IRMA SAARENMAA

Determinants of Sensitivity of Mammography

■

ACADEMIC DISSERTATION

To be presented, with the permission of
the Faculty of Medicine of the University of Tampere,
for public discussion in the auditorium of
Tampere School of Public Health of the University of Tampere,
Medisiinarinkatu 3, Tampere, on October 26th, 2001, at 12 o'clock.

Acta Universitatis Tamperensis 840
University of Tampere
Tampere 2001

ACADEMIC DISSERTATION
University of Tampere, School of Public Health
Finland

Supervised by
Professor Matti Hakama
University of Tampere

Reviewed by
Docent Ahti Anttila
University of Tampere
Docent Tapani Tikkakoski
University of Oulu

Distribution

University of Tampere
Sales Office
P.O. Box 617
33101 Tampere
Finland

Tel. +358 3 215 6055
Fax +358 3 215 7685
taju@uta.fi
http://granum.uta.fi

Cover design by
Juha Siro

Printed dissertation
Acta Universitatis Tamperensis 840
ISBN 951-44-5182-1
ISSN 1455-1616

Electronic dissertation
Acta Electronica Universitatis Tamperensis 135
ISBN 951-44-5183-X
ISSN 1456-954X
http://acta.uta.fi

Tampereen yliopistopaino Oy Juvenes Print
Tampere 2001

Contents

List of original publications

This dissertation is based on the following original publications, which are referred to by Roman numerals in the text.

I Saarenmaa I, Salminen T, Geiger U, Heikkinen P, Hyvärinen S, Isola J, Kataja V, Kokko M-L, Kokko R, Kumpulainen E, Kärkkäinen A, Pakkanen J, Peltonen P, Piironen A, Salo A, Talviala M-L and Hakama M (2001): Validity of radiological examinations of patients with breast cancer in different age groups in a population based study. Breast 10:78–81.

II Saarenmaa I, Salminen T, Geiger U, Heikkinen P, Hyvärinen S, Isola J, Kataja V, Kokko M-L, Kokko R, Kumpulainen E, Kärkkäinen A, Pakkanen J, Peltonen P, Piironen A, Salo A, Talviala M-L and Hakama M (2001): The effect of age and density of the breast on the sensitivity of breast cancer diagnostic by mammography and ultrasonography. Breast Cancer Res Treat 67:117–123.

III Saarenmaa I, Salminen T, Geiger U, Heikkinen P, Hyvärinen S, Isola J, Kataja V, Kokko M-L, Kokko R, Kumpulainen E, Kärkkäinen A, Pakkanen J, Peltonen P, Piironen A, Salo A, Talviala M-L and Hakama M (2001): The visibility of cancer on previous mammograms in retrospective review. Clin Radiol 56:40–43.

IV Saarenmaa I, Salminen T, Geiger U, Holli K, Isola J, Kärkkäinen A, Pakkanen J, Piironen A, Salo A and Hakama M (1999): The visibility of cancer on earlier mammograms in a population based mammography screening programme: Eur J Cancer 35:1118–1122.

In addition the thesis includes previously unpublished data.

Abbreviations

BSE = breast self examination
CI = confidence interval
CNB = core needle biopsy
DCIS = ductal carcinoma in situ
FNAB = fine needle aspiration biopsy
FNAC = fine needle aspiration cytology
HRT = hormone replacement therapy
MR = magnetic resonance imaging
OR = odds ratio
TIS = tumour in situ
US = ultrasonography

1 Introduction

Breast cancer is the most common cancer in Finland (Finnish Cancer Registry 2000) and worldwide the most common cancer in women (Kelsey and Bernstein 1996). Every year new breast cancer is diagnosed in over 3 400 Finnish women and 10–20 men (Finnish Cancer Registry 2000). Over 800 patients die yearly from breast cancer. However, the survival of breast cancer patients has been improving in the last four decades (Dickman et al. 1999). Early diagnosis is one means to lower breast cancer mortality. Breast self examination (BSE) has been shown to lower breast cancer mortality among compliers in one study in Finland (Gästrin 1994). In the UK Trial of Early Detection of Breast Cancer (TEDBC) the effect of mammography screening and education about breast self-examination was studied in a non-randomized setting. No reduction of breast cancer mortality was observed due to breast self-examination (Moss et al. 1999). Some other studies have also failed to show clear evidence to either confirm or deny the efficacy of BSE as an intervention to reduce breast cancer mortality (Auvinen et al. 1996, Hider and Nicholas 1999). However, BSE is a means to make women more "breast aware" (Austoker 1994) and perharps to avoid tumour growth to inoperable size. There is also little evidence of the effect of clinical breast examination alone on breast cancer mortality (Barton et al. 1999).

Mammography screening has been shown in many studies to decrease breast cancer mortality at least in women older than 50 years (Shapiro et al. 1982, Eddy 1989, Duffy et al. 1991, Nyström et al. 1993, de Koning et al. 1995, Kerlikowske et al. 1995, Larsson et al. 1996, van Dijk et al. 1996, Frisell et al. 1997, Tabar et al. 1997, Alexander et al. 1999, Given-Wilson et al. 1999, Blanks et al. 2000, de Koning 2000). The effectiveness of screening for breast cancer as a public health policy was demonstrated in Finland (Hakama et al. 1997). There are some trials which also show an effect on breast cancer mortality for younger women (Kopans et al. 1996, Thurfjell and Lindgren 1996, Bjurstam et al. 1997, Kopans 1997a, UK Trial of Early Detection of Breast Cancer Group 1999, Andersson 2000).

Tumours are found clinically in women not covered by or attending the screening programme. Clinical mammography also has major importance in detecting breast cancer when small and curable. Even if there is a palpable mass or verified breast cancer in the breast, it is important to use mammography to screen the remainder of both breasts for nonpalpable cancer (Robidoux et al. 1997, Rosen et al. 1999). Ultrasonography has a major role as a complementary examination to mammography (Moss et al. 1999, Zonderland et al. 1999). There may be considerable interobserver variation in the interpretation of both mammography and US (Skaane et. al. 1997, Skaane et al. 1999), thereby having a noteworthy effect on the sensitivity of these examinations. It is a well-

known fact in radiology that some originally undetected tumours are visible in retrospective review of mammograms (Harvey et al. 1993, Jones et al. 1996). These tumours may be of great medical-legal significance (Berlin 2001). The proportion reported of such tumours varies between studies.

With fine needle biopsy the cytology and with core needle biopsy the histology of the tumour can be verified before open biopsy. The sensitivity of these examinations determines how small the tumour is when found and how effective the diagnostic procedure is.

The purpose of this thesis is to evaluate the sensitivity of breast cancer diagnostics as well as the proportion of restrospectively visible tumours.

2 Review of the literature

2.1 Mammographic appearance of breast cancer versus benign breast abnormalities

Benign and malignant tumors are shown in mammography mainly as densities or microcalcifications. Breast cancer is a rare condition in male population. Mammographically the same characteristics are observed in men with breast cancer as in women (Barth and Prechtel 1991, Stewart et al. 1997), but the smaller breast size makes mammography imaging technically difficult and also complicates the interpretation of mammograms.

Densities

Dense, irregular mass with spiculated margin (stellate lesion) can be regarded as typical for malignancy. A postoperative scar may give almost a similar picture and sometimes differential diagnosis between scar and recurrent tumour may cause difficulties (Stomper et al. 1988, Kopans 1998). Radial scar and fat necrosis may also appear like a stellate lesion in mammography. New or growing masses are malignant suspect, but also some benign abnormalities, like fibroadenomas may appear or grow (Kopans 1998). Lesions with ill-defined or microlobulated margins and density increasing over time should arouse suspicion of malignancy (Kopans 1998). Halo-sign, a narrow radiolucent ring or a segment of a ring aroud the periphery of a lesion, is typical for a benign lesion with three exceptions: intracystic carcinoma, papillary carcinoma and carcinoma in a fibroadenoma may have a halo-sign (Tabar and Dean 1983). Solitary sharp circumscribed low density mass is probably benign. Mammography, however, cannot differentiate, for example, a cyst from a solid lesion and mammography can rarely be relied on to make a conclusive diagnosis of a benign process (Kopans 1998).

Microcalcifications

Casting type ductal calcifications are built of fine, linear, branching calcification fragments, that differ in density, length and outline and form the most typical and reliable sign of intraductal carcinoma (Tabar and Dean 1983). Clustered tiny, dotlike innumerable and irregularly grouped calcifications resembling fine grains of salt or

powdery calcifications are also malignant suspect (Tabar and Dean 1983). In differential diagnosis of benign and malignant calcification the form of the microcalcification cluster is also helpful and a triangular shape helps to differentiate comedocarcinoma from a calcifying fibroadenoma (Lanay 1986). Lesions with changing calcifications should arouse suspicion of malignancy (Kopans 1998). Round, regular clustered calcifications are probably benign.

Indirect signs

New architectural distortion and a distorted edge should arouse suspicion of malignancy. Asymmetric breast tissue, ducts or veins, skin and trabecular thickening, nipple retraction, deviation or inversion and enlarged axillary lymph nodes support the possibility of malignancy (Kopans 1998).

Mammographic-histologic correlation

Mammography finding of breast cancer depends on the histological subtype of the tumour (Le Gal et al. 1992, Newstead et al. 1992). There are three major subtypes of noninvasive, or in situ carcinoma (Rubin and Farber 1988). The most common of these is intraductal carcinoma or ductal carcinoma in situ (DCIS). Microcalcifications without or with a tumour mass are the most important mammographic findings of DCIS (Dershaw et al. 1989). Lobular carcinoma in situ has no specific mammography finding (Sonnenfeld et al. 1991) and it is found periodically when a palpable mass or benign mammographic abnormality is removed. Intraductal papillary carcinoma is a rare condition (Rubin and Farber 1988).

Almost 90% of all invasive breast cancers are of the invasive ductal type (Rubin and Farber 1988). Invasive ductal carcinoma appears in mammography as spiculated or ill-defined mass with or without microcalcifications or as microcalcifications only (Newstead et al. 1992, Rothman et al. 1996).

Because of the diffuse growth pattern and tendency to form lesions with opacity equal to or less that of the parenchyma, invasive lobular carcinoma can be extremely difficult to detect mammographically (Hilleren et al. 1991).

Medullary carcinoma is usually an uncalcified mass with indistinct or circumscribed borders in mammography (Liberman et al. 1996), and it may mimic a benign mass both in mammography and US (Meyer et al. 1989). Mucinous carcinoma may also resemble a benign mass and appears as a well-circumscribed, round or oval,

"soft" density in mammography. Tubular carcinoma shows a small stellate density in mammography.

Mammography techniques and interpretation

Special mammography views (spot compression-magnification and tangential views) may help in differentiating benign from malignant lesions. Faulk and Sickles (1992) evaluated the role of special mammographic views in examining palpable breast masses. Special views depicted 9% of palpable masses not seen on standard views (cranio-caudal and mediolateral or mediolateral oblique views). Special views allowed correct prediction of benign or malignant status in 77% of cases while standard views allowed correct prediction in 69% of cases. Berkowitz et al. (1989) found 8 (11%) lesions of 75 to be more suspicious on spot compression views than in standard projections and all of them were cancers on biopsy. Microfocus magnification mammogram is essential by analysing microcalcifications (Tabar and Dean 1983).

Mammography interpretation varies among individuals. Linver et al. (1992) concluded that dedicated mammography courses can help improve radiologists' performance and alter their interpretive approach. However, even with best technical quality of mammograms and skills of the radiologist not all breast cancers are visible in mammography. Coveney et al. (1994) reported 16.5% false negative mammograms in patients with palpable breast cancer.

2.2 Breast density

On the basis of its mammographic appearance, the breast can be considered to have two major components: fibroglandular tissue and fat. Fibroglandular tissue is a mixture of fibrous connective tissue (the stroma) and glandular tissue (epithelial cells that line the ducts). Fat is more radiolucent than fibroglandular tissue and appears darker in mammography. Regions of brightness associated with fibroglandular tissue are referred to as mammograpic density (Byng et al. 1998). The breast size and density of woman varies during lifetime and those differ considerably among individuals (Warner et al. 1992). In the breast there are both translucent and dense areas to an individually varying degree (Barth and Prechtel 1991).

The breast parechymal patterns or the breast density have been classified in different ways in different numbers of categories. In the 1970’s Wolfe (1976a, 1976b) classified breast into four density classes by visual pattern reading. Breast parenchymal

density has also been classified in percentages (Wolfe et al. 1986). Threatt et al. (1980) classified breast into three major density classes: lucent, intermediate and dense. Tabar classified breast parenchymal patterns into five groups based on anatomic-mammographic correlations, following three-dimensional (thick slice technique) histopathologic-mammographic comparisons (Gram et al. 1997). To evaluate the association between mammographic density and breast cancer risk, a simple, observer-assisted technique called interactive thresholding was developed. This allows reliable quantitative assessment of mammographic density from digitized mammograms using a computer workstation (Byng et al 1998, van Gils 1998).

The age of the patient influences breast parenchymal density significantly. Stomper et al. (1996) found in the age cohort 25 through 29 years old 38% predominantly fatty breasts compared to age cohort 75–79 years old with 76% fatty breasts. Salminen et al. (1998b) found incidence of fatty breast to increase with increasing age in a population based follow-up study of mammography screening. Incidence of fatty breast was almost two times higher among women aged 45 or more compared to women under 45.

Hormone contraceptives may reduce (Spicer et al. 1994) and hormone replacement therapy (HRT) may increase mammographic density (Berkowitz et al. 1990, Stomper et al. 1990, Kaufman et al. 1991, McNicholas et al. 1994, Laya et al. 1995, Persson et al. 1997, Salminen et al. 1999, Schairer et al. 2000). HRT may increase the size of fibroadenomas and cysts (Cyrlak and Wong 1993), the recall rate in screening (Litherland et al. 1997) and the amount of interval cancers (Litherland et al. 1999). In 1988–1993 in the West of Scotland 12% of women with screen-detected cancers and 22% of women with interval cancers were using HRT. The relative risk of an interval cancer arising in the first year after screening for a woman on HRT was over twofold compared to women without HRT (Litherland et al. 1999). Dense breast parenchymal pattern may be a contraindication for HRT (Salminen et al. 2000, Schairer et al. 2000), but there is no total agreement on the real influence of the risks and benefits of HRT. Further evaluation is needed (Colditz et al. 1990, Thurfjell et al. 1997, Colditz 1998).

