Abstract

Purpose: To analyse whether the screening performance parameters of the Maltese National Breast Screening Programme first screening round met requirements set by European standards. The association between screening age and results of screening performance parameters was also investigated.

Method: Quantitative methodology was used to review examinations of women who were recalled for a technical recall or further assessment rates. All accessible members of the population recalled during the first round were retrospectively reviewed resulting in a sample of 2300 recalled examinations.

Results: Malta's first screening round met the European Guidelines recommendations for technical repeat rate (0.26%), early recall rate (0.45%), breast cancer detection rate (13.77 per 1000 women) and Positive Predictive Value of screening test (7.58%). However, local recall rate (18.53%) and further assessment rate (18.27%) were higher than recommended.

The Chi square test showed a statistically significant difference ($p \leq 0.05$) in recall rates between the compared age groups, as younger women (51-55 years) were more likely to have a negative diagnosis after the initial mammogram whereas older women (56-60 years) were more likely to be recalled. There was no age discrepancy ($p \geq 0.05$) in local breast cancer detection rate and positive predictive value of screening test.

Conclusion: Although the Maltese first screening round performed well, this study found deficiencies in recall and further assessment rates, which according to literature may result in psychological morbidity and inefficient use of screening resources. This study also concluded that when a cohort is analysed, age is not as significant as the screening round itself (first/subsequent).

Keywords

Introduction

Breast cancer is the most common cancer among Maltese women and accounts for a third of all female cancer deaths. Every one in twelve Maltese women will get breast cancer in her lifetime; a significantly higher proportion than the European average.[1] In October 2009, Malta's Department for Health agreed to implement a National Breast Screening Programme (MNBSP).[2] The aim of screening was to reduce morbidity and mortality from the disease without adversely affecting the health of participants.[3,4] Key to achieving this aim were high levels of quality within the entire screening process.[4] The MNBSP strives to protect the dignity and privacy of women, while offering an effective service at the highest levels of quality to diagnose and treat breast cancers at the earliest possible stage, and also meeting the European clinical standards.[4]

The first MNBSP round commenced in October 2009 and ended in February 2013. The local screening programme provided free screening, every three years for all women aged 50 to 60, resident in the Maltese Islands. This age range was selected for the first screening round since this age group was deemed to be the 'most at risk' of developing breast cancer.[1] Malta's Department for Health additionally states that as soon as more human resources became available, the programme will be extended even to older women.[1]

Double reading of mammograms was performed locally, as recommended for programmes in their first round of screening.[4] When comparing dual to single reading, studies have shown that sensitivity increased by 5–10% without a significant effect on recall rates.[5] Consensus arbitration of discordant double read cases also improved reading performance since it decreased recall rates and increased Positive Predictive Value (PPV).[4,6,7] Since consensus arbitration was not practised locally, as per radiologists' choice, women who were recalled by one radiologist experienced additional examinations during further assessment clinics. This could have resulted in higher recall rates and lower PPV.

A high quality screening service can be achieved through the use of targets, performance parameters and audits.[4] Although audits of performance parameters were undertaken in European screening units, this research was the first of its kind in Malta. The findings offer valuable contribution for future advancement of the local screening unit as they helped to identify areas of strength, as well as areas needing improvement.

This study aimed to retrospectively audit results of the first screening round of the MNBSP and to assess whether the service fulfilled the European standard requirements. Thus the study objectives were to measure screening performance parameters (Recall Rate, Technical Recall (TR) Rate, Further Assessment Rate, Early Recall (ER) Rate, Breast Cancer Detection Rate, PPV of Screening Test) of the local first screening round; calculate the different types of clinical examinations performed during further assessment clinics; compare the rates of the screening performance parameters to the levels set by European and United Kingdom (UK) guidelines and investigate any association between screening age and results of screening performance parameters.

Literature Review

Several online databases such as Medline, Cumulative Index Nursing and Allied Health Literature, and ScienceDirect were utilised to access e-journals. Although most of the identified literature was undertaken after the inception of screening programmes, around 10– 20 years ago, it was seminal to the study and was therefore included in this review. Only peer-reviewed, European studies written in the English language were included due to the similarity of breast screening programmes.

