee of Ghent University hospital (B670201318961).

ACKNOWLEDGEMENTS

The authors especially would like to acknowledge the dedication of dr Margarete (Griet) Mortier, who performed the expert review reported in this paper. We thank Roos Colman for statistical advice, dr Luc Bleyen and dr Soetkin De Brucker for their assistance in data collection and handling, the Centre for Cancer detection (Centrum voor Kankeropsporing vzw, CvKO) for their cooperation and providing the data for this work.

FUNDING INFORMATION

This study has received funding by the Flemish agency for Innovation and Entrepreneurship (Vlaio) (grant number 130472). Apart from approving this study and providing financial sources Vlaio had no further involvement in this study.

COMPETING INTERESTS

TK is employee at Barco, Beneluxpark 21, 8500 Kortrijk, Belgium, which part funded the research grant for this project. All other authors have no competing interests.

AUTHOR AFFILIATIONS

Lieve Vandendaele orcid.org/0000-0003-3254-6430
Universiteit Gent & CvKO, Centrum voor Kankeropsporing, BE
Svetlana Jidkova orcid.org/0000-0002-4250-2215
Universiteit Gent & CvKO, Centrum voor Kankeropsporing, Data Science Institute UZ Gent, BE
Koen Van Herck orcid.org/0000-0003-0717-2406
Universiteit Gent, BE; Belgian Cancer Registry, BE
Tom Kimpe orcid.org/0000-0003-2117-0343
Barco, BE
Veerle Verschuere orcid.org/0000-0002-0424-5174
Universiteit Gent & CvKO, Centrum voor Kankeropsporing, BE

REFERENCES

1. Perry N BM, de Wolf C, Törnberg S, Holland R, von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. In: European Commission DHaC, editor. 4th edn. ed. Luxemburg: European Union Publications Office; 2006. http://www.euref.org/european-guidelines.
2. Timmermans L, Bleyen L, Bacher K, et al. Screen-detected versus interval cancers: Effect of imaging modality and breast density in the Flemish Breast Cancer Screening Programme. Eur Radiol. 2017; 27(9): 3810–9. DOI: https://doi.org/10.1007/s00330-017-4757-4
3. Houssami N, Hunter K. The epidemiology, radiology and biological characteristics of interval breast cancers in population mammography screening. NPJ Breast Cancer. 2017; 3(1): 1–13. DOI: https://doi.org/10.1038/s41523-017-0014-x
4. Evans KK, Birdwell RL, Wolfe JM. If you don’t find it often, you often don’t find it: Why some cancers are missed in breast cancer screening. PLOS ONE. 2013; 8(5): e64366. DOI: https://doi.org/10.1371/journal.pone.0064366
5. Bellio G, Marion R, Giudici F, et al. Interval breast cancer versus screen-detected cancer: Comparison of clinicopathologic characteristics in a single-center analysis.
Clin Breast Cancer. 2017; 17(7): 564–71. DOI: https://doi.org/10.1016/j.clbc.2017.04.001

6. Hoff SR, Samset JH, Abrahamsen A-L, Vigeland E, Klepp O, Hofvind S. Missed and true interval and screen-detected breast cancers in a population-based screening program. Acad Radiol. 2011; 18(4): 454–60. DOI: https://doi.org/10.1016/j.acra.2010.11.014

7. Hovda T, Tsuruda K, Hoff SR, Sahlberg KK, Hofvind S. Radiological review of prior screening mammograms of screen-detected breast cancer. Eur Radiol. 2021; 31(4): 2568–79. DOI: https://doi.org/10.1007/s00330-020-07130-y

8. Hofvind S, Skoane P, Vitak B, et al. Influence of review design on percentages of missed interval breast cancers: Retrospective study of interval cancers in a population-based screening program. Radiology. 2005; 237(2): 437–43. DOI: https://doi.org/10.1148/radiol.2372041174

9. Lekanidi K, Dilks P, Suaris T, Kennett S, Purushothaman H. Breast screening: What can the interval cancer review teach us? Are we perhaps being a bit too hard on ourselves? Eur J Radiol. 2017; 94: 13–5. DOI: https://doi.org/10.1016/j.ejrad.2017.07.005

10. Prummel MV, Muradali D, Shumak R, et al. Digital compared with screen-film mammography: Measures of diagnostic accuracy among women screened in the Ontario breast screening program. Radiology. 2016; 278(2): 365–73. DOI: https://doi.org/10.1148/radiol.2015150733

11. Séradour B, Heid P, Estève J. Comparison of direct digital mammography, computed radiography, and film-screen in the French national breast cancer screening program. AJR Am J Roentgenol. 2014; 202(1): 229–36. DOI: https://doi.org/10.2214/AJR.12.10419