Ever since Wolfe's publication there has been debate about the significance of breast parenchymal density. It is still unclear which patterns represent the normal or which, if any, are abnormal (Kopans 1998). In 1976 Wolfe suggested that the tissue pattern of the breast, as seen in mammography, might predict a risk for developing breast cancer. In 1986 Wolfe suggested that women with mammographic densities in at least 25% of the breast have a significantly elevated risk of 4.3 to 5.5 compared with women with less than 25% density (Wolfe et al. 1986). Besides him, many other authors suggest that dense breast parenchyma increases the risk of breast cancer. Salminen et al.

(1998a) found the age adjusted relative risk of breast cancer to be 2.5 among women with high-risk mammographic parenchymal patterns (dense to relatively dense) at the screening preceding cancer diagnosis compared with low-risk patterns (fatty to relatively fatty). Byng et al. (1998) suggested that mammographic density is one of the strongest risk factors for breast cancer. Some authors suggest that these patterns do not constitute a true risk factor, rather the effect is caused by the greater difficulty of detecting breast cancer in the dense patterns compared with the fatty patterns (Ma et al. 1992, Bird et al. 1992, Sala et al. 1998, van Gils et al. 1998). Whitehead et al. (1985) found both theories to be true; dense breast parenchyma did have a masking effect, but it was also a breast cancer risk factor. Boyd et al. (1998) pointed out in their review article, that studies using quantitative measurement to classify mammographic patterns have consistently found that women with dense tissue in more than 60–75% of the breast are at four to six times greater risk of breast cancer than those with no densities. However, Tabar concluded that even though there is a significantly higher risk of breast cancer in so called high-risk breasts (dense to relatively dense), more than 72% of the cancers were found in so-called low-risk breasts (fatty to relatively fatty) (Tabar and Dean 1982). There is no such combination of breast cancer risk factors which could be used in selective screening of breast cancer (Soini and Hakama 1978).

2.3 Effectiveness of mammography based breast cancer screening

There are a number of randomized and non-randomized studies of mammography screening for breast cancer for which mortality data have been published (Shapiro et al. 1982, Andersson et al. 1988, Eddy 1989, Roberts et al. 1990, Duffy et al. 1991, Nyström et al. 1993, de Koning et al. 1995, Kerlikowske et al. 1995, Larsson et al. 1996, van Dijck et al. 1996, Frisell et al. 1997, Tabar et al. 1997, Alexander et al. 1999, Given-Wilson et al. 1999, Blanks et al. 2000). The results of the studies are difficult to compare because of differences in their settings, attendances, penetrations of intervention into the control group, study design and the number of women years at follow-up. Despite differences in the design and settings between these studies, it seems reasonable to conclude that mammography based breast cancer screening results in significant reduction in breast cancer mortality at least for women over 50 years (Kopans 1997b). The studies demonstrated an average reduction of about 30% among women who were offered screening. Recently Tabar et al. (2001) published results from organized mammography screening in a clinical setting in the two county study in Sweden. The breast cancer mortality of women who were actually screened during the period 1988-1996 were compared with breast cancer mortality during the time period

when no screening was available (i.e. 1968–1977) and 63% mortality reduction was observed. No significant change in breast carcinoma mortality was observed either among young women (i.e. aged 20–39) or among the invited non-attenders in the 40–69 year age group.

There has been debate over the drawbacks and benefits of screening for over 30 years (Dilhuydy and Barreau 1997). Recently there has been debate over the effectiveness of mammography screening, especially because of the article by Gotzsche and Olsen in spring 2000. They concluded that mammography screening does not reduce breast cancer mortality. They based their criticism on the assumptions, that randomisation in mammography screening trials has been weak and on the Canadian trials (Miller et al. 1992a, Miller et al. 1992b), where no effect on mortality was found. Other investigators found the weaknesses in the trials to be non-significant, and that the effectiveness of mammography screening has been proven (de Koning 2000, Screening mammography re-evaluated 2000). Feig (1999) calculated that 18 900 deaths from breast cancer could be averted by screening annually 1 million women from age of 40 until age of 74. At that case at most 21.6 excess deaths might be caused by radiation from mammography.

The clinical value of mammography depends on the age of the patient and is doubtful among young women (Harris and Jackson 1989). Hindle et al. (1999) examined 1908 symptomatic women $\leq$ 35 years old in the period 1992–1995. They found no clinically unsuspected cancers by mammography alone; all of the 23 invasive cancers were palpable. Lannin et al. (1993) reported the sensitivity of mammography in women less than fifty years of age to be significantly lower than that of older women (68% versus 91%, $p < 0.005$). Therefore screening women under 50 years is more controversial and the results in different studies varied concerning mortality reduction rates. Andersson (2000) found in his review article that trials indicate that screening women between 40 and 49 with mammography results in a clear mortality benefit. Such results were obtained in the Gothenburg trial (Bjurstam et al. 1997), in the UK trial, where no evidence of less benefit in women aged 45–46 at the screening was found (UK Trial of Early Detection of Breast Cancer Group 1999) in Uppsala, Sweden (Thurfjell et al. 1996) and in New Mexico (Linver and Paster 1997). Kopans (1997a) concluded in his overview of the screening trials, that screening 40–49 year-old women is as important as screening older women; and that there is no abrupt change at the age of 50 years (Kopans et al. 1996). Fletcher et al. (1993), on the contrary, concluded that there is no benefit from screening in the first 5–7 years after entry for women offered screening in their forties. Fletcher (1997) stressed that women need age-related information about both the benefits and potential risks of screening in order to make appropriate decisions regarding mammography.

A shorter, i.e. one year screening interval for younger women's screening has been recommended (Day and Miller 1988, Tabar et al. 1995, Jansen and Zoetelief 1997). The upper age limit for breast cancer screening has also been discussed (Horton 1993, Boer et al. 1995, Chen et al. 1995, Gabriel et al. 1997, Kerlikowske et al. 1999, Edwards and Jones 2000) and some benefit has been found up to the age of 79.

The European Union recommends that Member States offer a organised mammography screening programme with quality assurance at all levels for asymptomatic women aged 50–69 years. The screening interval should be 2–3 years. If screening is offered to women 40–49 years, women should clearly be informed about the possible benefits and adverse effects of screening. Organised programmes should be set in order to discourage spontaneous screening in units without adequate quality control systems, two-view mammography with double reading and 12–18 months' interval should be used and data monitoring and proper evaluation should be mandatory (Advisory Committee on Cancer Prevention 2000).

In 1995 in at least 22 countries there was a national, regional or pilot based mammography screening programme (Shapiro et al. 1998). The programmes vary in how they have been organised; guidelines for the lower and upper age limit for mammography (lowest 40 to 50, upper limit most commonly 69 but in some countries, for instance in Germany, there is no upper age limit), in screening intervals (1–3 years) and whether there is breast self examination or clinical breast examination included in the programme (for instance both in the United States and in the United Kingdom). Most programmes began in late 1980's or early 1990's and the number of countries that have established or plan to establish breast cancer screening programmes is increasing. Depending on how the mammography screening programme is organised, the percentage of target population covered by the programme and the expectations of coverage vary (Shapiro et al. 1998).

The effectiveness of screening for breast cancer as a public health policy was demonstrated in Finland (Hakama et al. 1997). In Finland a nationwide population based screening programme for breast cancer was started in 1987 and gradually implemented (Hakama et al. 1991, Hakama et al. 1999). The effectiveness in reducing breast cancer mortality has been assessed initially for the main part of the screening programme in unique group-randomised design, built-in while implementing the programme in the late 1980's. The follow-up during the period 1987–1992, up to five years from the start, showed an effect of 24% reduction in breast cancer mortality attributable to screening among those invited. The effect was statistically significant only among women under 56 at entry (Hakama et al. 1997). There is no reliable data on the later developments of the effectiveness of the programme, nor data on other centres not included on the original design.

From 1987 to 1997 the participation rate in the mammography screening programme in Finland has been high, around 89% (Pamilo et al. 1993, Dean et al. 1999) as in the Swedish screening trials (Tabar et al. 1992) and higher than that (75%) in the United Kingdom (Blanks et al. 2000). No decrease in the course of the programme has been observed as, for instance in the Netherlands from 88% in the initial screen to over 60% until round eight (Scaf-Klompf et al. 1995). The recall rate decreased from initial 4.7% to 3.3% in the first three years, with little variation thereafter. The biopsy referral rate also decreased from 1.1% to 0.6% during the first three years with little variation thereafter. 5 595 breast cancers were operated on (0.4% of those participating) of which 11% were stage 0 or TIS, 29.7% were Stage I ≤10mm, 26.9% were Stage I 11–20 mm and 28.2% were Stage II–IV (Pamilo et al. 1993, Dean et al. 1999).

2.4 Interval cancers

The effectiveness of breast cancer screening measured by reduction in breast cancer mortality depends on how the screening programme is organised, i.e. how the population targetted for screening is defined, attendance rate in screening, cancer detection rate and number of interval cancers (cancers detected between two screening rounds). The quality of mammography work up and the clinical work up, i.e. how small clinically detected breast cancers are found and how effective the treatment of breast cancer is also have an effect on the success of screening. (de Wolf and Perry 1996)

Interval cancers are a direct indicator of the sensitivity of mammography and the success of a mammography screening programme (Day 1985). The sensitivity of mammography is also affected by the biology of the breast (breast parenchymal density), histology of the tumour (diffuse growth pattern), quality of mammography and its interpretation. The visibility of breast cancer in earlier mammograms can be also regarded as a direct indicator of the sensitivity of mammography.

Day et al. (1995) compared interval cancer rates in the Swedish two county study in the first, second and third years (17%, 30% and 56% of the expected underlying incidence in the absence of screening) to those in East Anglia, U.K. (24%, 59% and 79%) and concluded that because of higher interval cancer rate the mortality reduction following screening will be lower in East Anglia (21%) compared to that observed in Sweden (35%). Frisell et al. (1987) analysed the interval cancers in the first interval in the Stocholm study and found 1.8 interval cancers per 1 000 examinations in 24 months. Burrel et al. (1996) reported 8.2 interval cancers per 10 000 women screened in 24 months after screening in Nottingham, U.K. The interval cancer rate depends on the screening interval (Frisell et al. 1987, Woodman et al. 1995) and on the age group to

which screening is offered. Klemi et al. (1997) found in a population based screening 422 screen-detected and 104 (25%) interval cancers. In the age group 40–49 years the interval/screen detect cancer rate was 17/52 (25%) in one-year screening intervals and 17/32 (35%) in three-year screening intervals. Women aged 50–74 years were screened at 2-year intervals and the responding rate for them was 70/338 (17%).

The sensitivity of mammography screening improves with experience. Denton and Field (1997) reported improvement during the three years of the first screening round. Kan et al. (2000) found that the mean standardized abnormal interpretation ratios and cancer detection ratio improved gradually with the increasing annual reading volume of the radiologists.

Double reading has been reported to increase the number of breast cancers detected (Thurfell 1994) and decrease the number of women recalled i.e. to increase both sensitivity and specifity of screening. According to Anttinen et al. (1993) independent double reading of mammograms increased the breast cancer detection rate by 9% and decreased recall rate by 45%. Brown et al. (1996) found a consensus double reading policy to increase the number of breast cancers detected by 9 per 10 000 women screened and to decrease the recall rate significantly. There have been experiences suggesting that the use of a third reader to arbitrate on disagreements reduces unnecessary recalls (Mucci et al. 1999).

2.5 Retrospective review of mammograms

The amount of retrospectively visible cancers (cancers which were originally not noted but were visible in retropective review) can be estimated in different ways. Retrospective review of earlier mammograms can be made blinded or retrospectively, i.e. with the finding in diagnostic mammograms known or not known and by one or more radiologists. The number of retrospectively visible carcinomas depends on the design of the review.

Frisell et al. (1987) found 31/60 of interval cancers to be true interval tumours, i.e. no sign of them could be found in the first mammogram; the rest, non-true, could be traced on the first mammogram. Ikeda et al. (1992) reported 94 interval carcinomas with earlier mammograms available and found that 10 of them were missed (observer's error), 63 had no tumour in retrospective review and 21 had subtle signs of malignancy in earlier mammogram, mostly nonspecific densities or asymmetries. Burrel et al. (1996) analysed the earlier screening mammograms of 90 interval cancers. Of these 51 were true-positive (mammographic abnormality was seen at the time of symptoms, but screening mammography was regarded as normal by all four reviewer radiologists), 20

were false negative (some of the reviewers would have recalled the patient), 7 were mammographically occult (no mammographic abnormality at the time of symptoms) and 12 were unclassified (mammography was not obtained at the time of symptoms). The most common missed abnormality in the false-negative cases was architectural distortion. It was found, that prognosis of interval cancers is similar to that in symptomatic, unscreened tumours and statistically poorer than that of screen-detected tumours.

Jones et al. (1996) in a blinded review found that 30 (22.6%) out of 133 incident screening cancers were false negative at the prevalent screen. Of these 25 were earlier either not noticed or diagnosed as benign and 5 were misdiagnosed on further assessment. Harvey et al. (1993) reviewed 152 previous mammograms in patients in whom inpalpable breast carcinomas were subsequently detected on later mammograms. The earlier studies were interpreted both blinded (without knowledge that carcinoma was subsequently detected and retrospectively (with the mammogram showing the carcinoma for comparison). They found 30 (41%) out of 73 previous mammograms performed 6–39 months before the diagnostic mammograms to have evidence of carcinoma in blinded review. Out of the remaining 43 mammograms they found some signs of malignancy in 25 (34%) when preoperative mammograms were shown. Ma et al. (1992) analysed the mammograms of histologically confirmed breast cancer patients and found three variables to be associated with failure to detect breast cancer: extensive parenchymal densities in mammography, a tumour of lobular histology and tumours of small size.

