Available online at www.sciencedirect.com

ScienceDirect

Journal of the Chinese Medical Association 77 (2014) 531–534

www.jcma-online.com

Original Article

The outcome of a quality-controlled mammography screening program: Experience from a population-based study in Taiwan

Huay-Ben Pan [a,b], Kam-Fai Wong [c], Tsung-Lung Yang [a,b], Giu-Cheng Hsu [d], Chen-Pin Chou [a,b], Jer-Shyung Huang [a], San-Kan Lee [e], Yi-Hong Chou [b,f], Chia-Ling Chiang [a], Huei-Lung Liang [a,b,*]

[a] *Department of Radiology, Kaohsiung Veterans General Hospital, Kaohsiung, Taiwan, ROC*
[b] *National Yang-Ming University School of Medicine, Taipei, Taiwan, ROC*
[c] *Institute of Statistics, National University of Kaohsiung, Kaohsiung, Taiwan, ROC*
[d] *Department of Radiology, Tri-Service General Hospital, Taipei, Taiwan, ROC*
[e] *Department of Radiology, Taichung Veterans General Hospital, Taichung, Taiwan, ROC*
[f] *Department of Radiology, Taipei Veterans General Hospital, Taipei, Taiwan, ROC*

Received November 19, 2013; accepted March 4, 2014

Abstract

Background: A greater policy of emphasis on the early detection and treatment of breast cancer is prevalent among developed countries. To raise the screening performance with a potentially decreased mortality rate, it is crucial to evaluate and analyze the screening outcome after implementation. We report the clinical outcome of an 8-year nationwide mammography screening in Taiwan to help share our statistical information on breast screening worldwide, especially in Asia.

Methods: Taiwan has provided nationwide, free, biennial mammographic screening since 2004. A total of 2,392,789 consecutive screening mammography examinations were performed during this study period for women aged 50–69 years (2006–2009) and 45–69 years (from December 2009 onwards). The screening covers 33.2% of the target population in the most recent 2 years. The workload of every screening radiologist, the overall recall rate, positive predictive value (PPV1), cancer detection rate (CDR), cancer incidence rate (CIR) from the screening, 1-year interval cancer, sensitivity, and specificity of the screening mammography are calculated, and compared with the American College of Radiology (ACR) recommendation level and/or those of other screening mammographic series.

Results: The CDRs (%) and CIRs (%) increased from 3.94–4.08 and 4.80–5.04 to 4.71–5.04 and 5.71 after 2009, implying a high occurrence of breast cancer in the younger age group of 45–49 years. The recall rates (9.3–10.0%) in this review are within the ACR recommendation range (<10%) and the PPV1 has also reached the ACR recommended level (>5%) in the most recent 2 years. The improvement of the screening performance may be attributed to our peer auditing review and education program. The sensitivity of our screening mammography is slightly lower than that of the ACR recommended level (>85%), which is still comparable to the results of the Vermont area in the USA. Although the workload (screenees/screeners) for every radiologist each year has increased from 150 in 2004 to 1360 in 2012, it does not seem to worsen the quality outcome of this screening program.

Conclusion: From the outcome review of this national mammography screening, there is still room to ameliorate our performance through comprehensive and continued education, to improve the competence of cancer detection and decrease false negative (FN) cases.

Keywords: breast screening; cancer detection rate; mammographic screening; medical audit; sensitivity

Conflicts of interest: The authors declare that there are no conflicts of interest related to the subject matter or materials discussed in this article.

* Corresponding author. Dr. Huei-Lung Liang, Department of Radiology, Kaohsiung Veterans General Hospital, 386, Ta-Chung 1st Road, Kaohsiung 813, Taiwan, ROC.