Quality Assurance (QA)

Ensuring the quality of a screening service is vital. This could be achieved by early monitoring of screening performance parameters of the unit, potentially optimising the use of resources and ultimately producing an observable reduction in mortality.[4] Strict adherence to quality assurance and quality control guidelines must be practised in all mammography facilities to ensure accurate diagnosis, thus minimising false positive mammograms.[8,9] False-positive rate refers to recalls for further assessment which turn out to be normal or benign.[10] This is one of the reasons for the ongoing screening debate, since it gives rise to negative effects, namely: financial costs to the health service and psychological strain on the women.[10,11] Since the majority of screening mammograms are normal, radiologists' record of reporting should demonstrate high specificity avoiding false-positive mammograms.[12]

In addition to false positive rates, several other performance parameters were identified by '*The European Guidelines for QA in Breast Cancer Screening and Diagnosis*'.

Recommended acceptable levels for each parameter were also set.[4] Table 1 defines the screening performance parameters audited in this study and indicates the acceptable EU and UK levels for programmes in their first screening round. [1,4,13,14,15]

Several research studies investigated these screening performance parameters.

Recall Rate

No increase in cancer detection rates and in screening sensitivity beyond a recall rate of 4.8% were recorded.[16] High recall rates were found to signify that resources are used inefficiently in women undergoing unnecessary follow-up procedures.[17,18] Conversely, rates lower than 1% were associated with reduced cancer detection and increased interval cancers.[4] Previous research states that recall rates were influenced by several factors including training and experience of radiologists, image quality, the volume of mammograms interpreted and the age of screened women.[19,20,21]

Early Recall Rate

ER was found to be associated with a low predictive value for malignancy and thus every effort should be made to obtain a definitive diagnosis at initial assessment. [4,22,23,24] Additionally, a study analysing 110 women who were recalled early, revealed that 3.6% had invasive cancer, 0.9% had DCIS, while 84% had benign findings.[25]

PPV of Screening Tests

The European guidelines did not define an acceptable PPV of Screening Test, whilst the UK National Health Service Breast Screening Programme (NHSBSP) defined the minimum standard PPV as 2.7% or more at first screenings.[26] Due to the similarity between the UK NHSBSP and the MNBSP, this value was taken and accepted as the recommended level.[27] Although recall rates of first screening rounds are generally high, the PPVs are usually low. In subsequent screenings the previous investigations could be referred to without the need for recall. Therefore only new findings would be fully investigated, thus lowering recall rates and further increasing PPV.[28, 20]

Performance in European Breast Screening Programmes

Performance indicators for mammography screening in 17 European countries showed some discrepancies, with recall rates ranging from 1.3% to 18.4%.[29] First screening rounds resulted

in detection rates varying from 10.7 per 1000 women screened in Copenhagen to 3.6 per 1000 in Finland. [29, 30] This difference between countries should be interpreted with caution due to variations in screening and interpretation methods used within the various programmes. For instance the UK NHSBSP and the Netherlands followed very different recall policies; the latter deliberately aiming at a very low recall rate. Other differences included the use of one versus two-view mammography, screening interval, double-reading and methods for arriving at a resolution when double-reading led to different conclusions. [29, 30]

Effect of Age

Age is another independent factor predicting the accuracy of screening mammography.[31, 32] As screening age increased, recall rates decreased. Sensitivity, specificity and PPV increased with age thus increasing the accuracy of screening. [31, 32]

A study that analysed 215,665 mammograms revealed that the PPV was inversely related to the recall rates for age. [31] Recall rates decreased from 7.3% for the youngest women to 4.9% for the oldest women, whilst the PPV rose from 1.9% to 12.7%. This outcome was supported by an extensive study of 1.5 million examinations reporting an increase in PPV by age. [33] This was reasonable given the relatively higher breast density and lower incidence of disease in younger women. [34, 20]

Alternatively, studies analysing cohorts of women undergoing successive screening rounds indicated no clear association between age and risks of false-positive mammograms, concluding that such risks are much higher in first screening rounds.[35, 36] Nevertheless, there were no considerable reductions from second rounds onwards, regardless of women's increasing age. Therefore, the association between age and false-positive recall rates seen in several cross-sectional studies could be overestimated because of the proportion of young women without a previous mammogram in first screening rounds. This implies that age is not as significant as the first screening round itself. [35]

Psychological Impact of Recall

One of the reasons for maintaining screening performance parameters within the acceptable EU/UK levels is that women undergoing further assessment suffer significantly greater

adverse psychological consequences when compared to those with an immediate normal result.[37] One study found that 30% of recalled women felt very anxious on receipt of their recall letter, whilst another study reported that 40% of women reported feeling very frightened when they received their recall letter.[38, 39] 16% of respondents felt certain that they had breast cancer.[38,39] Such a psychological impact was reported by recalled women before[40], during[41], and immediately after recall.[42] Women who had clear results after an invasive investigation and those who were placed on ER experienced the highest levels of anxiety.[43,44]