The Northern Region Breast Screening Audit Group (Simpson et al. 1995) has developed a method of classifying interval cancers. Each member of the group independently reads a set of films which contains both interval cancers and controls from centres other than their own. A cancer is classified as "false negative" only when it is correctly identified, on the previous screening film, by at least two members of the group. Analysis of 167 interval cancers showed 46% "true intervals" (cancer was detected on pre-operative mammogram and was not present on the earlier screening mammogram), 26% "false negatives", 1.2% technical false negatives (screening mammogram was technically inadequate), 11% "occult" (clinically detectable but not shown on pre-operative mammogram), and 16% not classifiable because of incomplete data. Van Dijck et al. (1993) reviewed the previous screening mammograms of 44 screen-detected and 40 interval cancer cases. They classified 13% of all cases as "screening error", 38% as "minimal sign present", 43% "radiographically occult" (no changes at all from the previous mammogram) and 6% as "radiographically occult at diagnosis".

Bird et al. (1992) found that missed cancers occured in women with denser breasts, were less likely to demonstrate malignant microcalcifications and were more likely to demonstrate a developing opacity than cancers correctly diagnosed at sceening mammography. Moberg et al. (2000) reported reviewing mammograms of 119 women with breast cancer and the previous round attendance. The earlier screening mammograms were reviewed by both single and double reading, first mixed with other screening images and then non-mixed. The proportion of cancers detected on retrospective review varied between 5% and 50% depending on the reviewing method used and the number of reviewers included to classify a case as truly identified. Generally more cancers were detected when non-mixed samples of mammograms were reviewed than when mixed samples were reviewed (mean increase 23%) and when interpreted by double reading compared with single reading (mean increase 14%). Goergen et al. (1997) concluded that missed cancers were statistically significantly lower in density and more often seen on only one of two views than detected cancers.

Coveney et al. (1994) reviewed 291 mammograms taken prior to biopsy of a palpable tumour. False negative reports occured in 16.5%. Retrospective review of false negative mammograms showed 20% oversights i.e. missed carcinomas, 50% showed subtle radiographic abnormalities (27% asymmetric breast density, 14% microcalifications initially considered benign, 7% minor architectural distortion of the breast parenchyma and 2% benign looking postmenopausal cyst, which histology revealed to be a fibroadenoma with adjacent minimal breast carcinoma) and 30% were normal with no abnormality

Kavanagh et al. (2000) examined the sensitivity, specificity and the positive predictive value of screening mammography by symptomatic status in first round mammography. They divided symptomatic status into the following categories: asymptomatic; significant symptoms, if the woman reported a breast lump and/or blood stained or watery nipple discharge; and other symptoms, if reported. They found the sensitivity lower for women with other symptoms (60.0%) than asymptomatic women (75.6%) or women with significant symptoms (80.8%). They regarded increased breast density and poor image quality as possible explanations for lower sensitivity for women with other symptoms.

Poplack et al. (2000) described measures of mammography performance in a geographically defined population. Screening mammography had a sensitivity of 72.4%, specificity of 97.4% and positive predictive value of 10.6%. Diagnostic mammography had higher sensitivity, 78.1%, lower specificity, 89.3% and better positive predictive value, 17.1%. The cancer detection rate with screening mammography was 3.3 and interval cancer rate 1.2 per 1 000 women.

2.6 Ultrasonography

In the 1970's ultrasonography (US) was used mainly in the identification of breast cysts. In the 1980's investigators reported on the features of benign and malignant breast masses in US, which, however, were not yet reliable enough to avoid biopsy (Rahbar et al. 1999). Recently Rahbar et al. (1999) investigated the general applicability and interobserver variability of US features in differentiating benign from malignant solid breast masses. Benign features in US were round or oval in shape, and displayed circumscribed margins and width-to-anteroposterior ratio greater than 1.4. Malignant features were irregular shape, microlobulated or spiculated margins and width-to-anteroposterior ratio 1.4 or less. Enhancement of the beam is usually associated with benign lesions, particularly cysts; reduction or shadowing with malignancy (Egan 1988).

Zonderland et al. (1999) investigated the value of US as an adjunct to mammography for the diagnosis of breast cancer. With US the sensitivity of breast cancer diagnostics increased from 83% to 91% and the increase was highest among women younger than 50 years. Moss et al. (1999) investigated the value of US adjunct to mammography in symptomatic women. They concluded that extensive use of US increased the cancer detection rate by 14%. US is also widely used as a guidance method for cyst aspiration and fine or core needle biopsies (Rumack et al. 1991, Rissanen 1994).

The members of the Eurorean Group for Breast Cancer Screening recommended US to be used as a complementary method to mammography and clinical examination in the further assessment of both palpable and impalpable breast abnormalities (Teh and Wilson 1998). They found that there is little evidence to support the use of ultrasound in population based screening at any age. The International Breast Ultrasound School (IBUS) has given guidelines for the ultrasonic examination of the breast. The quality of ultrasonic imaging depends on equipment, examination technique and interpretation and only high resolution instrumentation producing high quality images should be used by experient investigators (Majdar et al. 1999).

Kopans et al. (1985) in a 4-year follow-up found no cancers in the group that had suspicious findings by whole-breast US imaging only. However, some physicians have used US as a screening method in women with dense breasts. Kolb et al. (1998) reported the results of a total of 11220 consecutive patients prospectively examined. All 3626 women with dense breasts and normal mammographic and physical examination findings underwent screening US performed by a physician. Out of 136 cancers 11 (8%) were identified by US alone. US and mammographically detected cancers regarding

mean surgical size and stage were not statistically different and were both smaller and lower in stage than palpable cancers. Buchberger et al. (1999) found 23 malignancies in 21 patients by high-resolution sonography as an adjunct to mammography (prevalence 0.31%) found in 6113 asymptomatic women with dense breasts. Berg and Gillbreath (2000) evaluated preoperative whole-breast US in the ipsilateral breast of 40 patients with known breast cancer, or high suspicion of it, to find out the multicentricity or multifocality of the tumour. Nine out of 64 malignant foci were seen with US alone and required wider excision than by mammography finding.

Metastatic involvement of the axillary lymph node is the primary prognostic factor for breast cancer and US may be more sensitive than mammography or clinical examination in detecting it (Pamilo et al. 1989). Bruneton et al. (1986) compared clinical examination of axillary lymph nodes to US. The sensitivity was 45% versus 72.7%. However, US depicts lymph node enlargement, but provides no information on the histology, thus metastases must be confirmed by biopsy.

Tumour size varies according to the time of diagnosis. The correct determination of tumour size is important for patient management, particularly for breast conservation therapy. US has shown the best correlation with pathologic cancer size when compared with mammographic and clinical measurements (Solbiati et al. l995).

2.7 Fine needle aspiration biopsy

Fine needle aspiration cytology (FNAC) assessed by palpation, US (Rissanen et al. 1998a), mammographic or sterotactic guidance (Evans and Cade 1989) is a valuable diagnostic method for both palpable and non-palpable breast lesions. Hann et al. (1989) found 61 out of 96 cytology aspirates taken with standard mammographic technique adequate for diagnosis, and cytologic examination permitted accurate diagnosis of 21 out of 23 carcinomas. Hinton et al. (1999) published their FNAC data recently. Out of a total 138 cancers the FNAB result was malignant (class 5) in 125 (90.6%).

The accuracy of mammography in differentiating benign and malignant subclinical breast lesions is not high, and unnecessary biopsies of benign lesions are the cost to be paid for the detection of subclinical cancers. FNAB of nonpalpable mammographic lesions under sonographic or stereotactic guidance can achieve a sharp reduction in unnecessary benign biopsies in cases of low suspicion of malignancy in mammography. FNAC also has a major role in diagnosing recurrences (Rissanen et al. 1997).

Ciatto et al. (1989) reported the results of their institution where the benign-to-malignant biopsy ratio of nonpalpable lesions was greatly reduced after all nonpalpable

mammographic lesions were assessed by means of stereotactic aspiration cytology. The sensitivity and specificity of FNAC can be high (96.7% and 77.7%) and false-negative/ inadequate cytology low (27/ 2444 cases) with low suspicion of malignancy at mammography as in Ciatto et al. (1997). With FNAB the need for follow-up mammography can be reduced and so radiation exposure, patient anxiety, unnecessary biopsies and costs will be likewise reduced (Franquet et al. 1992). Dowlatshahi et al. (1989) reached 95% sensitivity, 91% specificity and 92% accuracy with stereotactic guided FNAB of clinically occult breast cancers.

Insufficient specimens may be a major problem in FNAB (Pisano et al. 2001). Hayes et al. (1995) decreased the rate of insufficient specimens for FNAC from 27% to 8% by modification of techniques and implementation of a quality assurance programme. Another problem with FNAB is that it may cause suspicious density in mammography and malignant-looking features in US and should always be performed after other diagnostic approaches (Svensson et al. 1992).

2.8 Core needle biopsy

The limitations of FNAB, especially the false negative rate due to insufficient material and the absence of information concerning the invasive characteristic of the lesion has led to the use of core needle biopsy (CNB) in many mammography units (Dronkers 1992). CNB can be performed by palpation, US or stereotactic guidance (Hall 1997). Special needles and automated biopsy devices are used to obtain adequate specimens for histological diagnosis (Bernardino 1990). The combination of FNAB and CNB has decreased the amount of unnecessary surgery on abnormal but benign mammographic lesions (Dowlatshahi et al. 1991). In long-term follow-up studies the false negative rate with stereotactic, automated large-core needle biopsy has been as low as 1.2%–4.0% (Jackman et al. 1999).

We can expect that the rate of mammography follow-up studies on probably beingn lesions (Varas et al. 1992, Sickles 1991) can be lowered through FNAB and CNB. These were earlier used otherwise successfully, but with patient anxiety.

3 Aims of the study

The general aim of the study is to evaluate the sensitivity of mammography screening in relation to its determinants. The more detailed aims are to evaluate the effect of age of the patient, breast parenchymal density, the tumour characteristics, the interpretation of the mammograms on the sensitivity of mammography; and to compare sensitivity of mammography to other diagnostic methods.

The thesis consists of four articles which specifically cover the following issues:

1) to examine the effect of age of the patient on the sensitivity of mammography and how much extra benefit US and FNAB offer for mammography (Publications I and II).
2) to find out whether both patient's age and density of the breast have an independent effect on the sensitivity of mammography or whether one of these correlated variables, age or density, accounts for all the variation in the sensitivity of mammography. Breast pattern is not uniform over the breast and therefore the effects of both the overall density of the breast and the density at the site of the tumour on the sensitivity were examined (Publication II).
3) to examine the visibility of screening and clinically detected cancers in earlier mammograms by a retrospective review without and with preoperative mammograms shown, how much age of the patient and breast parenchymal density influenced the retrospective visibility and what kind of changes originally remained unnoticed (Publication III).
4) to estimate how many screening and interval cancers could have been detected in the earlier screening round by improved interpretation of the mammogram and whether there are any radiological or clinical differences between retrospectively visible and invisible cancers (Publication IV).

4 Materials and methods

The thesis is based on two different materials. The first material (screening results in general; Publication IV) originates from the mammographically screened women in Tampere and its surroundings (about 440 000 inhabitants) in the period 1987–1992 (Table 1).

Screening mammography was performed in two mammography units: women living in the city of Tampere were screened in Tampere City Hospital and women living in the surroundings of Tampere were screened in Pirkanmaa Cancer Society. The mammograms were taken with Mamex DC (Soredex, Helsinki, Finland), later Alfa 3 (Instrumentarium, Helsinki) or Mamex Dc S (Soredex, Helsinki, Finland) using Kodak Min Rm screen film combination for mammography. The developers were Kodak X-omat 35 S, later Kodak Day-light M 35-M, and Cronex.

Regular screening is organised and paid for by the municipalities. The recommendation for screening interval is 24 months. Cancers diagnosed at screening, between screens, and among non-attenders were checked by linkage of the screening data to the Finnish Cancer Registry. Therefore the study material includes all new breast cancer cases in the defined population in the age groups screened. A woman was classified as a non-attender if she had not responded to the invitation for the screening round preceding the diagnosis of breast cancer. Cancers diagnosed between two consecutive screening rounds were considered to be interval cancers. To further confirm the validity of the patient data, all medical records were retrospectively reviewed by a senior radiologist and oncologist.

Table 1. Population based mammography screening results in numbers (N) and percentages (%) in the Tampere screening programme 1987–1992 (material I, Publication IV).

	N	%
Invited	63 731	
Screened	56 158	88.00[1]
Recalled	2 269	4.00[2]
Open surgical biopsies	441	0.79[2]
Benign	218	0.39[2]
Screen-detected cancers	213	0.38[2]
1. round cancers	136	0.24[2]
2. round cancers	52	0.09[2]
3. round cancers	25	0.04[2]
Interval cancers	54	0.10[2]
Cancers among non-attenders	9	0.02[1]
All cancers	276	0.49[2]
Recurrences	9[3]	0.02[2]

[1] % among those invited
[2] % among those screened
[3] excluded from the data

In 130 (76 screen-detected and 54 interval cancer) cases the radiologist who diagnosed the breast cancer in or outside of the screening reviewed the mammograms taken in an earlier screening round together with the subsequent mammograms and recorded radiological data concerning the visibility and size of the tumour (Publication IV).

No blinded review of mammograms was carried out. The tumour was regarded as visible in retrospect if there were any, even minimal, abnormalities in the location where cancer was subsequently detected. All kind of densities, even benign-looking tumour masses, structural distortions and microcalcifications were regarded as abnormalities. The radiological signs and tumour characteristics were analysed from the mammogram taken at the time the tumour was diagnosed.