E-mail address: hlliang@vghks.gov.tw (H.-L. Liang).

http://dx.doi.org/10.1016/j.jcma.2014.06.013

1726-4901/

1. Introduction

Breast cancer is the most frequent cancer among women in both developing and developed countries.[1,2] Screening mammography contributes to the substantial decrease of 6% in breast cancer mortality, to 63% through early detection.[3,4] Mortality increases in regions without early detection programs.[5] Even for women in their 40s, screening mammography was reported to efficiently reduce breast cancer mortality by 29%.[6] Despite the fact that correlation of screening and mortality is still debatable,[7] a greater emphasis on the early detection and treatment of breast cancer is prevalent among developed countries.[8] To raise the screening performance and potentially decrease the mortality rate, it is crucial to evaluate and analyze the screening outcome after implementation.[9] Thereafter, we report the clinical outcome [including positive predictive value (PPV1), recall rate, cancer detection rate (CDR), early CDR, sensitivity, and specificity] of an 8-year nationwide mammography screening in Taiwan, to help share our statistical information on breast screening worldwide, especially in Asia.

2. Methods

2.1. Strategy of screening mammography

Since 2004, a nationwide, free, biennial screening mammography program has been offered for asymptomatic women aged 50−69 years (before December 2009) and aged 45−69 years thereafter. For the convenience of screenees and to increase the percentage coverage of participants, screening mobile vans have also been used since 2007. Four views of bilateral breasts in craniocaudal and mediolateral oblique projections are the standard of our screening mammography. Assessments are based on the Breast Imaging Reporting and Data Systems (BI-RADS) established by the American College of Radiology (ACR). In that system, categories 0, 4, and 5 are considered as positive assessments, representing either requiring further study or increasing risk of malignancy. Categories 1, 2, and 3 are considered as negative assessments, with requirement of a short follow up for the latter.[10]

To be qualified to attend the screening mammograms, all board-certified radiologists and radiographers are asked to attend a mammogram-interpretation or quality control education program yearly. The former are also required to have interpreted at least 1000 mammograms within 2 years.

2.2. Definitions[10]

True positive (TP) is tissue diagnosis of cancer within 1 year after a positive examination (BI-RADS Category 0, 4, or 5 for screening). False negative (FN) or interval cancer is tissue diagnosis of cancer within 1 year of a negative examination (BI-RADS Category 1 or 2 for screening, BI-RADS Category 1, 2, or 3 for diagnostic). Recall rate is reported as the percentage of positive interpretation (BI-RADS Category 0, 4, or 5; ACR recommended level to be <10%). PPV1 is the percentage of all positive screening examinations (BI-RADS Categories 0, 4, and 5) that result in a tissue diagnosis of cancer within 1 year (ACR recommended level to be >5%). CDR from screening is the number of cancers correctly detected/1000 patients examined on mammography (=number of TP/number of screening). Cancer incidence rate (CIR) from screening is the number of cancers diagnosed from screening mammography in 1 year [=(number of TP + number of FN)/number of screening]. Sensitivity [TP/(TP + FN)] is the probability of detecting a cancer when a cancer exists or the number of cancers diagnosed after being identified at mammography in a population within 1 year of the imaging examination, divided by all cancers present in that population in the same time period (ACR recommended level to be >85%). Specificity [TN/(TN + FP)] is the probability of interpreting an examination as negative when cancer does not exist, or the number of true negative mammograms in a population divided by all actual negative cases (those for which there is no tissue diagnosis of cancer within 1 year of the mammogram) in the population (ACR recommended level to be >90%).

2.3. Data collection

The study was approved by the Institutional Review Board at Kaohsiung Veterans General Hospital, Kaohsiung, Taiwan and was performed according to the Declaration of Helsinki principles. We collected data from the results of a free, nationwide population-based mammography-screening program funded and coordinated by Taiwan's Bureau of Health Promotion (received on May 2012) and National Cancer Registry. As the population of the participant women in the first 2 years (2004 and 2005) was small (<100,000) and/or was contaminated with symptomatic patients, we excluded the data from these 2 years for later analysis.