The adverse psychological impacts of false-positive recalls could discourage women from re-attending for subsequent screening.[45] Absenting from screening consequently negates any benefits of the service and increases the mortality rate.[46] To ensure accurate and low false-positive recall rates it is important to regularly measure screening performance parameters and provide feedback to screening personnel.[47]

Methods

Following ethical approval, a quantitative, non-experimental, descriptive, comparative, retrospective design was used to audit the examinations of women who were recalled for either a TR or further assessment. The study was conducted at the Maltese screening unit on the basis of data obtained from computerised databases. After every recall session, radiographers are responsible for recording all pertinent data on a database at the unit. This included recommended and performed TR and further assessments and outcomes of the recall clinics. All the women recalled during the first screening round (October 2009 to February 2013) were included in this study (n=2300). Retrospective data collection was performed between the 4th and 30th August 2014.

The information was audited by means of three data record sheets derived from the 4th edition of the European guidelines for QA in breast cancer screening and diagnosis.[4] To provide a more sensitive measure of performance and to be able to investigate the effect of age on each parameter, data was further classified in age groups (51 to 55 and 56 to 60 years).[4]

Data Record Sheet 1 (Table 2) was used to audit the possible outcomes of the initial screening mammogram, namely *Negative Examinations, Radiologist Recommended TR,*

Actually Performed TR, Radiologist Recommended Further Assessment and *Actually Performed Further Assessment.*

Data Record Sheet 2 (Table 3) was utilised to audit investigations performed during further assessment clinics, namely the *Number of Clinical Breast Examinations,* the *Number of Additional Imaging,* which was subdivided into *Number of Further Views and Ultrasounds,* and *Number of Interventional Procedures,* which was subdivided into *Number of Radiologist Recommended Cytology and Actually Performed Cytology* and *Number of Radiologist Recommended WCNB and Actually Performed WCNB.*

Data Record Sheet 3 (Table 4) was used to audit the outcome of the overall screening process, namely the *Total Negative Examinations, ER following Further Assessment, Number of DCIS cases, Number of Invasive Cancers detected* and *Unknown Outcome.* Data Record Sheet 3 was also used to audit the timing of breast cancer detection, namely *At Initial Screen* and *During ER.*

The researcher undertook the following process for data collection:

Recalled women were classified either as *recommended TR* or as *recommended recall for further assessment*. Such classification was opted for, for the sake of consistency with terminology used at the unit. The number of recommended and performed TRs and further assessments were audited. Data from each recalled woman was then reviewed. This served to audit: the number of women who had a clinical breast examination, the number of women who had additional imaging (further views or ultrasound); the number of women who underwent an interventional procedure (FNA or WCNB).

The number of women recalled for further assessment and classified as requiring an ER was also audited. Cases that had undergone interventional procedures were reviewed to determine whether any malignancy had been diagnosed. The type of cancer diagnosed was also documented and diagnosed cancers were audited according to the time of their detection (at the initial screen or during ER).

Descriptive statistics were used to present the audit results and to calculate screening performance parameters for the MNBSP.[48] The *chi-square test* was used to establish whether

there was an association between categorical variables.[49] One of the categorical variables was the age group (51-55; 56-60) whilst others described the screening performance parameters. The *difference between two proportions test* was used to determine whether local screening performance parameters that failed to meet acceptable levels set by European guidelines, differed significantly from the guidelines.

Results

The audit results and rates of screening performance parameter rates of the Maltese first screening round are summarised in Table 2-5. The MNBSP met the recommendations given by European Guidelines for TR, ER and breast cancer detection rates which were found to be 0.26%, 0.45%, 13.77 per 1000 women respectively.[4] The PPV of the Screening Test for the Maltese first screening round was 7.58%, thus within the recommendations by the UK NHSBSP.[26] However, local recall and further assessment rates (18.53% and 18.27% respectively) were higher than those recommended by European guidelines (Table 6).

The majority of recalled women (n=2251, 83.3%) underwent additional imaging (further views or ultrasound); whilst only 23 (0.9%) were examined clinically, 427 (15.8%) women underwent interventional procedures (FNA and WCNB). A total of 27 (0.2%) women were diagnosed with DCIS, whilst 144 (1.2%) women were diagnosed with invasive breast cancer. The majority of cancers (n=169, 98.8%) were detected at the initial screen whilst 2 (1.2%) cancers were detected in women for whom on ER assessment was recommended at further assessment.

