European Journal of Radiology 101 (2018) 24–29

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Research article

Combined screening with mammography and ultrasound in a population-based screening program

Wolfgang Buchberger[a,b], Sabine Geiger-Gritsch[b,c], Rudolf Knapp[d], Kurt Gautsch[e], Willi Oberaigner[b,c,*]

[a] Medical Department, Tirol Kliniken Ltd., Anichstr. 35, 6020 Innsbruck, Austria
[b] Department of Public Health, Health Services Research and Health Technology Assessment, UMIT – University for Health Sciences, Medical Informatics and Technology, Eduard Wallnöfer-Zentrum 1, 6060 Hall i.T., Austria
[c] Department of Clinical Epidemiology of Tirol Kliniken Ltd., Anichstr. 35, 6020 Innsbruck, Austria
[d] Department of Radiology, Kufstein County Hospital, Endachstr. 27, 6330 Kufstein, Austria
[e] Department of Radiology, Innsbruck Medical University, Anichstr. 35, 6020 Innsbruck, Austria

ARTICLE INFO

Keywords:
Breast cancer screening
Mammography
Ultrasound
Breast density
Sensitivity
Specificity

ABSTRACT

Objectives: To compare the performance of screening with mammography combined with ultrasound versus mammography alone in women at average risk for breast cancer.

Methods: 66,680 women underwent physician-performed ultrasound as an adjunct to screening mammography. Histological results and follow-up at one year were used as reference standard for sensitivity. Main outcome measures were cancer detection rate, sensitivity, recall rate, biopsy rate, and positive predictive value of biopsy for combined screening with mammography plus ultrasound versus mammography alone.

Results: The overall sensitivity of mammography only was 61.5% in women with dense breasts and 86.6% in women with non-dense breasts. The sensitivity of mammography plus ultrasound combined was 81.3% in women with dense breasts and 95.0% in women with non-dense breasts.

Adjunctive ultrasound increased the recall rate from 10.5 to 16.5 per 1000 women screened, and increased the biopsy rate from 6.3 to 9.3 per 1000 women screened. The positive predictive value of biopsy was 55.5% (95% CI 50.6%–60.3%) for mammography alone and 43.3 (95% CI 39.4%–47.3%) for combined mammography plus ultrasound.

Conclusions: Supplemental ultrasound improves cancer detection in screening of women at average risk for breast cancer. Recall rates and biopsy rates can be kept within acceptable limits.

1. Introduction

Mammography is the only screening modality that has been proved to reduce breast cancer mortality [1]. However, its ability to depict small non-calcified carcinomas varies greatly with breast tissue composition. While mammography detects up to 98% of carcinomas in fatty breasts, sensitivity declines significantly with increasing breast density and may be as low as 30%–48% in extremely dense breasts [2,3]. Dense breast tissue is a frequent finding in all age groups. In a recent analysis of more than 7000 screening mammograms, heterogeneously dense or extremely dense breast tissue was found in 74% of women 40–49 years old, in 57% of women 50–59 years old, and in 44% of women 60–59 years old [4]. Hence, improving cancer detection in dense breasts is crucial for increasing the effectiveness of mammography screening.

Ultrasound is a promising adjunctive screening modality, because it is widely available, relatively inexpensive, and well tolerated by patients. In addition, suspicious breast lesions can be readily biopsied under ultrasound guidance. Data from various single-center cohort studies, see for example [3,5–8] and one multi-center study suggest that the addition of hand-held ultrasound screening to mammography for women with dense breasts increases cancer detection rates at the expense of lower specificity and lower positive predictive values. However, these results were obtained in a few specialized institutions with relatively small numbers of participants. In addition, most studies included a significant proportion of women at elevated risk for breast cancer. Therefore, the results cannot be directly applied to women participating in a population-based screening program.

The Tyrolean Breast Cancer Screening Program implemented

* Corresponding author at: Department of Public Health, Health Services Research and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Eduard Wallnöfer-Zentrum 1, 6060 Hall i.T., Austria.
E-mail address: Wilhelm.oberaigner@umit.at (W. Oberaigner).

https://doi.org/10.1016/j.ejrad.2018.01.022
Received 25 October 2017; Received in revised form 20 January 2018; Accepted 22 January 2018

ultrasound as a second-line screening procedure in addition to mammography. The purpose of this retrospective analysis was to evaluate the benefits and adverse effects of combined screening with mammography and ultrasound compared to mammography alone in the framework of a population-based screening program.

2. Methods

2.1. Screening program

The Tyrolean breast cancer screening program started in Innsbruck and the surrounding area in June 2007 and was rolled out to the whole state of Tyrol in June 2008. The target population included all women aged 40–69, who lived in Tyrol and were covered by compulsory social insurance, which is about 98% of the total population. Women aged 40–59 were invited annually, and women aged 60–69 were invited biannually. Women with clinical symptoms, a personal or first-degree family history of breast cancer, or a previous high-risk biopsy result were excluded from routine screening, and were referred for diagnostic mammography or high risk screening. Screening was offered in 22 screening units; thirteen of them were run by radiologists in private practice and nine by hospital outpatient departments. Detailed information about the screening program was published earlier [9,10]. In a prior study [10] we reported on 42,834 women screened, including 32,322 women who underwent both mammography and ultrasound. The current study expands on this by having a larger patient number, and including new analyses of performance characteristics of mammography alone versus mammography plus ultrasound.