2.4. Statistical analysis

As the latest data (after 2010) of FN from the National Cancer Registry is lacking at the start of this study, we reported the numbers of screening, recall rate, PPV1, and CDR (TP) data from the years 2004−2012, and the CIRs, 1 year interval cancer (FN), sensitivity, and specificity from the years 2004 to 2010. Due to the size of the observations, the standard error was small, and thus there was little need to test statistical significance.

3. Results

Since the initiation of the screening program, the enrolled mammographic units in our country increased from 102 units in 2004 to 195 units in 2012, including 68 mobile vans. The digital mammographic system (CR or DR) was adopted at around 20%, initially to reach >95% in 2012 (163 DR, 86 CR, and 11 screen-film combinations). A total of 2,473,608 consecutive mammographic examinations for breast cancer screening were performed between 2004 and 2012 and 2,392,789 between 2006 and 2012. The target population had increased from 17,272 women in 2004 to 670,528 in 2012. In the years 2011 and 2012,

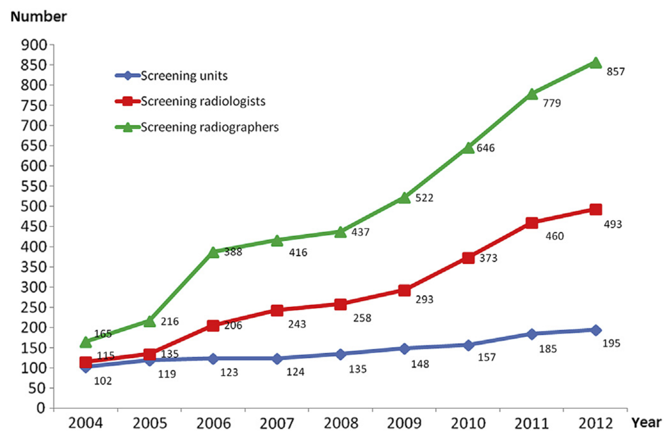

Fig. 1. The numbers of screening units, radiologists, and radiographers who participated in the screening mammography in Taiwan from 2004 to 2012.

the percentage of eligible participants (aged 45–69 years) receiving screening mammography was 15.3% and 17.9%, respectively. In other words, for a biennial screening program, the adjusted coverage had reached 33.2% of the target population in the most recent 2 years. For a corresponding service with increasing participants, the number of qualified screening radiographers and radiologists also progressively increased from 165 and 115, respectively, in 2004, to 857 and 493, respectively, in 2012 (Fig. 1). The presumed average workload (participant women/screening radiologists) for every radiologist was 150 in 2004, 486 in 2005, 512 in 2006, 521 in 2007, 611 in 2008, 843 in 2009, 1417 in 2010, 1215 in 2011, and 1360 in 2012.

In this study, the overall recall rates of screening mammography in Taiwan after 2006 are between 9.28 and 10.0, which is within the ACR recommendation level. The PPV1 values, albeit slightly <5% (4.10–4.88%) in the initial years, reached the ACR recommended range in the latest 2 years (5.00% and 5.16%). The CDRs and CIRs of the screening mammography were 3.94–4.39/1000 women and 4.80–5.04/1000 women, respectively, for participants aged 50–69 years before 2010; these increased to 4.71–5.04/1000 women and 5.71/1000 women following the enrollment of younger women (aged 45–49 years). The percentage of early cancer in this screening increased from 15.7% in 2004 to around 40% in these years. The 1-year interval cancer rates ranged from 0.73/1000 to 1.12/1000 screening women. The overall sensitivities in our screening mammography after 2006 ranged from 79.6% to 87.0%, and the specificities ranged from 90.5% to 91.1%. All of the above data are summarized in Table 1.

4. Discussion

The major goal of a screening mammography is to detect occult breast cancer with an acceptable range of recommendations for false positives, by raising both the CDR and sensitivity. The influencing factors of CDR vary and include, but are not limited to, social, economic, culture, race, and the improvement of diagnostic modalities.[11] In this retrospective study, we found a trend of increase in the CDR and CIR from the screening mammography after 2009, when younger women (45–49 years old) were enrolled. This may be due to the fact that breast cancer occurs in younger patients in Taiwan. Mammography is of benefit for cancer detection in young women with even higher breast density.