The Chi square test showed a statistically significant difference (p≤0.05) in recall rates between the compared age groups, as younger women (51 to 55 years) were more likely to have a negative diagnosis after the initial mammogram, whereas older women (56 to 60 years) were more likely to be recalled. There was no age discrepancy (p≥0.05) in the local breast cancer detection rate and PPV of the screening test.

Discussion

Higher recall rates were expected for first screening rounds when compared to subsequent ones. The local Recall Rate for the first screening round was higher than the acceptable rate for first screening rounds [28, 35] and also higher than European recall rates.[29] The different

screening and/or interpretation methods utilized locally may account for the differences in recall rates leading to issues of comparability of findings to other studies and European recall rates.

Differences in data collection may have also affected comparability of recall rates. Although this research did not investigate the effect of breast density on mammography performance, studies have shown that this factor could affect the performance parameters.[32] Screening programmes recording breast density demonstrated a great variation in breast density classification affecting comparisons of screening results.

The presence of symptoms at screening was another important modifier of performance estimates, as symptomatic women may have higher recall and PPV rates.[31] A few programmes, including the local one, recorded symptoms present at the time of screening. This factor was not included in the pre-established data record sheets and was therefore not investigated in this research.

Whilst low recall rates may decrease detection rates and increase incidence of subsequent interval cancers, high recall rates alternatively cause psychological morbidity and indicate inefficient use of screening resources.[4,35]

Additionally, differences in recall rates may have been affected by factors which were beyond the researcher's control, such as the prevalence of cancer in the population, radiologists' training, image quality, and in some instances, malpractice concerns. The training of radiologists is an important factor affecting the recall rate.[21] The researcher had no evidence whether this factor affected the recall rate locally. However, the fact that this was the first experience for local radiologists to work within a screening program might have had an impact on the recall rate as acknowledged in literature.[21] Local radiologists might have also recalled women with suspicious findings, even in the absence of a possibility of cancer, to avoid possible malpractice litigation. Additionally, reading an average of 3600 mammograms yearly, local radiologists are considered to be low volume readers compared with the stipulated 5000 yearly mammograms which European guidelines recommend.[3,38] As this was the first unit on the island using direct digital mammography, radiologists were new to the equipment possibly leading to increased recall rates.[39] Such high recall rates may also arise from the local absence of consensus or arbitration of double reading. Other studies have

suggested that double reading by consensus or arbitration, although being more labour intensive, decreases the recall rate.[7]

The local Further Assessment Rate exceeded European guidelines for the same reasons since the recall rate is the summation of further assessment and TR rates.[13] The small number of invasive investigations performed locally is considered as a strength since women with benign breast disease undergoing invasive investigations report higher levels of stress and anxiety lasting for several months after the additional investigations. [43, 50.]

A TR rate of 0.26% indicated high image quality locally, as the objective of a screening programme is to minimise the number of women undergoing technical recalls while achieving the optimum image quality.[20] To obtain such an excellent TR Rate careful selection and training of staff is necessary.

Low predictive value for malignancy in ER has been expressed in various studies[24,25] being consistent with this study's findings as the majority of cancers were detected at initial screen, with only 2 cancers detected at ER stage. Nevertheless the latter could represent late diagnosis and failure of the screening and assessment process since women were falsely reassured, and ultimately diagnosed with cancer.[4, 51]

Such a promising Breast Cancer Detection Rate could be due to the benefits of direct digital mammography when compared to film screen mammography, namely improving detection of calcification related malignancies and non-symptomatic cancers. Although European guidelines do not specify an upper limit on the desirable levels, a high breast cancer detection rate may suggest a longer sojourn time (the period in which the tumour is asymptomatic but detectable by screening) than expected and diagnosis of cancers that would not have been detected without screening.[52] However, over diagnosis has to be evaluated carefully with tumour characteristics in mind.

A PPV of the Screening Test of 7.58% can be interpreted as 13 women recalled per cancer detected. Thus 13 women had a false positive mammogram for each cancer detected during the local first screening round. Studies have shown that false positive mammograms negatively affect women's well-being; experiencing breast cancer related anxiety, worry and distress, when compared with women who receive normal mammogram results.[10,11] Whilst

false positive mammograms are undesirable, they are necessary to maximise the number of small cancers detected.