Clinical and histopathological data were recorded from patient files of the Tampere University Hospital or Tampere City Hospital, where all the patients were treated and followed up. Clinicopathological characterization of tumours was done according to TNM classification and histopathology classification was performed according to the WHO scheme. TNM system offers details regarding the size of the tumour (T), the status of lymph nodes (N), and any distant metastases (M) (Kirby and Copeland 1991). It can be used for clinical diagnostic, surgical evaluation or postsurgical pathological staging. The TNM system classifies tumours by their anatomical extent using designations and then combines these groups into prognostically similar categories called stages. The stages are numbered from the best (Stage 0) to the poorest (Stage IV) regarding prognosis based on survival rates: i.e., Stage 0, carcinoma in situ; Stage I, localized cancer; Stage II, limited local or regional spread; Stage III, extensive local or regional spread; and Stage IV; distant metastasis. Grading is a measurement of the degree to which the tumour tissue lacks normal histological differentiation; this may be recorded as general grade or nuclear grade. In general, the better the differentiation or grade, the better the prognosis. Grade I means well-differentiated; Grade II moderately differentiated and Grade III poorly differentiated. (Kirby and Copeland 1991.) All cytological samples were taken using fine needle aspiration technique under mammographic, ultrasonographic or palpation guidance.

For the second study material oncologists collected clinical data and radiologists group radiological data on 659 histologically and 6 cytologically verified new consecutive breast cancers in the period 1996–1997 in two areas of Finland (Pirkanmaa and North-Karelia). The study areas covered 227 000 and 89 000 women respectively. Yearly 300 and 100 new breast cancers were diagnosed in the study areas. Each area has only one oncological clinic and almost all new breast cancer patients were referred there for treatment or consultation. Most municipalities offer organised mammography screening for women aged 50–59 years. For women older than that local cancer

societies have offered screening at the patients' expense. Both screenings are based on personal invitation, two-view, double reading mammograms at two-year intervals. Study material II consists of 90% of all new breast cancers in the study areas during the study period. (Table 2)

The youngest patient in material II was 26, the oldest 92, the mean age being 60. Tumour size varied between 3 and 130 mm, mean 19mm. A total of 62% of cancers were node negative; 76% of screen-detected cancers, 77% of screen-detected at own expense, 55% of interval cancers, 77% of cancers in non-attenders and 50% of cancers in women under and 53% in women over screening age. In only 21 cases (3%) were there distant metastases by the time breast cancer was diagnosed; seven in young women, one in a woman with interval cancer, one in a woman in whom cancer was detected by screening at own expense and 11 in women older than screening age. (Table 3)

Table 2. Study material II according to age of the patient and detection mode of the tumour (Not in original Publications)

	N	%
Under screening age		
26–39 years	18	3
40–49 years	105	16
Screen-detected (50–59/64 years)	192[1]	29
Interval cancer	78	12
Cancers among non-attenders	17[2]	3
Screening at own expense (60/65–70 years)	44	7
Other over screening age	207[3]	31
Men	4	1
Total	665	100

[1]4, [2]1 and [3]5 women had bilateral cancer

Table 3. Study material II according to stage and detection mode of the tumour (Not in original Publications)

	TIS	Stage I		Stage II–IV, TX[1]	Total
		<11 mm	11–20 mm		
	%	%	%	%	N
Under screening age	4	16	21	59	123
Mammography screening	14	26	31	30	192
Interval cancer	8	19	18	55	78
Cancers among non-attenders	0	35	41	24	17
Screening at own expense	9	25	23	43	44
Other over screening age	5	8	20	67	207
Men	0	25	0	75	4
Total	8	18	24	51	665

[1]Tumour size cannot be assessed or it is missing.

Of these tumours 24% were operated on in two weeks, 53% in two to four weeks and 17% in one to two months after radiological examination.

Interval cancers amounted to 27% in material II of all cancers in screening age groups including non-attenders during the study period. Stage distribution was more favourable in screen-detected tumours and in material II they were more often node negative local tumours than cancers among women over or under screening age (Table 3). Out of the four men one had nodal and one nodal and distant metastases.

The study area covers over 10 mammography units. The original radiological work-up was done by over 50 radiologists. However, 76% of the examinations were done by 5 radiologists.

Publication I included all patients of material II.

Publication II comprised those 577 breast cancer patients from material II on whom both mammography and US was performed preoperatively and in whom breast parenchymal density could be classified both in total breast and in the tumour area.

Publication III included those 320 patients from material II from whom earlier and preoperative mammograms were available for review and breast parenchymal density was classified both in total breast and in tumour area.

Radiologists use a 5-point rating scale for grading probability of malignancy in mammography and ultrasonography findings (1 = normal, 2 = benign, 3 = indeterminate, 4 = probably malignant, 5 = malignant). Classes 1–2 indicate benign findings, classes 3–5 malignant findings. Cytological and histopathological examinations were done by several cytopathologists whose experience varies. FNAC findings were classified originally by a pathologist according to Papanicolau into classes 0–5, where class 0 is an acellular or inadequate aspirate and classes 1–5 the same as those used for the radiological scale. No reclassifications were done. Performing core needle biopsies began during the study period. 25 core biopsies were obtained with Bard TruGuide Coaxial Biopsy Needles (C.R.Bard Inc., Covington, KY, USA) using a biopsy gun with a 22-mm throw (Bard Magnum, Bard Radiology Division) under stereotactic or US guidance.

The sensitivity of the examinations was calculated by the model:

sensitivity (S) =a / (a+c),

where a= true positives i.e. patients with the disease who are correctly diagnosed with the test

c= false negatives i.e. patients with the disease who were not diagnosed with the test

This means that sensitivity defines how many of the diseased patients the test diagnoses. We considered classes 1–2 as negative and classes 3–5 as positive for the disease. An ideal test has high sensitivity (correctly identifies a high proportion of truly diseased individuals) (dos Santos Silva 1999).

Table 4. Patients in Publications I–III originating from study material II, and purpose of publications

	Publication I	Publication II	Publication III
Gender			
Women	651[1]	577	320
Men	4	0	0
Age			
Minimum	26	26	36
Maximum	92	92	76
Mean	60	59	59
Purpose	Sensitivity of mammography, US and FNAB	Sensitivity of mammography and US by age and density	Visibility of cancers in earlier mammograms by age, density and type of finding

[1] 10 women had bilateral cancer

The material consists of breast cancer patients only, and no false positive cases or total number of examinations made during the study period are included. That is why we cannot estimate specificity.

Retrospective evaluation of mammograms was made by groups of local radiologists. In both areas one radiologist was present at each reading session (IS or PP), and the rest consisted of at minimum two and at maximum seven radiologists with long experience in mammography work-up. In practice it was not always possible for the whole group to participate in all reading sessions.

The groups of radiologists classified breast density in mammogram in total breast and in tumour area in three classes (fatty, mixed and dense) by comparison with model mammograms. Tumours were classified as not visible (finding normal or benign) and visible (finding malignant or suspect) in mammogram. Whether the tumour was originally visible or not was decided on the basis of original patient files and reviewing results on the consensus opinion of the radiologist group.

For both materials statistical analyses were made by SPSS for Windows 95 and INSTAT softwares. The statistical test of significance was based on two-tailed Fisher's exact test, and Chi-square test was used for testing trends (Motulsky 1995). Because of the dependency between observations McNemar test was used to find out whether there is statistically significant differencies in the sensitivities of different diagnostic methods. Chi-square tests were done to find out whether the sensitivity differs between age groups and breast density classes and whether the screen-detected cancers differ from interval cancers. Logistic regression was used to analyse both univariate and multivariate effects of age, breast density and time lag on the sensitivity of both mammography and ultrasonography (Hosmer and Lemeshow 1989).

5 Results

5.1 Process measures of the screening programme

Material I consisted of 50–59 year old women to whom mammography screening was offered in the period 1987–1992 according to the national recommendations. The attendance rate was 88% of those invited, the recall rate was 4% among those screened and cancer detection rate was 0.4% of those screened. The most common screening interval was 23–25 months, the shortest interval was 12 months and the longest five years eight months, the mean was 24.4 months. There were 136 first round, 52 second round, 25 third round screen-detected cancers and 54 interval cancers. Nine cancers among non-attenders were diagnosed. (Table 1)

Half of the tumours in material I were palpable, interval cancers (91%) statistically significantly more often than screen-detected ($p < 0.0001$) and first round screen-detected (48%) more often than those detected in subsequent rounds (35%) ($p < 0.08$). Three out of nine non-attenders' cancers (30 %) had axillary metastases (Table 5). First round screen-detected cancers had axillary metastases more often (24%) than later rounds cancers (19%), but the difference was not statistically significant ($p = 0.6$). Interval cancers had axillary metastases significantly more often (38%) than screen-detected ones ($p = 0.04$). Only two interval cancers had distant metastases by the time of detection of the tumour.

Tumour size was <10mm in 44% and <20mm in 82% of all cancers. Only 7% of cancers exceeded 40mm in size. In subsequent screening rounds more tumours (41%) were detected as <10mm in size than in first round screening (29%) or among interval cancers (16%). Most of the first round screening cancers were slow growing since 59% of them were Grade 1 and only 4% were Grade 3 compared with 50% Grade 1 and 13% Grade 3 of subsequent round and 36% Grade 1 and 19% of Grade 3 of interval cancers. On the other hand more (28%) first round screening cancers were Stage 2–4 (size >20mm and / or axillary or distant metastases) than in later rounds (20%).

Out of 54 interval cancers, 16 (30%) appeared within 12 months of a previous screening, 15 (28%) within 12 to 18 months, 16 (30%) within 18 to 24 months and 7 (13%) more than 24 months after the earlier screening round; mean interval was 16.8 months. The shortest time from normal interpreted screening mammography to interval cancer operation was two months and the longest three years seven months. Interval cancer rates in the first and second years after a negative screen were 18% and 54%, and in total the interval cancer rate was 61% of the expected underlying incidence in the absence of screening which in 1986 – just before screening started – was 168/100 000 (Finnish Cancer Registry 1990).

Table 5. Tumour size measured from the surgical specimen by a pathologist, tumour grade, axillary nodal status and postsurgical tumour stage determined by a clinician. Tampere screening programme 1987–1992 (Not in original Publications).

	Screening round		Interval	Non-attenders	Total
	1.	2.–3.			
	N (%)	N (%)	N (%)	N (%)	N (%)
Tumour size					
1– 9 mm	20 (29)	19 (41)	5 (16)	0 (0)	44 (30)
10–14 mm	18 (26)	11 (24)	8 (26)	0 (0)	37 (25)
15–19 mm	8 (12)	7 (15)	4 (13)	0 (0)	19 (13)
20–29 mm	10 (15)	4 (9)	8 (25)	1 (25)	23 (15)
30–49 mm	8 (12)	4 (9)	5 (16)	1 (25)	18 (12)
50–55 mm	0 (0)	0 (0)	0 (0)	2 (50)	2 (1)
No measurable tumour	4 (6)	1 (2)	1 (3)	0 (0)	6 (4)
Total with data known	68 (100)	46 (100)	31 (100)	4 (100)	149 (100)
Data missing	68	31	23	5	127
Tumour grade					
1	40 (59)	27 (50)	13 (36)	0 (0)	80 (49)
2	25 (37)	19 (35)	16 (44)	2 (50)	62 (38)
3	3 (4)	7 (13)	7 (19)	2 (50)	19 (12)
Total with data known	68 (100)	53 (100)	36 (100)	4 (100)	161 (100)
Data missing	68	24	18	5	115
Axillary nodal status					
pN0	84 (76)	55 (81)	30 (63)	1 (25%)	170 (74)
pN+	26 (24)	13 (19)	18 (38)	3 (75%)	60 (26)
Total with data known	110 (100)	68 (100)	48 (100)	4 (100%)	230 (100)
Data missing	26	9	6	5	46
Postsurgical tumour stage					
0	13 (11)	2 (3)	0 (0)	0 (0)	15 (7)
1	71 (61)	53 (77)	24 (63)	1 (17)	149 (65)
2	28 (24)	12 (17)	12 (32)	3 (50)	55 (24)
3	3 (3)	2 (3)	2 (5)	2 (33)	9 (4)
4	1 (1)	0 (0)	0 (0)	0 (0)	1 (1)
Total with data known	116 (100)	69 (100)	38 (100)	6 (100)	229 (100)
Data missing	20	8	16	3	47
Total study material I	136	77	54	9	276

Only two interval cancer patients had distant metastases by the time of detection of the breast cancer.

5.2 Mammographic appearance of tumours detected in and outside of screening

In mammogram most tumours occurred as spiculated tumour masses and there were no major differences depending on the detection mode. Screen-detected cancers had microcalcifications statistically significantly more often (35% in first and 31% in later screening rounds) than interval cancers (8%) ($p = 0.005$). In Table 6 the mammographical characteristics of cancers in material I are shown.

Table 6. Main mammography findings depending on detection mode in numbers (N) and percentages (%). Tampere screening programme 1987–1992 (Not in original Publications).

	Screening round		Interval	Total
	1. N (%)	2.–3. N (%)	N (%)	N (%)
Main finding				
Mass	89 (66)	53 (68)	21 (70)	163 (67)
Microcalcifications	24 (18)	14 (18)	1 (3)	39 (16)
Mass with microcalcifications	14 (10)	8 (10)	3 (10)	25 (10)
No abnormality	2 (2)	0 (0)	1 (3)	3 (1)
Other[1]	6 (4)	2 (4)	4 (13)	12 (5)
Total with data known	135 (100)	77 (100)	30 (100)	242 (100)
Data missing	1	0	24	25
Tumour borders				
Spiculations	84 (62)	43 (56)	18 (67)	145 (60)
Unsharp defined	26 (19)	18 (23)	5 (19)	49 (20)
Well defined	1 (1)	3 (4)	0 (0)	4 (2)
No tumour	25 (18)	13 (17)	4 (15)	42 (18)
Total with data known	136 (100)	77 (100)	27 (100)	240 (100)
Data missing	0	0	27	27
Microcalcifications				
No microcalcifications	89 (65)	53 (69)	24 (92)	166 (69)
Microcalcifications	47 (35)	24 (31)	2 (8)	73 (31)
Total with data known	136 (100)	77 (100)	26 (100)	239 (100)
Data missing	0	0	28	28
Total study material	136	77	54	267

[1]Other mammography findings include asymmetry, architectural distortion and skin retraction. There was very little information available about non-attenders' cancers; one out of nine had microcalcifications and all other radiological data were missing.