The recall rate in a screening examination is defined as the percentage of screening studies for which further work-up is recommended. A wide range of recall rates (1–15%) has been reported for screening mammography in the literature.[11,12] A high recall rate means that screening program resources are probably used inefficiently and women undergo unnecessary follow-up procedures, although a low recall rate can potentially result in lower rates of detecting incident breast cancers. The recall rates of screening mammography in Taiwan since 2005 are between 9.28% and 10.0%, which is within the ACR recommended range. Moreover, through the peer audit and education programs, we find some extreme cases (i.e., recall rate >15% or <5%) have been gradually eliminated in recent years.[12]

Recent publications have also used PPV1 as one of the primary indicators of the quality of screening mammography programs.[11,13,14] PPV1 has been generally thought to represent the "necessary balance" between sensitivity and specificity. Low PPV1 usually indicates a high percentage of unnecessary recall, which increases additional medical cost and a patient's anxiety. However, a high PPV1 may still miss a certain number of potentially detectable cancers if sensitivity is not incorporated into the analysis. The PPV1 in our series reached the ACR recommendation level of 5% in recent years.

Sensitivity is simplified as the probability of detecting a cancer when a cancer exists on the film. Hence, high quality

Table 1
The clinical outcome of the screening mammography in Taiwan from 2004 to 2012.

Year	RR (%)	PPV1 (%)	CDR (%)	ECDR (%)	CIR (%)	1 y ICR (%)	Sensitivity (%)	Specificity (%)
2004	10.3	5.2	5.38	15.7	6.25	0.97	86.1	90.1
2005	11.9	3.4	4.02	20.2	4.91	1.01	81.9	88.4
2006	9.8	4.1	4.03	18.3	4.80	0.85	84.0	90.6
2007	9.3	4.4	4.08	19.8	4.90	0.91	83.2	91.1
2008	9.9	4.4	4.39	30.3	5.04	0.73	87.0	90.5
2009	9.6	4.1	3.94	36.4	4.95	1.12	79.6	90.7
2010	9.7	4.9	4.71	41.7	5.71	1.1	82.6	90.8
2011	10.0	5.0	5.04	39.7	—	—	—	—
2012	9.6	5.2	4.96	39.9	—	—	—	—

CDR = cancer detection rate; CIR = cancer incidence rate from screening; ECDR = early cancer detection rate; FN = false negative; PPV1 = number of TP/number of recall; RR = recall rate; TP = true positive; 1 y ICR = 1st year interval cancer rate.

examination and high competence of film interpreters that avoid FNs contribute to identifying the existing cancer. In a recent study, Hofvind et al[15] reported that the sensitivity of final assessment for a 2-year follow-up was 83.3% for a single reading in Vermont (USA) and 90.7% for a double reading in Norway with the interval cancer (FN) of 1-year follow-up of 0.8/1000 and 1.2/1000 screening women, respectively. The sensitivity and 1-year interval cancer in our series are 79.6−87.0% and 0.73−1.12/1000 screening women, respectively. This is comparable with the results of Vermont area for a single reading study.

The workload (screening women/screening radiologists) of the screening program in Taiwan has increased from 150 in 2004 to >1200 in 2010−2012. The increased workload does not seem to worsen the screening quality, including the indicators of recall rate, PPV1, CDR, incidence rate, sensitivity, and specificity. The increase of early CDRs is partly due to the reduction of contamination of symptomatic patients which occurred in the early years. It is also due to the improvement of interpretation accuracy, which may be attributed to more radiologists willing to be full-time breast imaging professionals in recent years in Taiwan, and/or the successful education programs feedback from the peer auditing system.[13] To achieve higher sensitivity while lowering false positive rates, further studies are needed to elucidate the interrelationships between training, experience, workload, and performance measures.[16]