Literature demonstrated that the age of the screened population has an impact on recall rates documenting a decline in rates across age and in subsequent screens.[21, 32] In contrast this research showed lower recall rates in the 51 to 55 years age group. Moreover, results revealed that younger women were more likely to have a negative diagnosis after the initial mammogram, and older women were more likely to be recalled. In this research the age of the screened population was 50 to 60 years, whilst in literature the age of the screened population was of a wider age span. However, comparing local results to other studies with wider age span was still appropriate.

This may be due to the fact that the programme commenced with the older client spectrum, and invited those aged 56 to 60 in the first years and then moved down the age group and invited women aged between 51 and 55 in the following years. Thus the reduction in recall rates over the years can be attributed to accumulated experience in screening, allowing individual radiologists to attain a greater level of experience and competence. However there was no age discrepancy in the local cancer detection rate and PPV of the screening test, signifying that mammograms had similar accuracy in both younger and older women.

Limitations

Limitations included the lacking or varying definitions of recall found in literature. Definition evidently affects the recall rate and thereby its relationship with PPV and breast cancer detection rate, making comparison of findings difficult. The performance parameters measured in this research were based on definitions stated by the European Guidelines[4], thus permitting comparison of findings to acceptable levels set by these guidelines.

Although European Guidelines were much more comprehensive, measuring additional screening performance parameters was impractical due to timing restrictions and available resources.

Participant related factors such as breast density and the use of HRT were a further limitation. These were not utilised for analysis and could have influenced screening performance parameters results. Further studies are required to elucidate this issue.

Conclusions

Through analysis of the first round of the newly implemented MNBSP, it was concluded that the programme performed well since most of the performance parameters complied with European guidelines. This successful performance may be the result of joint efforts from all personnel, comprehensive QA measures and the use of digital techniques. Such positive performance can be used to build and maintain a high state of confidence among local radiologists and those responsible for the execution of the screening programme. A successful performance in the first screening round encourages new and repeat attendance to the Maltese screening programme.

This research has established that the local recall and further assessment rates did not comply with the acceptable levels set by the European guidelines, uncovering a problem in local clinical performance since according to literature high recall rates can cause psychological morbidity [53] and indicating inefficient use of screening resources.[35] In attempt to reduce recall and further assessment rates, the ongoing education of personnel particularly screening radiologists should be assured. Other recommendations include double reading by consensus or arbitration and interpretation of a high volume of mammograms.

This study also investigated any association between age and screening performance parameters and demonstrated that younger women were more likely to have a negative diagnosis after the initial mammogram whereas older women were more likely to be recalled. In addition there was no age discrepancy in the local cancer detection rate and PPV of the screening test, signifying that mammograms had similar accuracy in both younger and older women. Based on these outcomes, it was established that when a cohort is analysed, age is not as significant as the screening round itself.

Screening performance results highlight the importance of continuous surveillance of the local screening unit in order to sustain and improve performance and effectiveness. Thus it is suggested that regular QA is maintained to offer high quality mammographic screening.

Future research focusing on the effects of mammography equipment used, training of radiologists, volumes of mammograms read, and double-reading by arbitration is recommended.

Conflict of Interest Statement

Conflicts of Interest: None

3803 Words

References

References

1. Ministry for Health. Breast Care. Available at: https://ehealth.gov.mt/healthportal/health_institutions/hospital_services/mater_dei_hospital/surgical_outpatients/breast_care/bstcare_default.aspx 2013 (Accessed: 06 January 14).
2. Bezzina P. Exploring the Practice of Radiographers in the Mammography Unit of a General Hospital in Malta: A case study. PhD Thesis. University of East Anglia; 2007.
3. Michell M. J. Contemporary Issues in Cancer Imaging: Breast Cancer. New York: Cambridge University Press; 2010.
4. Perry N, Broeders M, de Wolf C, Tornberg S, Holland R, von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. 4th ed. Belgium: European Communities; 2006.
5. Azavedo E, Zackrisson S, Mejàre I, Heibert Arnlind M. Is single reading with computer-aided detection (CAD) as good as double reading in mammography screening? A systematic review. *BMC Medical Imaging* 2012;12:22.
6. Shaw C, Flanagan F, Fenlon H, McNicholas M. Consensus review of discordant findings maximizes cancer detection rate in double-reader screening mammography: Irish national breast screening program experience. *Radiology* 2009;250:354-62.
7. Hofvind S, Geller B, Rosenberg R, Skaane P. Screening-detected breast cancers: Discordant independent double reading in a population-based screening program. *Radiology* 2009;253:652-60.
8. Patnick J. NHS Breast Screening Programme Annual Review 2012. Sheffield: NHS Breast Screening Programme; 2012.
9. Giordano L, von Karsa L, Tomatis M, Majek O, de Wolf C, Lancucki L, et al. Mammographic screening programmes in Europe: organization, coverage and participation. *J Med Screen* 2012;19:72-82.