5.3 Sensitivity of mammography by age of the patient

In material II 92% of cancers were diagnosed by mammography (mammograpahy originally classified malignant or suspect, classes 3 to 5) and thus the sensitivity of mammography was 0.92. The sensitivity of mammography increased by increasing age of the patient and was 0.83 among patients aged 26–49, 0.91 among patients aged 50–59 and 0.96 among patients aged 60–92 years ($p < 0.0001$ for trend). (Publication I) 3% of breast cancer cases were operated on without any radiological examinations and 11% without cytological or histological verification.

By logistic regression analysis (Publication II) the relative sensitivity of mammography was 0.1 (95% CI 0.1–0.3) in the youngest age group and 0.3 (95% CI 0.1–0.8) in the middle age group compared to the oldest age group.

5.4 Effect of density of the breast on the sensitivity of mammography

In material II the density of the breast was heavily dependent on the age of the patient and decreased with age both in total breast and in tumour area. In most cases overall density and density in tumour area were classified to be the same, there was a difference in only 12% of cases. The relative sensitivity of mammography was 0.3 (95% CI 0.1–0.6) in dense and 0.9 (95% CI 0.4–2.3) in mixed breast compared to fatty breast by using overall density. Density in the tumour area had slightly more effect on the sensitivity. The relative sensitivity of mammography was 0.2 (95% CI 0.1–0.4) in dense and 0.7 (95% CI 0.3–1.6) in mixed parenchyma compared to fatty. (Publication II)

5.5 Retrospective visibility of cancer in penultimate mammograms

In material II out of 320 earlier mammograms of consecutive new breast cancer cases 14% (45) of cancers were retrospectively visible in earlier mammograms without the preoperative mammograms having been shown and 29% (95) when preoperative mammograms were shown to the group of radiologists who were reclassifying the findings. The percentage of retrospectively visible cancers increased by age and was 12%, 28% and 30% among premenopausal, perimenopausal and postmenopausal women respectively. (Publication III)

In material I out of the 130 later screening rounds or interval cancers cases 43 (33%) of cancers were visible, 84 (66%) invisible and three (2%) not included on the mammogram in a retrospective review. Later-round screen-detected cancers were

statistically significantly more often visible in earlier screening mammograms (43%) than interval cancers (19%) (p = 0.002). (Publication IV)

5.6 Sensitivity of mammography by radiological and histological appearance of the tumour

In material II the most common reasons as assessed by the radiologists for non-detection were that the lesion was overlooked (55 cases), diagnosed as benign (33 cases) or was visible only in one projection (26 cases). Growing density was the most common (37%) feature of those lesions originally overlooked or regarded as benign. (Publication III)

In material I primary screen-detected cancers showed microcalcifications more frequently (28%) than interval cancers (13%), but the difference was not statistically significant (p = 0.13). Retrospectively invisible carcinomas showed microcalcifications more often (19%) than the visible ones (15%), but the difference was not statistically significant either (p = 0.15). Tumours missed by screening mammography but which were visible on retrospective review were often histologically well-differentiated and more often diagnosed by the subsequent screening round than by clinical diagnosis as interval cancers. If all retrospectively visible interval cancers had been diagnosed by screening 19% (10/54) of the interval cancers could have been avoided. If all retrospectively visible cancers had been diagnosed at the time of false negative screening or assessment 65% (84/130) of all patients would have benefitted from an earlier diagnosis compared to the actual figure of 31% (41/130). (Publication IV)

5.7 Joint effects of determinants on the sensitivity of mammography

Age and density are correlated (Table 7) and therefore we considered the effect of age and density simultaneously. The density-adjusted sensitivity of mammography in the youngest age group was OR = 0.2 (95% CI 0.1–0.5) and 0.4 (95% CI 0.1–1.0) in the middle age group compared to that in the oldest age group. Sensitivity of mammography increased with fattiness of the breast, especially in the tumour area. Age-adjusted OR was 0.4 (95% CI 0.1–1.0) for dense breast in tumour area compared to fatty breast. (Publication II)

Table 7. Original interpretation of mammogram by overall density of the breast and age.

Age	Overall density of the breast	Interpretation of mammogram 1–2	3–5	Total
26–49	Fatty	1	8	9
	Mixed	5	37	42
	Dense	12	44	56
50–59	Fatty	3	36	39
	Mixed	5	105	110
	Dense	8	49	57
60–92	Fatty	4	133	137
	Mixed	2	102	104
	Dense	1	22	23
All	Fatty	8	177	185
	Mixed	12	244	256
	Dense	21	115	136

The retrospective visibility of breast cancers in earlier mammograms increased with age because the breast parenchymal density decreased with age. The retrospective visibility was affected by the time period between the two mammogram examinations, and this should be considered as a confounding factor. We therefore considered the effect of age, density and lag simultaneously. The restrospective visibility of cancer was 2.9 times higher among women aged over 55 compared to women aged under 55 (95% CI 1.6–5.2) and 1.5 times higher among women with fatty parenchyma compared to women with mixed or dense parenchyma. The effect of age was stronger than that of density. (Publication III)

5.8 Sensitivity of mammography vs. US and FNAB

Mammography was significantly more sensitive (Se = 0.92) at the cutoff level of class 3 than ultrasonography (US) (Se = 0.86, p = 0.004) or fine needle aspiration biopsy (FNAB) (Se 0.85, p = 0.004). In 4% of cases the mammography finding was normal or benign but US finding was malignant and thus US increased sensitivity by 4%. US was more sensitive in younger than in older women. US increased sensitivity 11% in the youngest age group, 3% in the middle age group and 2% in the oldest age group.

Mammography finding was normal or benign but the FNAB finding was malignant in 7%, and thus FNAB increased sensitivity by 7%. In 3% the mammography and US findings were normal or benign but FNAB was malignant. In 0.7% the mammography and FNAB finding were benign but the US finding was malignant. In

4% of cases US and FNAB findings were benign but the mammography finding was malignant. The sensitivity of FNAB was not dependent on the age of the patient.

All three examinations failed to show the malignant characteristics of the tumour in two cases, which were operated on without delay because of a palpable lump after the false negative examinations. (Publication I)

6 Discussion

Our first study material consists of all cancers among screen-aged women and our second material of 90% of all breast cancers in the study areas during the study period. Diagnostic work up was made in several mammography units by many physicians. Thus the results reflect everyday diagnostic work well. In routine work radiologists use a 5 point rating scale for grading the probability of malignancy in mammography findings resembling the grading system by Roche et al. (1998), which makes retrospective review of patient files easy and reliable. On the other hand in many cases important data regarding tumour size and stage was not available. It was especially difficult to get information about interval cancer cases, which were often diagnosed outside of the screening units.

The attendance rate in screening material was 88%, which is as high as in Sweden (Tabar et al. 1985) and higher than in the United Kingdom (Blanks et al. 2000) and has not fallen with time. In the Netherlands the attendance rate in the initial screening round was 82% –89% but it decreased approximately 12% or more in later screening rounds (Scaf-Klomp et al. 1995).

Recall rate was 4%, which is on approximately the same level as in other screening units in Finland, likewise the benign to malignant ratio for surgical biopsies in the first years of screening (Dean et al. 1999).

The age-specific and age-adjusted incidence rate of breast cancer in screen-aged women in 1986 – just before screening started – was 168/100 000 (Finnish Cancer Registry 1990). The cancer detection rate in screening was two times higher than the expected underlying incidence in the absence of screening and on the same level as elsewhere in Finland.

It is difficult to compare sensitivity to results from other studies because of differences in age groups screened, screening intervals, screening modes (one versus two-view mammography, single versus double reading), accuracy in reporting of interval cancers, the length of the follow-up and different methods in calculations. Interval cancer rates are expressed in different ways in different studies and it is difficult to compare these.

In our material interval cancer rates in the first and second years after a negative screen were 18% and 54% of expected underlying incidence in the absence of screening. These figures are slightly higher than those in Sweden in the two county study (with corresponding figures for first and second years' interval cancer rates 17% and 30%) and slightly lower than in East Anglia (with corresponding figures for first and second years' interval cancer rates 24% and 59%) (Day et al. 1995).

Advanced cancers were rare, among 276 cancers only two had distant metastases, both of them in non-attenders. During the period 1985–1989, 58% of breast cancers were found as localised node negative disease in screening age groups (Dickman et al. 1999). In our study 74% of all cancers in the age groups screened were localised. Axillary metastases were present in 24% of the first and 19% of subsequent round screen-detected cancers and in 38% of interval cancers. Interval and non-attenders' cancers were larger, they more often had axillary or distant metastases and tumour grade and stage was higher than screen-detected cancers. Subsequent rounds' screen-detected cancers were more often <10mm and less often >40mm than first round screen-detected cancers. In the later rounds the radiologist had earlier mammograms to compare and this may have helped to detect smaller lesions. First round screen-detected cancer had more time to grow and, according to lower tumour stage and grade, they were not as aggressive as later round cancers. The result is similar to earlier publications (Burrel et al. 1996, Fracheboud et al. 2001).

In material II 62% of all cases were node negative and the stage distribution as a whole was more favourable than that reported by the Romagna Cancer Registry (1997). This may be due to wide use of BSE in Finland. On the other hand our TIS rate was 8%, which is much lower than the 17% reported by Silverstein in 1998 in the United States. This difference may depend on different material selection and of diagnostic criteria. The stage distribution is more favourable in screen-detected cases than in younger or older women, which is to be expected and similar to screening trials.

In our material II 1.5% of the cancers were bilateral, which is in the range of earlier publications (Bruenner et al. 1984). Men accounted for 0.5% of cancer patients, which is on the same level as in earlier reports (Stewart et al. 1997).

6.1 Sensitivity of mammography by age of the patient and density of the breast

In our study material II the sensitivity of mammography increased statistically significantly with increasing age of the patient. This result concurs with earlier studies (Barth and Prechtel 1991, Horton 1993, Lannin et al. 1993, Kerlikowske et al. 1996, Harris 1997, Kerlikowske and Barclay 1997, Muslin et al. 1998, Hider and Nicholas 1999). Age seemed to be an independent factor to the sensitivity of mammogram. In the youngest age group (<50) 17% of cancers remained undetected or looked benign in mammography. This result is similar to earlier publications (Basset et al. 1991, Lannin et al. 1993) and Baines and Miller (1997) recommended combining at least clinical examination to mammography for women aged less than 50 years.

In most studies, as also in ours, the density of breast parenchyma was defined visually with more or less subjective evaluations, but in future digital mammography equipment will make it possible to define the density more objectively (Byng et al. 1998**)** and if needed also give recommendations for treatment and follow-up (Colditz 1998).

In our material there was a substantial increase in sensitivity of mammography by decrease in the density of the breast. The density of breast decreased significantly with age, which is to be expected, and concurs with earlier papers (Stomper et al. 1990). The effect of age was larger than the effect of density on the sensitivity when considered simultanously. Hence, the effect of density could not be totally accounted for by the age of patient.

Breast parenchyma density may vary considerably in quantities in the same patient and most breasts have both totally translucent and dense areas. However, in our material in most cases the density in total breast and in tumour area was the same and a difference was found only in 12%. The sensitivity of mammography was statistically significantly higher in fatty than in dense breast. The effect of density was slightly larger using density in tumour area than overall density.

6.2 Retrospective visibility of cancer in penultimate mammograms

When breast carcinoma is evident mammographically, a corresponding lesion can be detected on a prior mammography in up to 75% of cases (Bird et al. 1992, Harvey et al. 1993, van Dijk et al. 1993). Wolverton and Sickles (1996) found most of the benign looking mammography findings to be truly benign in follow-up, but the differentiation between benign and malignant lesions is complicated by overlapping of the findings (Kopans 1998) and variation in interpretation of mammograms (Parham et al. 1996). In our studies the retrospective visibility of cancers was estimated in three different ways. Reading A: "blinded", i.e. the radiologists knowing that all patients had breast cancer, but not the exact location of the tumour, Reading B: the preoperative mammograms were shown and thus the radiologists knew the exact location of the tumour, if it was mammographically visible by the time of the diagnosis (Reading A and Reading B in Publication III) and Reading C: the radiologist who made the breast cancer diagnosis in or outside of screening estimated the retropective visibility of the tumour in earlier mammograms (Publication IV). No true blinded review was done, i.e. the reviewed mammograms were not mixed with normal screening mammograms and radiologists were aware in each reading, that all patients had been diagnosed breast cancer – they had just to find its' location. In this sense our study differs from most earlier papers, where true blinded review has been done by several radiologists individually (Harvey et

al. 1993, van Dijck et al. 1993, Jones et al. 1996, Amos et al. 2000). The reading method did influence results remarkably and by Reading A 14%, by Reading B 29% and by Reading C 33% of the tumours were classified as visible in retrospect, which is consistent with earlier publications (Moberg et al. 2000).

In Publication IV, 41 (31%) of the 130 reviewed tumours were screen-detected tumours not visible on earlier mammograms and therefore in fact early detected. All the other 89 (68%) tumours were late detected, as they were either interval cancers detected clinically or screen-detected cancers visible on earlier mammograms (with 3 exceptions, where the tumour was not included in the mammogram). Later round screen-detected cancers were significantly more often visible in retrospect (43%) than interval cancers (19%) (Publication IV). This is less than in Amos et al. (2000), who found 38% of interval cancers to be visible in retrospect. Had all the 43 visible interval and screen-detected cancers been detected by the previous screening the proportion for early detection would have been substantially higher (65% instead of 32%). If all retrospectively visible interval cancers had been diagnosed by screening, 19% of the interval cancers could have been avoided. Our result concurs with Moberg et al. (2000), who reported the number of interval cancers identified on previous screening mammograms to vary between 7% and 34% depending on the reviewing method used and the number of reviewers included to classify a case as truly detectable.