Two-view digital mammograms were obtained either with direct radiography or with computed radiography. All mammograms were read by specially trained radiologists who had been certified by the Austrian medical association. Mammography results were recorded according to the BI-RADS classification scheme [11]. Breast density was visually assessed by the radiologist interpreting the mammogram, namely according to the following BI-RADS density categories: grade 1 (fatty); grade 2 (scattered fibroglandular elements); grade 3 (heterogeneously dense); grade 4 (extremely dense) [11]. The mammograms were read immediately at the time of study, not batch read, so that a decision for supplemental ultrasound could be made. For organisational reasons, double reading was not performed.

According to the screening policy, all women with heterogeneously dense or extremely dense breasts underwent supplemental ultrasound. Ultrasound could also be performed in women with non-dense breasts per specific radiologist's request. The ultrasound examinations included both breasts and were performed during the same visit and by the same radiologist who read the mammograms. The reports were recorded according to the BI-RADS classification scheme [12].

After both examinations a final BI-RADS category was reported. Possible changes of diagnostic classification after ultrasound screening were handled at the judgement of each site. Women with BI-RADS 1 or 2 went back to routine screening. Women with BI-RADS 3 were invited for intermediate screening in six months, and women with BI-RADS 0, 4 or 5 were referred for further assessment. Assessment was offered at nine hospital radiology departments and included additional mammographic views, ultrasound, MRI, or core-needle biopsy, as needed.

All screening units and assessment units registered relevant data in a database. Screening and assessment information was electronically transferred to a central screening database at the Department of Clinical Epidemiology of Tirol Kliniken Ltd. after pseudonymising the woman's social insurance number. For reasons of data privacy restrictions, women had to sign a written consent to permit data transfer to the screening database. Linkage of cancer registry data with the screening database provided information on tumour characteristics (e.g., tumour size and stage) and therapy, and permitted identification of interval cancers [9].

2.2. Study population

From June 2008 to May 2010, 136,335 women were invited for mammography screening, and a total of 87,455 screening examinations were performed. 67,265 women underwent both mammography and bilateral breast ultrasound screening. 585 cases were excluded because of missing or incomplete data; in total, 66,680 cases were included in the final analysis.

The study was conducted in conformity with the Helsinki Declaration [13]. The Ethics Committee of the Medical University of lnnsbruck confirmed in writing that no formal ethics committee approval was required for this retrospective analysis.

2.3. Ascertainment of interval cancers

An interval cancer is by definition a primary breast cancer diagnosed in a woman who had a screening test, with/without further assessment, which was negative for malignancy [15]. All women included in the final analysis have residence in Tyrol. In addition, all incident breast cancer cases diagnosed in women living in Tyrol are registered in the cancer registry of Tyrol (CRT) with a completeness of $\geq 97\%$, independently from attending the screening program [14]. We ascertained interval cancers by linking incident breast cancer cases to women who have attended the screening program and were classified negative for malignancy in the time period of up to 356 days before the date of breast cancer diagnosis.

2.4. Data analysis

The study is based on observational data from a population-based screening program. Therefore, no sample size considerations were applied. The primary unit of analysis was the participant, with the breast imaging assessment on mammography alone or on mammography plus ultrasound. A BI-RADS assessment of 0, 3, 4 or 5 was considered positive, and a BI-RADS assessment of 1 or 2 was considered negative, which is a common rule in the assessment of screening outcomes [12].

Histological results obtained from the screening database and follow-up at one year were used as reference standard. Follow-up information was obtained by linking the screening dataset to all breast cancer cases collected in the Cancer Registry of Tyrol. Completeness of registration of cancer cases in the cancer registry in general and of breast cancer cases in particular has been shown to be high [14].

Absence of a recorded diagnosis of breast cancer in the screening database or in the cancer registry within 356 days after screening was considered disease-negative.

Cancer detection rate (i.e., the number of pathologically proven malignant lesions detected in screening per 1000 women screened), recall rate (i.e., the rate of referral for further assessment after initial screening), intermediate mammogram rate (i.e., the rate of intermediate mammogram in six months), biopsy rate (i.e., the rate of biopsies performed in screening positive cases), and positive predictive values (PPV) for screening, referral for assessment, and biopsy were calculated for combined screening with mammography and ultrasound based on the final recommendation after both modalities.

For mammography alone, recall rates and positive predictive values were calculated assuming that cases with BI-RADS 1, 2 would have returned to routine screening, cases with BI-RADS 3 would have undergone intermediate mammogram in six months, and cases with BI-RADS 0, 4 and 5 would have been recalled for further assessment.

In order to compare differences in the performance characteristics of mammography only and mammography combined with ultrasound, we applied the McNemar test in order to account for the natural pairing of tests within a participant. Statistical significance was established at $P < 0.05$. All data handling and statistical analyses were performed with STATA Version 13 [16].