10. DeFrank J, Rimer B, Bowling J, Earp J, Breslau E, Brewer T. Influence of false-positive mammography results on subsequent screening: do physician recommendations buffer negative effects? *J Med Screen* 2012;19:35-41.

11. Keyzer-Dekker CMG, De Vries J, van Esch L, Ernst MF, Nieuwenhuijzen GAP, Roukema JA, et al. Anxiety after an abnormal screening mammogram is a serious problem. *Breast J* 2012;21:83-8.

12. Tabar L, Tot T, Dean PB. Breast Cancer - The Art and Science of Early Detection with Mammography: Perception, Interpretation, Histopathologic Correlation. 1st ed. USA: Thieme; 2004.

13. Liston J, Wilson R. Clinical Guidelines for Breast Cancer Screening Assessment. NHSBSP Publication no 49. 3rd ed. Sheffield: NHS Cancer Screening Programmes; 2010

14. National Quality Assurance Coordinating Group for Radiography. Quality Assurance Guidelines for Mammography including Radiographic Quality Control. NHSBSP Publication No. 63. Sheffield: NHS Cancer Screening Programmes; 2006.

15. Ferlay J, Shin H, Bray F, Forman D, Mathers C, Parkin D. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 2010;127: 2893-917.

16. Feig S. Adverse effects of screening mammography. *Radiol Clin North Am* 2004;42:807-19.

17. Otten J, Karssemeijer N, Hendricks J, Groenewound J, Fracheboud J, Verbeek A, et al. Effect of recall rate on earlier screen detection of breast cancers based on the Dutch performance indicators. *J Natl Cancer Inst* 2005;97:748-54.

18. Pharoah P, Sewell B, Fitzsimmons D, Bennet H, Pashayan N. Cost effectiveness of the NHS breast screening programme: life table model. *BMJ* 2013;346:15.

19. Esserman L, Cowley H, Eberle C, Kirkpatrick A, Chang S, et al. Improving the accuracy of mammography: volume and outcome relationships. *J Natl Cancer Inst* 2002;94:369-75.

20. Hofvind S, Ponti A, Patnick J, Ascunce N, Njor S, Broeders M, et al. False-positive results in mammographic screening for breast cancer in Europe: a literature and survey of service screening programmes. *J Med Screen 2012;*19:57-66.

21. Yankaskas BC, Klabunde CN, Ancelle-Park R, Renner G, Wang H, Fracheboud J, et al. International comparison of performance measures for screening mammography: can it be done?. *J Med Screen* 2004;11:187-193.

22. Macmillan Cancer Support. Possible results of breast assessment. Available at: http://www.macmillan.org.uk/Cancerinformation/Testsscreening/Breastscreening/Assessmentresults.aspx 2014 (Accessed: 26 April 2014).

23. Wilson R, Liston J. Quality Assurance Guidelines for Breast Cancer Screening Radiology. NHSBSP Publication No 59. Sheffield: NHS Cancer Screening Programmes; 2011.

24. Yasmeen S, Romano PS, Pettinger M, Chlebowski RT, Robbins JA., Lane DS, et al. Frequency and Predictive Value of a Mammographic Recommendation for Short-Interval Follow-Up. *J Natl Cancer Inst* 2003;95:429-36.

25. Pritchard MG, Kearins O, Lawrence G. Audit of short-term recall in the National Health Service Breast Screening Programme in the West Midlands. *Breast Cancer Res* 2006;8:48.

26. NHS Breast Screening Programme. Consolidated guidance on standards for the NHS breast screening programme. Sheffield UK: NHSBSP Publications; 2005.

27. Bennett R. Blanks R. Should a standard be defined for the Positive Predictive Value (PPV) of recall in the UK NHS Breast Screening Programme?, *Breast,* 2007;16:55-9.

28. Cancer Research UK. Breast Screening other issues. Available at: http://www.cancerresearchuk.org/cancer-info/cancerstats/types/breast/screening/Other-Issues/ 2012 (Accessed: 10 April 2014).

29. Lynge E, Olsen AH, Fracheboud J, Patnick J. Reporting of performance indicators of mammography screening in Europe. *Eur J Cancer Prev* 2003;12:213-22.