In Publication IV, 11 (26%) patients with retrospectively visible tumours had had further diagnostic assessment in an earlier screening round with a false negative result. Jones et al. (1996) found 17% false negatives at assessment.

In Publication IV there were no major differences in the histological types between the retrospectively visible and invisible cancers. Hollingsworth et al. (1993) found diffuse histology and Ma et al. (1992) lobular histology as a principal cause of false negative mammograms. Most (79%) of the interval cancers were invisible retrospectively and some of them were very fast growing with high malignant potential. Our results were consistent with previous studies (Peeters et al. 1989, Ikeda et al. 1992, Burrell et al. 1996).

Retrospectively visible carcinomas had lower tumour grade and more favourable tumour stage than retrospectively invisible carcinomas. Retrospectively visible carcinomas were less aggressive than invisible ones and most of them (33 out of 43) were found even in the next screening round and at small size (28 had ≤15 mm tumours). (Publication IV) This result is concordant with Jones et al. (1996). Screen-detected cancers were more often visible in earlier mammograms in retrospective interpretation than interval cancers, even though the time interval was longer for screen-detected cancers. The tumour size in earlier mammograms was measurable less often in

the interval than in the screen-detected cancers i.e. their tumour borders were indistinct and unspecific. (Publication IV)

An analysis of retrospectively visible lesions (Publication III) found that overlooking was more common than misinterpretation of a perceived abnormality (55/33). This is opposite to the result of Bird et al. (1992). In our study the most important reason for missing the lesion was its visibility in only one projection (27%), which is more than the 10% reported by Bird et al. (1992) and concurs with Goergen et al. (1997). The next important reason was the presence of misleading further examinations, i.e. false interpretation of further mammography work up like coned down compression views, ultrasonography and probably fine or core needle biopsies (14%). It was estimated that in 4% of cases an extra projection or some other procedure could have been helpful, which is less than the 9% extra benefit of special views for palpable lesions in Faulk and Sickes (1992). Quality of mammogram was estimated to be low in only 3% of cases, which is not much considering that our study included mammograms on a routine basis from all mammography units in the study areas.

The most common missed lesion in earlier mammograms in Reading B (Publication III) was growing density and in Reading C (Publication IV) tumour mass, next were architectural distortions and asymmetry. In Burrell et al. (1996) architectural distortion, in Coveney et al. (1994) and Harvey et al. (1993) asymmetric density (27%), in Ikeda et al. (1992) nonspecific density, in Jones et al. (1996) ill-defined opacity and in van Dijck et al. (1993) vague density was the most often missed lesion. Definitions of mammographic lesions used by different study groups are not consistent, but all such lesions could also be called asymmetric densities, which can often be detected only by reviewing or follow-up. Maes et al. (1997) found such non-specific minimal signs in approximately 10% of a normal screening population with 0.5% additional breast cancer risk. Most of these cancers were found on the next screening round still with favourable stage. 14% of the retrospectively visible lesions were microcalcifications, which is the same as in Coveney et al. (1994) and slightly less than the 18% in Bird et al. (1992).

In the youngest age group we found 4 retrospectively visible cancers. This means that 12% of younger women with previous examinations could theoretically have been diagnosed for breast cancer earlier by mammography. This percentage is much lower than the 35% in Joensuu et al. (1994) and the 36% in Lannin et al. (1993) estimated as false negatives in mammograms for young women. On the other hand, our study protocol was not identical to the earlier studies. Most radiologists were aware of the poor results of diagnosing young women in the earlier Finnish study, and we made extensive use of complementary examinations and may have been more careful when diagnosing young women in the clinical parctice. Hence the small percentage may truly indicate improvement in the routine practice.

In Publication III, 41 out of 95 retrospectively visible cancers were in fatty breast areas, 30 of those in the age group 60–77. Most (35/46) of the retrospectively visible cancers among screen aged women (50–59) were in mixed or dense breast. Bird et al. (1992) and Harvey et al. (1993) found most of the missed lesions in women with dense breasts and few of those in fatty breasts, which is contrary to our result. We estimated that a dense breast parenchyma obscured the lesion in only 13% of cases, the effect of density was less than that of age and there was no substantial difference in the sensitivity whether density was evaluated overall or at the site of tumour. This is contrary to what we could expect, since the unfavourable effect of dense breast parenchyma on mammography screening result has been earlier shown (Simpson et al. 1995, Kolb et al. 1998, Sala et al. 1998, van Gils et al. 1998).

6.3 Sensitivity of mammography vs. US and FNAB

The extra benefit of US was highest among young women (11%), as in many earlier studies (Duijm et al 1998, Zonderland et al. 1999) The extra benefit of US in mammography may further depend on the histological type of breast cancer. According to the studies by Pointon and Cunningham (1999) and Skaane and Skjorten (1999) US was more sensitive than mammography in finding invasive lobular carcinoma. Rissanen et al. (1998b), on the contrary, found US statistically not more sensitive than mammography in finding lobular invasive cancer. They found these methods complementary, especially with regard to evaluating dense breasts and lesions which were difficult to assess clinically and mammographically.

Neither mammography nor US is accurate enough for the surgeon to decide whether to perform a breast cancer operation or lumpectomy. With FNAB cytological diagnosis and with core needle biopsy (CNB) histological diagnosis of the tumour can be made before open biopsy. In this study the extra benefit to mammography was higher with FNAB than with US and it was not dependent on the age of the patient. In 7% the FNAB finding was malignant and the mammography finding was benign. This was much higher than in Westinghouse Logan-Young et al. (1992), who detected 3 cancers among 222 (1%) screened patients with benign-looking opacities in mammography and 8 unsuspected malignancies in 2 248 (0.4%) symptomatic patients with FNAB. However, in 13% the FNAB finding was insufficient or benign when the mammography finding was malignant. Insufficient FNAB samples are a major problem in breast cancer diagnostics (Teixidor et al. 1992) and this has been the reason for many breast imaging units to increase the use of CNB (Britton et al. 1997, Zonderland et al. 1998, Roth et al. 1999, Philpotts et al. 2000). In our study areas performing of CNB also began during the study period. However, there were only 25 CNB examinations with 68% sensitivity,

which is much lower than that of Lee et al. (1999), who had a false negative rate as low as 2%. Later experiences have been much better also in units of study regions and technical difficulties in beginning a new examination may explain the low sensitivity of CNB in our study material.

In our material II in Publication I there were 2 cases in which mammography, US and FNAB examinations yielded benign findings, but surgery showed the malignant nature of the tumour. Reinikainen et al. (1999) had no cancer which was negative for malignancy both in mammography, US and FNAB in their material. Watson and Given-Wilson (1999) also had no false negatives in such a triple test of palpable probably benign lumps.

New technologies like US power doppler (Reinikainen et al. 2001), three dimensional imaging, digital mammography (Weinreb and Newstead 1995, Pisano et al. 2000), computer assisted breast cancer detection (Kegelmeyer et al. 1994, Thurfjell et al. 1998) and combining MR imaging (Hall 1997, Mueller-Schimpfe et al. 1997, Sickles 2000) or scintigraphy (Özdemir et al. 1997) to mammography and US could further increase the sensitivity of breast cancer diagnostics, but we have no experiences of these.

7 Conclusions

High participation rate, moderate cancer detection rate, favourable stage distribution in screen-detected cancers and moderate interval cancer rate permit the conclusion that mammography screening has been successfully performed in the study areas.

The sensitivity of breast cancer diagnostics is important both in screening and in clinical practice. The proportion of interval cancers was 24% in material I and 27% in material II in age groups screened. This fact should also be known by clinicians, particularly because of some of them appeared in a very short time after a normal interpretation of a screening mammogram.

The sensitivity of mammography is significantly dependent on the age of the patient. The density of the breast, especially in tumour area, also influences sensitivity considerably. Retrospectively visible cancers are common. However, most of them have only unspecific signs of abnormality and can only be detected when the tumour grows or otherwise changes or by reviewing when the exact location of the tumour is known. Thus in routine mammographic practice which involves prospective interpretation it is not possible to differentiate most of them from normal parenchymal densities. If all unspecific densities were classified as screening positives the specificity of the screening programme would decrease remarkably and the costs would increase unreasonably.

Visibility of breast cancer in the earlier screening mammogram may be a favourable prognostic sign however, because such tumours are relatively low grade and therefore slow growing with low malignant potential. In fact this is what one would expect, since only slow growing cancers would have been sufficiently large to have been visible at previous screening mammography.

The failure of mammography to detect cancers in younger women may lead to great size of the tumour. US has an important role as a complementary examination to mammography, especially in younger women and those with dense breast parenchyma. With needle biopsies it is possible not only to ascertain the cytology or histology of the tumour but also to improve the sensitivity of breast cancer diagnostics in mammographically low suspect cases.

In our material II there were two cases with false negative result in mammography, US and FNAB and we advise against trusting negative findings if there is a true palpable tumour suggestive of malignancy. In the screening material there were 11 patients with false negative further assessments in earlier screening rounds. One has to take into account the possibility of false negative clinical radiology and cytology.

All diagnostic methods have their limitations which should be commonly known. The limitations of mammography should be especially familiar to radiologists and clinicians so that suitable additional examination methods can be used if necessary. Also the public should be informed about the limitations of mammography.

8 Summary

Background: Breast cancer is the most common and increasing cancer in Finland. The survival of breast cancer patients has increased remarkably in the last four decades. This has been assumed to be due to early diagnosis and in wide mammography based screening trials it has been shown that breast cancer mortality can be reduced by 30%. However, not all breast cancers are visible or remain undetected in mammogram, some of these cancers could be found by ultrasonography or the malignant character of these tumours may show up by FNAB. The sensitivity of radiological investigation methods determines how small breast cancers are found and thus how good the survival of breast cancer patients is. Some breast cancers are visible in earlier mammograms by retrospective review. The number of originally unnoticed but retrospectively visible cancers can be regarded as a sensitivity measure for the diagnostic methods in mammography screening.

Aims of the study: The general aim of the study is to evaluate the sensitivity of mammography screening in relation to determinants possibly affecting it. The more detailed aims were to evaluate the effect of age of the patient, breast parenchymal density, the tumour characteristics and the interpretation of the mammograms to the sensitivity of mammography; and to further compare sensitivity of mammography to other diagnostic methods.

Materials and methods: Material I: The clinical and radiological data of all screen-aged new breast cancer patients in the Pirkanmaa area in 1987–1992 (N = 277). The radiological and histopathological characteristics of cancers detected in and outside of screening were recorded. Material II: Clinical and radiological data of 665 consecutive new breast cancer patients in 1996–1997 in Pirkanmaa and North-Karelia. The sensitivity of mammography, US and FNAB was compared by age group and the sensitivity of mammography and US was also compared in density classes. The visibility of cancer in earlier mammograms was evaluated by either the radiologists who found the cancer (Material I) or by the groups of radiologists (Material II) in non-blinded manner.

Results: The sensitivity of mammography was heavily dependent on the age of the patient and increased with increasing age of the patient, being 83% in the youngest, 91% in middle and 96% in the oldest age group. The density of the breast, especially in the tumour area, also influences sensitivity considerably. The sensitivity of US was inversely related to age and directly related with fattiness of the breast. In material I, 33% of cancers were visible in retrospective review and in material II, 14% of cancers were visible in retrospective review without the preoperative mammograms having been

shown and 29% when the preoperative mammograms were shown. Mammography (Se = 0.92) was significantly more sensitive than US (Se = 0.86) or FNAB (Se = 0.85).

Discussion: The sensitivity of breast cancer diagnostics is very important both in screening and in clinical practice. A significant number of breast cancers was visible in earlier mammograms by retrospective review which is concordant with other studies. However, most of those are unspecific densities and they can only be detected when they grow or otherwise change.

Conclusions: Mammography was the most sensitive method but US and FNAB had an important complementary role. US was more sensitive in younger age groups than mammography and this underlines the role of US as a complementary examination. The number of retrospectively visible tumours was remarkable but in clinical practice inevitable. On the whole the limitations of mammography and other examination methods should be familiar to radiologists, clinicians and women.

9 Acknowledgements

The present study was carried out at the Tampere School of Public Health, University of Tampere and Pirkanmaa Cancer Society.

I would like to express my deepest gratitude to my supervisor, Professor Matti Hakama, Sc.D., Tampere School of Public Health, University of Tampere for his patient guidance and warm support during all stages of this study. He introduced me to scientific writing and encouraged and advised on numerous problems during the course of the research.

Sincere thanks are due to Docent Ahti Anttila, Ph.D and radiologist, Docent Tapani Tikkakoski, M.D., Ph.D for their constructive criticism and valuable advice that I appreciatively applied to improve the final outcome of this study.

This work would not have been possible without the help and support of Tiina Salminen, Ph.D. She has taught me everything I know about computers, a lot about epidemiology and helped me with data processing and scientifical writing. She really made working and discussing all sorts of problems fun and enjoyable.

I extend my wamest thanks to my friends and co-authors radiologist Alpo Kärkkäinen, M.D., Ph.D and radiologist Pirkko Peltonen M.D. and oncologist Riitta Kokko, M.D. who were involved in the planning of the study and also in data collection. The data collection was carried out as a team-work, which was much fun and hopefully taught us all a lot. I extend my wamest thanks to radiologists Ulla Geiger, M.D., Pirkko Heikkinen M.D., Seppo Hyvärinen, M.D., Ph.D., Marja-Liisa Kokko, M.D., Juhani Pakkanen, M.D., Anneli Piironen, M.D., Ph.D. Anja Salo, M.D. and Marja-Liisa Talviala, M.D, and oncologists, Professor Kaija Holli, M.D., Ph.D., Vesa Kataja, M.D., Ph.D and Eero Kumpulainen, M.D., Ph.D., and pathologist, Professor Jorma Isola M.D., Ph.D. For good advice I thank also radiologist Tarja Rissanen M.D., Ph.D. and pathologist Kalle Salo M.D.