The main strength of the present study is to assess the screening mammography outcome by the use of a national representative sample. Limitations of this study include several variables that may affect the result, e.g., age, breast density, asymptomatic or symptomatic, and prevalence (initial screen) or incidence (subsequent screen). Risk factors have not been analyzed due to data limitation provided from the Bureau of Health Promotion. Further analyses are needed to better interpret the results and trends of mammography screening.[9]

In conclusion, the medical audit data of screening mammography in Taiwan in recent years has reached the recommendation of ACR levels, except the screening sensitivity, which is slightly lower but comparable with that of Vermont's area in the USA. From the outcome review of this national mammography screening, there is still room to ameliorate our performance through comprehensive and continued education, to improve the competence of cancer detection and decrease FN cases.

Acknowledgments

This work was supported by the Bureau of Health promotion, Department of Health, (grant number 9910027C). We thank participating mammography facilities, and radiologists for the data they provided for this study.

References

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 2010;**127**:2893−917.
2. http://www.bhp.doh.gov.tw/BHPNet/Web/News/News.aspx?No=201304300001. Attachment.doc [in Chinese].
3. Aubard Y, Genet D, Eyraud JL, Clavère P, Tubiana-Mathieu N, Philippe HJ. Impact of screening on breast cancer detection. Retrospective comparative study of two periods ten years apart. *Eur J Gynaecol Oncol* 2002;**23**:37−41.
4. Otto SJ, Fracheboud J, Looman CW, Broeders MJ, Boer R, Hendriks JH, et al. Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systematic review. *Lancet* 2003;**361**:1411−7.
5. Smith RA, Caleffi M, Albert US, Chen TH, Duffy SW, Franceschi D, et al. Global Summit Early Detection and Access to Care Panel. Breast cancer in limited-resource countries: early detection and access to care. *Breast J* 2006;**12**(Suppl 1):S16−26.
6. 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**:714−22.
7. Tannenbaum SL, Koru-Sengul T, Miao F, Byrne MM. Disparities in survival after female breast cancer diagnosis: a population-based study. *Cancer Causes Control* 2013;**24**:1705−15.
8. Chen BK, Yang CY. Temporal trend analysis of avoidable mortality in Taiwan, 1971-2008: overall progress, with areas for further medical or public health investment. *BMC Public Health* 2013;**13**:551.
9. Giordano L, Giorgi D, Ventura L, Castagno R, Paci E, Segnan N. Time trends of process and impact indicators in Italian breast screening programmes (1999−2009). *Epidemiol Prev* 2011;**35**(5−6 Suppl 5):28−38.
10. 2003 ACR *BI-RADS®* *Mammography*, 4th ed. www.acr.org/~/media/.../BIRADS/MammoFollowUp.pdf.
11. Elmore JG, Nakano CY, Koepsell TD, Desnick LM, D'Orsi CJ, Ransohoff DF. International variation in screening mammography interpretations in community-based programs. *J Natl Cancer Inst* 2003;**95**:1384−93.
12. Pan HB, Hsu GC, Yang TL, Huang JS, Chou CP, Liang HL, et al. Peer reviewing of screening mammography in Taiwan: its reliability and the improvement. *Chin Med J (Engl)* 2013;**126**:68−71.
13. 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.
14. Blanks RG, Moss SM, Wallis MG. Monitoring and evaluating the UK National Health Service Breast Screening Programme: evaluating the variation in radiological performance between individual programmes using PPV-referral diagrams. *J Med Screen* 2001;**8**:24−8.
15. Hofvind S, Geller BM, Skelly J, Vacek PM. Sensitivity and specificity of mammographic screening as practised in Vermont and Norway. *Br J Radiol* 2012;**85**:e1226−32.
16. Carney PA, Sickles EA, Monsees BS, Bassett LW, Brenner RJ, Feig SA, et al. Identifying minimally acceptable interpretive performance criteria for screening mammography. *Radiology* 2010;**255**:354−61.