30. Biesheuvel C, Weigel S, Heindela W. Mammography Screening: Evidence, History and Current Practice in Germany and Other European Countries. *Breast Care* 2011;6:104-9.

31. Yankaskas BC, Cleveland RJ, Schell MJ, Kozar R. Association of recall rates with sensitivity and positive predictive values of screening mammography. *AJR Am J Roentgenol* 2001;177:543-9.

32. Carney PA, Miglioretti DL, Yankaskas BC, Kerlikowske K, Rosenberg R, Rutter CM, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. *Ann Intern Med* 2003;138:168-75.

33. Dean P, Pamilo M. Screening Mammography in Finland - 1.5 Million Examinations with 97 percent specificity. Mammography Working Group, Radiological Society of Finland. *Acta Oncol* 1999;38:47-54.

34. Broeders M, Scharpantgen A, Ascunce N, Gairard B, Olsen A, et al. Comparison of early performance indicators for screening projects within the European Breast Cancer Network: 1989-2000. *Eur J Cancer Prev* 2005;14:107-16.

35. Castells X, Molins E, Macia` F. Cumulative false positive recall rate and association with participant related factors in a population based breast cancer screening programme. *J Epidemiol Community Health* 2006;60:316-21.

36. Hofvind S, Thoresen S, Tretli S. The cumulative risk of a false-positive recall in the Norwegian breast cancer screening program. *Cancer* 2004;101:1501-7.

37. Bond M, Pavey T, Welch K, Cooper C, Garside R, Dean S, et al. Systematic review of the psychological consequences of false-positive screening mammograms. *Health Technol Assess* 2013;17:1-70.

38. Austoker J, Ong G. Written information needs of women who are recalled for further investigation of breast screening: Results of a multicentre study. *J Med Screen* 1994;1: 238–44.

39. Scaf-Klomp W, Sanderman R, van de Wiel HB, Otter R, van den Heuvel WJ. Distressed or relieved? Psychological side effects of breast cancer screening in The Netherlands. *J Epidemiol Com Health* 1997;51:705–10.

40. Sandin B, Chorot P, Valiente RM, Lostao L, Santed MA. Adverse psychological effects in women attending a second-stage breast cancer screening. *J Psychosom Res* 2002;52:303-09.

41. Ekeberg O, Skjauff H, Kåresen R. Screening for breast cancer is associated with a low degree of psychological distress. *Breast J* 2001;10:20-4.

42. Lampic C, Thurfjell E, Sjödén PO. The influence of a false-positive mammogram on a woman's subsequent behavior for detecting breast cancer. *Eur J Cancer* 2003;39:1730-7.

43. 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;23:292-300.

44. Chiarelli AM, Moravan V, Halapy E, Majpruz V, Mai V, Tatla RK. Effects of False-positive result and re-attendance in the Ontario Breast Screening Program. *J Med Screen* 2003;10:129-33.

45. Schou Bredal I, Kåresen R, Skaane P, Engelstad KS, Ekeberg Ø. Recall mammography and psychological distress. *Eur J Cancer* 2013;49:805-11.

46. Álamo-Junquera D, Murta-Nascimento C, Macià F, Baré M, Galcerán J, Ascunce N, et al. Effect of false-positive results on re-attendance at breast cancer screening programmes in Spain. *Eur J Pub Health* 2011;22:404-8.

47. Sim M, Prema Siva S, Ramli I, Fritschi L, Tresham J, et al. Effect of false-positive screening mammograms on rescreening in Western Australia. *Med J Aust* 2012;196:693-5.

48. Bowling A. Research Methods In Health. United Kingdom: Open University Press; 2009.

49. Munro B. Statistical methods for health care research. 5th ed. Philadelphia: Lippincott Williams and Wilkins; 2005.

50. Witek-Janusek L, Gabram S, Mathews HL. Psychologic stress, reduced NK cell activity, and cytokine dysregulation in women experiencing diagnostic breast biopsy. *Psychoneuroendocrinology* 2007;32:22-35.

51. Ong G, Austoker J, Michell M. Early rescreen/recall in the UK National Health Service breast screening programme: epidemiological data. *J Med Screen* 1998;5:146-55.

52. Hofvind S, Geller B, Vacek PM, Thoresen S, Skaane P. Using the European guidelines to evaluate the Norwegian Breast Cancer Screening Program. *Eur J Epidemiol* 2007;22:447-55.

53. Tosteson A, Fryback D, Hammond C, Hanna L, Grove M, et al. Consequences of false-positive screening mammograms. *JAMA Intern Med 2014;*174:954-61.