I am thankful to Ms Marita Hallila for technical advice in the final preparation of thesis and Ms Virginia Mattila for revising the English language of the thesis.

I am very grateful to the entire staff of Pirkanmaa Cancer Society for their many kinds of help during my work. Thanks likewise to the staff of other mammography units who kindly helped by sending mammograms for reviewing and especially to Mammography Screening Unit of Cancer Society of Finland in Joensuu and Tampere City Hospital.

Warmest thanks also to the breast cancer patients who gave permission to use their patient files and mammograms as study material.

My family I want to thank for love, support and care.

The financial support of Radiological Society of Finland, the Scientific Foundation of the City of Tampere and Maija Sulamaa Foundation is gratefully acknowledged.

Tampere, September 2001

Irma Saarenmaa

10 References

Advisory Committee on Cancer Prevention (2000): Recommendations on cancer screening in the European Union. Eur J Cancer 36:1473–1478.

Alexander FE, Anderson TJ, Brown HK, Forrest APM, Hepburn W, Kirkpatrick AE, Muir BB, Prescott RJ and Smith A (1999): 14 years of follow-up from the Edinburgh randomised trial of breast-cancer screening. Lancet 353:1903–1908.

Amos AF, Kavanagh AM, Cawson J and the Radiology Quality Assurance Group of BreastScreen Victoria (2000): Radiological review of interval cancers in an Australian mammographic screening programme. J Med Screen 7:184–189.

Andersson I (2000): Mammographic screening under age 50: a review. Breast 9:125–129.

Andersson I, Aspegren K, Janzon L, Landberg T, Lindholm K, Linell F, Ljunberg O, Ranstam J and Sigfusson B (1988): Mammographic screening and mortality from breast cancer: the Malmö mammographic screening trial. Brit Med J 297:943–948

Anttinen I, Pamilo M, Soiva M and Roiha M (1993):Double reading of mammography screening films - one radiologist or two? Clin Radiol 48:414–421.

Austoker J (1994): Screening and self examination for breast cancer. Brit Med J 309:168–174.

Auvinen A, Elovainio L and Hakama M (1996): Breast self-examination and survival from breast cancer. A prospective follow-up study. Breast Cancer Res Treat 38:161-168.

Baines CJ and Miller AB (1997): Mammography versus clinical examination of the breast. J Natl Cancer Inst 22:125–129.

Barth V and Prechtel K (1991): Atlas of breast disease. B.C. Decker Inc, Philadelphia.

Barton MB, Harris R and Fletcher SW (1999): Does this patient have breast cancer? The screening clinical breast examination: should it be done? How? JAMA 282:1270–1280.

Bassett LW, Ysrael M, Gold RH and Ysrael C (1991): Usefulness of mammography and sonography in women less than 35 years of age. Radiology 180:831–835.

Berg WA and Gillbreath PL (2000): Multicentric and multifocal cancer: whole-breast US in preoperative evaluation. Radiology 214:59–66.

Berkowitz JE, Gatewood OMB and Gayler BW (1989): Equivocal mammographic findings: evaluation with spot compression. Radiology 171:369–371.

Berkowitz JE, Gatewood OMB, Goldblum LE and Gayler BW (1990): Hormonal replacement therapy: mammographic manifestations. Radiology 174:199–201.

Berlin L (2001): The missed breast cancer redux: time for educating the public about the limitations of mammography? Am J Roentgenol 176:1131–1134.

Bernardino ME (1990): Automated biopsy devices: significance and safety. Radiology 176:615–616.

Bird RE, Wallace TW and Yankaskas BC (1992): Analysis of cancers missed at screening mammography. Radiology 184:613–617.

Bjurstam N, Björneld L, Duffy SW, Smith TC, Cahlin E, Erikson O, Lingaas H, Mattson J, Persson S, Rudenstam C-M and Säwe-Söderberg J (1997): The Gothenburg breast cancer screening trial: Preliminary results on breast cancer mortality for women aged 39–49 years. J Natl Cancer Inst 22:53–55.

Blanks RG, Moss SM, McGahan CE, Quin MJ and Babb PJ (2000): Effect of NHS breast screening programme on mortality from breast cancer in England and Wales, 1990-8: comparison of observed with predicted mortality. Brit Med J 321:665–669.

Blanks RG, Moss SM and Patnick J (2000): Results from the UK NHS screening programme 1994–1999. J Med Screen 7:195–198.

Boer R, de Koning HJ, van Oortmanssen GJ and van der Maas PJ (1995): In search of the best upper age limit for breast cancer screening. Eur J Cancer 31:2040–2043.

Boyd NF, Lockwood GA, Byng JW, Tritchler DL and Yaffe MJ (1998): Mammographic densities and breast cancer risk. Cancer Epidem Biomar 7:1133–1144.

Britton PD, Flower CDR, Freeman AH, Sinnatamby R, Warren R, Goddard MJ, Wight DGD and Bobrow L (1997): Changing to core biopsy in an NHS breast screening unit. Clin Radiol 52:764–767.

Brown J, Bryan S and Warren R (1996): Mammographic screening: an incremental cost-effectiveness analysis of double versus single reading of mammograms. Brit Med J 312:809–812.

Bruenner S, Langfeld B and Andersen PE (1984): Early detection of breast cancer. Springer-Verlag. Berlin.

Bruneton JN, Caramella E, Hery M, Aubanel D, Manzino JJ and Picard JL (1986): Axillary lymph node metastases in breast cancer: preoperative detection with US. Radiology 158:325–326.

Buchberger W, DeKoekkoek-Doll P, Springer P, Obrist P and Dunser M (1999): Incidental findings on sonography of the breast: clinical significance and diagnostic workup. Am J Roentgenol 173:921–927.

Burrell HC, Sibbering DM, Wilson ARM, Pinder SE, Evans AJ, Yeoman LJ, Elston CW, Ellis IO, Blamey RW and Robertson JFR (1996): Screening interval breast cancers: mammographic features and prognostic factors. Radiology 199:811–817.

Byng JW, Yaffe MJ, Jong RA, Shumak RS, Lockwood GA, Tritchler DL and Boyd NF (1998): Analysis of mammographic density and breast cancer risk from digitized mammograms. Radiographics 18:1587–1598.

Chen H-H, Tabar L, Fagerberg G and Duffy SW (1995): Effect of breast cancer screening after age 65. J Med Screen 2:10–14.

Ciatto S, del Turco MR and Bravetti P (1989): Nonpalpable breast lesions: stereotaxic fine-needle aspiration cytology. Radiology 173:57–59.

Ciatto S, Del Turco MR, Ambrogetti D, Bravetti P, Catarzi S, Morrone D and Cariaggi MP (1997): Solid nonpalpable breast lesions. Success and failure of guided fine-needle aspiration cytology in a consecutive series of 2444 cases. Acta Radiol 38:815–820.

Colditz GA (1998): Relationship between estrogen levels, use of hormone replacement therapy and breast cancer. J Natl Cancer Inst 90:814–823.

Colditz GA, Stampfer MJ, Willett WC, Hennekens CH, Rosner B and Speizer FE (1990): Prospective study of estrogen replacement therapy and risk of breast cancer in postmenopausal women. JAMA 264:2648–2653.

Coveney EC, Geraghty JG, O'Laoide R, Hourihane JB and O'Higgins NJ (1994): Reasons underlying negative mammography in patients with palpable breast cancer. Clin Radiol 49:123–125.

Cyrlac D and Wong CH (1993): Mammographic changes in postmenopausal women undergoing hormonal replacement therapy. Am J Roentgenol 161: 1177–1183.

Day NE (1985): Estimating the sensitivity of a screening test. J Epidemiol Commun H 39:364–366.

Day NE and Miller AB (1988): Screening for breast cancer. Hans Huber Publishers. Toronto

Day N, McCann J, Camilleri-Ferrante C, Britton P, Hurst G, Cush S, Duffy S and the Quality Assurance Management Group of the East Anglian Breast Screening Programme. (1995): Monitoring interval cancers in breast screening programmes: the East Anglian experience. J Med Screen 2:180–185.

de Koning HJ (2000). Assessment of nationwide cancer-screening programmes. Lancet 355:80–81.

de Koning HJ, Fracheboyd J, Boer R, Verbeek ALM, Collette HJA, Hendriks JHCL, van Ineveld BM, de Bruyn AE and van der Maas PJ (1995): Nationwide breast cancer screening in the Netherlands: support for breast cancer mortality reduction. Int J Cancer 60:777–780.

de Koning HJ (2000). Assessment of nationwide cancer-screening programmes. Lancet 355:80–81.

de Wolfe CJM and Perry NM (eds.) (1996). European guidelines for quality assurance in mammography screening. Europe Against Cancer. European Commission, Brussels, Luxembourg.

Dean PB, Pamilo M and the Mammography Working Group of Radiological Society of Finland (1999): Screening Mammography in Finland – 1.5 million examinations with 97 percent specificity. Acta Oncol Suppl 13:47–54.

Denton ERE and Field S (1997): Just how valuable is double reporting in screening mammography. Clin Radiol 52: 466–468.

Dershaw DD, Abramson A and Kinne DW (1989): Ductal carcinoma in situ: Mammographic findings and clinical implications. Radiology 170:411–415.

Dickman PW, Hakulinen T, Luostarinen L, Pukkala E, Sankila R, Söderman B and Teppo L (1999): Survival of cancer patients in Finland 1955–1994. Acta Oncol Suppl 12:1–103.

Dilhuydy MH and Barreau B (1997): The debate over mass mammography: is it beneficial for women? Eur J Radiol 24: 86–93.

dos Santos Silva I (1999): Cancer Epidemiology. Principles and Methods. International Agency for Research on Cancer. Lyon.

Dowlatshahi K, Gent HJ, Schmidt R, Jokich PM, Bibbo M and Sprenger E (1989): Nonpalpable breast tumours: diagnosis with stereotaxic localization and fine-needle aspiration. Radiology 170:427–433.

Dowlatshahi K, Yaremko ML, Kluskens LF and Jokich PM (1991): Nonpalpable breast lesions: Findings of stereotaxic needle-core biopsy and fine-needle aspiration cytology. Radiology 181:745–750.

Dronkers DJ (1992): Stereotaxic core biopsy of breast lesions. Radiology 183:631–634.

Duffy SW, Tabar L, Fagerberg G, Gad A, Gröntoft O, South MC and Day NE (1991): Breast screening, prognostic factors and survival – results from the Swedish two county study. Br J Cancer 64:1133–1138.

Duijm LEM, Guit GL, Hendriks JHCL, Zaat JOM and Mali WPTM (1998): Value of breast imaging in women with painful breasts: observational follow-up study. Brit Med J 317: 1492–1495.
Eddy DM (1989): Screening for breast cancer. Ann Intern Med 111:389–399.
Edwards NI and Jones DA (2000): Uptake of breast cancer screening in older women. Age Ageing 29:131–135.
Egan RL (1988): Breast imaging: diagnosis and morphology of breast diseases. W.B Saunders Company, Philadelphia.
Evans WP and Cade SH (1989): Needle localization and fine-needle aspiration biopsy of nonpalpable breast lesions with use of standard and stereotactic equipment. Radiology 173:53–56.
Faulk RM and Sickles EA (1992): Efficacy of spot compression-magnification and tangential views in mammographic evaluation of palpable breast masses. Radiology 185:87–90.
Feig SA (1999): Risk, benefit and controversies in mammographic screening. RSNA Categorial course in breast imaging: 99–107.
Finnish Cancer Registry (1990): Cancer Incidence in Finland 1986. Cancer Society of Finland, Publication No.46, Helsinki.
Finnish Cancer Registry (2000): Cancer Incidence in Finland 1996 and 1997. Cancer Society of Finland, Publication No.61, Helsinki.
Fletcher SW (1997): Breast cancer screening among women in their forties: an overview of the issues. J Natl Cancer Inst 22:5–9.
Fletcher SW, Black W, Harris R, Rimer BK and Shapiro S (1993): Report of the international workshop on screening for breast cancer. J Natl Cancer Inst 85:1644–1656.
Fracheboud J, de Koning HJ, Boer R, Groenewoud JH, Verbeek ALM, Broeders MJM, van Ineveld BM, Hendriks JHCL, de Bruyn AE, Holland R and van der Maas PJ (2001): Nationwide breast cancer screening programme fully implemented in the Netherlands. Breast 10:6–11.
Franquet T, Cozcolluela R and De Miguel C (1992): Stereotaxic fine-needle aspiration of low-suspicion, nonpalpable breast nodules: valid alternative to follow-up mammography. Radiology 183:635–637.
Frisell J, Eklund G, Hellström L and Somell A (1987): Analysis of interval breast carcinomas in a randomized screening trial in Stockholm. Breast Cancer Res Treat 9:219–225.
Frisell J, Lidbrink E, Hellström L and Rutqvist L-E (1997): Follow-up after 11-years-update of mortality results in the Stockholm mammographic screening trial. Breast Cancer Res Treat 45:263–270.
Gabriel H, Wilson TE and Helvie MA (1997): Breast cancer in women 65–74 years old: earlier detection by mammographic screening. Am J Roentgenol 168:23–27.
Given-Wilson R, Blanks RG, Moss SM, Ansell J, Carter JAR, Cooke J, Dabon L, Horton P, Kissin M, Rockall L, Rust A, Smee S, Toon E and Vecchi P (1999): An evaluation of breast screening in the South Thames (West) Region of the UK NHS breast screening programme: the first 10 years. Breast 8:66–71.
Goergen SK, Evans J, Cohen GPB and MacMillan JH (1997): Charactersitics of breast carcinomas missed by screening radiologists. Radiology 204:131–135.
Gotzsche PC and Olsen O (2000): Is screening for breast cancer with mammography justifiable? Lancet 355: 129–134.
Gram IT, Funkhouser E and Tabar L (1997): The Tabar classification of mammographic parenchymal patterns. Eur J Radiol 24:131–136.
Gästrin G (1994): The mama programme for breast cancer control. Acta Universitatis Tamperensis, ser A vol 394, Tampere.
Hakama M, Elovainio L, Kajantie R and Louhivuori K (1991): Breast cancer screening as public health policy in Finland. Br J Cancer 64:962–964.
Hakama M, Pukkala E, Heikkilä M and Kallio M (1997): Effectiveness of the public health policy for breast cancer screening in Finland: population based cohort study. Brit Med J 314:864–867.
Hakama M, Pukkala E, Söderman B and Day N (1999): Implementation of screening as a public health policy: issues in design and evaluation. J Med Screen 6:209–216.
Hall FM (1997): Technologic advances in breast imaging. Current and future strategies, controversies, and opportunities. Surg Oncol Clin N Am 2:403–409.
Hann L, Ducatman BS, Wang HH, Fein V and McIntire JM (1989): Nonpalpable breast lesions: evaluation by means of fine-needle aspiration cytology. Radiology 171:373–376.
Harris R (1997): Variation of benefits and harms of breast cancer screening with age. J Natl Cancer Inst 22:139–143.
Harris VJ and Jackson VP (1989): Indication for breast imaging in women under age 35 years. Radiology 172:445–448.