Table 1: Screening Performance Parameters Definitions and Acceptable Levels for Programmes in their First Screening Round

Screening Performance Parameters	Definition	Acceptable EU/UK Level
Recall Rate	The number of women who have to physically return to the screening unit for a technical recall or for further assessment as a proportion of all women who had a screening examination.	< 7
Technical Recall Rate	The number of women who have a repeat mammogram because of a technical inadequacy of the screening mammogram as a proportion of all women who had a screening examination.	< 3
Further Assessment Rate	The number of women undergoing additional diagnostic techniques (clinical breast examination, additional imaging and invasive investigations) to clarify the nature of a perceived abnormality detected at the screening examination, as a proportion of all women who had a screening examination.	< 7
Early Recall Rate	The number of women who are recalled for further assessment at an interval shorter than the normal interval (locally 3 years) as a proportion of all women who had a screening examination. Early recalls are performed when the initial further assessment examination leads to an equivocal diagnosis.	< 1
Breast Cancer Detection Rate	The number of in situ and invasive malignant breast lesions detected in a screening round per 1000 women. The acceptable European breast cancer detection rate for screening programmes in their first round, is 3 times the underlying, expected, breast cancer incidence rate in the absence of screening. The age standardised breast cancer incidence rate in 2008 was 0.72 per 1000 Maltese females.	≥2.16
Positive Predictive Value of Screening Test	The number of cancers detected as a proportion of the women undergoing further assessment (excluding technical recalls).	> 2.7

Table 2: Data Record Sheet 1 – Outcome of the Initial Screening Mammogram

	Age Group		Total
	51 to 55 years	**56 to 60 years**	
Number of Women Screened	6781	5634	12415
Outcome of the Screening test			
Negative	5623	4492	10115
Radiologist Recommended Technical Recalls	9	23	32
Actually Performed Technical Recalls	9	23	32
Radiologist Recommended Further Assessment	1149	1119	2268
Actually Performed Further Assessment	1143	1114	2257

Table 3: Data Record Sheet 2 – Investigations performed during Further Assessment Clinics

	Age Group		Total
	51 to 55 years	**56 to 60 years**	
Number of additional imaging	1142	1109	2251
Types of additional imaging			
Further views	907	874	1781
Ultrasound	567	614	1181
Clinical examination	8	15	23
Interventional procedures	213	214	427
Cytology			
Radiologist Recommended Cytology	3	9	12
Actually Performed Cytology	3	9	12
Core Biopsy			
Radiologist Recommended Core Biopsy	210	206	416
Actually Performed Core Biopsy	210	205	415

Table 4: Data Record Sheet 3 – Outcome of the Overall Screening Process and Timing of Breast Cancer Detection

	Age Group		Total
	51 to 55 years	56 to 60 years	
Outcome of screening process			
Negative Examinations	6678	5554	12232
Early Recalls following Further Assessment	28	28	56
Ductal Carcinoma in situ detected	10	17	27
Invasive Cancer Detected	87	57	144
Unknown Outcome	6	6	12
Timing of Breast Cancer Detection			
At Initial Screen	95	74	169
At Early Recall	2	0	2

Table 5: Summary of screening performance parameter rates of the Maltese first screening round

Screening Performance Parameters	51 to 55 years	56 to 60 years	Total Rates
Recall Rate (%)	17.07	20.27	18.53
Technical Repeat Rate (%)	0.13	0.41	0.26
Further Assessment Rate (%)	16.94	19.86	18.27
Early Recall Rate (%)	0.41	0.50	0.45
Breast Cancer Detection Rate (per 1000 women)	14.30	13.13	13.77
Positive Predictive Value of Screening Test (%)	8.49	6.64	7.58

Table 6: MNBSP Screening Performance Parameters Results compared to Acceptable EU/UK Levels

Screening Performance Parameters	MNBSP Results	Acceptable Level EU/UK Guidelines
Recall Rate (%)	18.53	< 7
Technical Repeat Rate (%)	0.26	< 3
Further Assessment Rate (%)	18.27	< 7
Early Recall Rate (%)	0.45	< 1
Breast Cancer Detection Rate (per 1000 women)	13.77	≥2.16
Positive Predictive Value of Screening Test (%)	7.58	> 2.7

Highlights

- The Maltese technical and early recall rates complied with European guidelines.

- Breast cancer detection rate and positive predictive value conformed to guidelines.

- The recall and further assessment rates exceeded recommended European guidelines.

- Effect of age on screening parameters is not as significant as the screening round.