Harvey JA, Fajardo LL and Innis CA (1993): Previous mammograms in patients with impalpable breast carcinoma: retrospective vs blinded interpretation. Am J Roentgenol 161:1167–1172.
Hayes MK, DeBruhl ND, Hirschowitz S, Kimme-Smith C and Bassett LW (1996): Mammographically guided fine-needle aspiration cytology of the breast: reducing the rate of insufficient specimens. Am J Roentgenol 167:381–384.
Hider P and Nicholas B (1999): The early detection and diagnosis of breast cancer: a literature review – an update. NZHTA Report 2(2).
Hilleren DJ, Andersson IT, Lindholm K and Linnell FS (1991): Invasive lobular carcinoma: mammographic findings in a 10-year experience. Radiology 178:149–154.
Hindle WH, Davis L and Wright D (1999): Clinical value of mammography for symptomatic women 35 years of age and younger. Am J Obstet Gynecol 180:1484– 1490.
Hinton CE; Grainger JM and Hinton CP (1999): FNAC in the diagnosis of symptomatic breast cancer: a reminder. Breast 8:26–27.
Hollingsworth AB, Taylor LDH and Rhodes DC (1993): Establishing a histologic basis for false-negative mammograms. Am J Roentgenol 166:643–647.
Horton DA (1993): Breast cancer screening of women 65 or older – a review of the evidence on specificity, effectiveness and compliance. Breast 2:64–66.
Hosmer DW and Lemshow S (1989): Applied logistic regression. John Wiley & Sons Inc., New York.
Ikeda DM, Andersson I, Wattsgård C, Janzon L and Linell F (1992): Interval carcinomas in the Malmö mammographic screening trial: Radiographic appearance and prognostic considerations. Am J Roentgenol 159: 287–294.
Jackman RJ, Nowels KW, Rodriquez-Soto J, Marzoni FA, Finkelstein SI and Shepard MJ (1999): Stereotactic, automated, large-core needle biopsy of nonpalpable breast lesions: false-negative and histologic underestimation rates after long-term follow-up. Radiology 210:799–805.
Jansen JThM and Zoetelief J (1997): Optimization of mammographic breast cancer screening using a computer simulation model. Eur J Radiol 24:137–144.
Joensuu H, Asola R, Holli K, Kumpulainen E, Nikkanen V and Parvinen L-M (1994): Delayed diagnosis and large size of breast cancer after false negative mammogram. Eur J Cancer 30:1299–1302.
Jones RD, Mclean L, Young JR, Simpson W and Neilson F (1996): Proportion of cancers detected at the first incident screen which were false negative at the prevalent screen. Breast 5:339–343.
Kan L, Olivotto IA, Burhenne LJW, Sickles EA and Coldman AJ (2000): Standardized abnormal interpretation and cancer detection ratios to assess reading volume and reading performance in a breast screening program. Radiology 215:563–567.
Kaufman Z, Garstin WIH, Hayes R, Michel MJ and Baum M (1991): The mammographic parenchymal patterns of women on hormonal replacement therapy. Clin Radiol 43:389–392.
Kavanagh AM, Giles GG, Mitchell H and Cawson JN (2000). The sensitivity, specificity and predictive value of screening mammography and symptomatic status. J Med Sceen 7:105–110.
Kegelmeyer WP, Pruneda JM, Bourland PD, Hillis A, Riggs MW and Nipper ML (1994): Computer-aided mammographic screening for spiculated lesions. Radiology 191:31–337.
Kelsey JL and Bernstein L (1996): Epidemiology and prevention of breast cancer. Ann Rev Publ Health 17:47–67.
Kerlikowske K and Barclay J (1997): Outcomes of modern screening mammography. J Natl Cancer Inst 22:105–111.
Kerlikowske K, Grady D, Rubin SM, Sandrock C and Ernster VL (1995): Efficacy of screening mammography. A meta-analysis. JAMA 273:149–154.
Kerlikowske K, Grady D, Barclay J, Sickles EA and Ernster V (1996): Likelihood ratios for modern screening mammography. Risk of beast cancer based on age and mammographic interpretation. JAMA 276:39–43.
Kerlikowske K, Salzmann P, Phillips KA, Cauley JA and Gummings SR (1999): Continuing screening mammography in women aged 70 to 79 years. Impact on life expectancy and cost-effectiveness. JAMA 282:2156–2163.
Kirby IB and Copeland EM (1991): The Breast: comprehensive management of benign and malignant diseases. W.B.Saunders Company, Philadelfia.
Klemi PJ, Toikkanen S, Räsänen O, Parvinen I and Joensuu H (1997): Mammography screening interval and the frequency of interval cancers in a population-based screening. Br J Cancer 75:762–766.
Kolb TM, Lichy J and Newhouse JH (1998): Occult cancer in women with dense breasts: detection with screening US – diagnostic yield and tumour characteristics. Radiology 207:191–199.
Kopans DB (1997a): An overview of the breast cancer screening controversy. J Natl Cancer Inst 22:1–3.
Kopans DB (1997b): Updated results of trials of screening mammography. Surg Oncol Clin N Am 6:233–263.

Kopans DB (1998): Breast Imaging. Lippincott-Raven publishers, Philadelphia.
Kopans DB, Meyer JE and Lindfors KK (1985): Whole-Breast US imaging: four-year follow-up. Radiology 157:505–507,
Kopans DB, Moore RH, McCarthy KA, Hall DA, Hulka CA, Whitman GJ, Slanetz PJ and Halpern EF (1996): Positive predictive value of breast biopsy performes as a result of mammography: there is no abrupt change at 50 years. Radiology 200:357–360.
Lanay M (1986): Diagnosis and differential diagnosis of breast calcifications. Springer-Verlag. Berlin.
Lannin DR, Harris RP, Swanson FH and Pories WJ (1993): Difficulties in diagnosis of carcinoma of the breast in patients less than fifty years of age. Surgery 177:457–462.
Larsson L-G, Nyström L, Wall S, Rutqvist LE, Andersson I, Bjurstam N, Fagerberg G, Frisell J and Tabar L (1996):The Swedish randomised mammography screening trials: analysis of their effect on the breast cancer related excess mortality. J Med Screen 3:129–132.
Laya MB, Gallagher JC, Schreiman JS, Larson EB, Watson P and Weinstein L (1995): Effect of postmenopausal hormonal replacement therapy on mammographic density and parenchymal pattern. Radiology 196:433–437.
Le Gal M, Ollivier L, Asselain B, Meunier M, Laurent M, Vielh P and Neuenschwander S (1992): Mammographic features of 455 invasive lobular carcinomas. Radiology 185:705–708.
Lee CH, Philpotts LE, Horvath LJ and Tocino I (1999): Follow-up of breast lesions diagnosed as benign with stereotactic core-needle biopsy: frequency of mammographic change and false-negative rate. Radiology 212:189–194.
Liberman L, LaTrenta LR, Samli B, Morris EA, Abramson AF and Dershaw DD (1996): Overdiagnosis of medullary carcinoma: a mammographic-pathologic correlative study. Radiology 210:443–446.
Linver MN and Paster SB (1997): Mammography outcomes in a practice setting by age. Prognostic factors, sensitivity and positive biopsy rate. J Natl Cancer Inst 22:113–117.
Linver MN, Paster SB, Rosenberg RD, Key CR, Stidley CA and King WV (1992): Improvement in mammography interpretation skills in a community radiology practice after dedicated teaching courses: 2-year medical audit of 38,633 cases. Radiology 184:39–43.
Litherland JC, Evans AJ and Wilson ARM (1997): The effect of hormone replacement therapy on recall rate in the national health service breast screening programme. Clin Radiol 52:276–279.
Litherland JC, Stallard S, Hole D and Cordiner C (1999): The effect of hormone replacement therapy on the sensitivity of screening mammograms. Clin Radiol 54:285–288.
Logan-Young WW, Hoffman NY and Janus JA (1992): Fine-needle aspiration cytology in the detection of breast cancer in nonsuspicious lesions. Radiology 184:49–53.
Ma L, Fishell E, Wright B, Hanna W, Allan S and Boyd NF (1992): Case-control study of factors associated with failure to detect breast cancer by mammography. J Natl Cancer Inst 84:781–785.
Maes RM, Dronkers DJ, Hendriks JHCL, Thijssen MAO and Nab HW (1997): Do non-specific minimal signs in biennal mammographic breast cancer screening programme need further diagnostic assessment? Br J Radiol70: 34–38.
Madjar H, Rickard M, Jellins J, Otto R and the IBUS International Faculty (1999): IBUS guidelines for the ultrasonic examination of the breast. Eur J Ultras 9:99–102.
McNicholas MMJ, Heneghan JP, Milner MH, Tunney T, Hourihane JB and MacErlaine DP (1994): Pain and increased mammographic density in women receiving hormone replacement therapy: a prospective study. Am J Roentgenol:163:311–315.
Meyer JE, Amin E, Lindfors KK, Lipman JC, Stomper PC and Genest D (1989): Medullary carcinoma of the breast: mammographic and US appearance. Radiology 170:79–82.
Miller AB, Baines CJ, To T and Wall C (1992a): Canadian national breast screening study: 1. Breast cancer detection and death rates among women aged 40 to 49 years. Can Med Assoc J 147:1459–1476.
Miller AB, Baines CJ, To T and Wall C (1992b): Canadian national breast screening study: 2. Breast cancer detection and death rates among women aged 50 to 59 years**.** Can Med Assoc J 147:1477–1488.
Moberg K, Grundström H, Lundquist H, Svane G, Havervall E and Muren C (2000): Radiological review of incidence breast cancers. J Med Screen 7:177–183.
Moss HA, Britton PD, Flower CDR, Freeman AH, Lomas DJ and Warren RML (1999): How reliable is modern breast imaging in differentiating benign from malignant breast lesions in the symptomatic population? Clin Radiol 54:676–682.
Motulsky H (1995): Intuitive biostatistics. Oxford University Press, New York, NY.
Mucci B, Athey G and Scarisbrick G (1999): Double reading of screening mammograms: the use of a third reader to arbitrate on disagreements. Breast 8:63–65.

Mueller-Schimpfle M, Stoll P, Stern W, Kurz S, Dammann F and Clausen CD (1997): Do mammography, sonography, and MR mammography have a diagnostic benefit compared with mammography and sonography? Am J Roentgenol 168:1323–1329.

Mushlin AI, Kouides RW and Shapiro DE (1998): Estimating the accuracy of screening mammography: a meta-analysis. Am J Prev Med 14:143–153.

Newstead GM, Baute PB and Toth HK (1992): Invasive lobular and ductal carcinoma: mammographic findings and stage at diagnosis. Radiology 184:623–627.

Nyström L, Rutqvist LE; Wall S, Lindgren A, Lindqvist M, Ryden S, Andersson I, Bjurstam N, Fagerberg G, Frisell J, Tabar L and Larsson L-G (1993): Breast screening with mammography: overview of Swedish randomised trials. Lancet 341:973–978.

Özdemir A, Öznur II, Vural G, Atasever T, Karabacak NI, Gökcora N, Isik S and Unlu M (1997): TL-201 scintigraphy, mammography and ultrasonography in the evaluation of palpable and nonpalpable breast lesions: a correlative study. Eur J Radiol 24:145–154.

Pamilo M, Soiva M and Lavast E-M (1989): Real–time ultrasound, axillary mammography, and clinical examination in the detection of axillary lymph node metastases in breast cancer patients. J Ultras Med 8: 115–120.

Pamilo M, Dean PB, Räsänen O and Mammography working group of Radiological Society of Finland (1993): Mammographic screening for breast cancer: four years of nationwide screening in Finland. Eur Radiol 3:44–45.

Parham DM, Creagh-Barry M, Hill RA and Nicholas DS (1996): Observer variation in the grading of screen-detected mammographic abnormalities. 1. Assessment of reproducibility. Breast 5:422–424.

Peeters PHM, Verbeek ALM, Hendriks JHCL, Holland R, Mravunac M and Vooijs GP (1989): The occurence of interval cancers in the Nijmegen screening programme. Br J Cancer 59:929–932.

Persson I, Thurfjell E and Holmberg L (1997): Effect of estrogen and estrogen-progestin replacement regimens on mammographic breast parenchymal density. J Clin Oncol 15:3201–3207.

Philpotts LE, Shaheen NA, Jain KS, Carter D and Lee CH (2000): Uncommon high-risk lesions of the breast diagnosed at stereotactic core-needle biopsy: clinical importance. Radiology 216:831–837.

Pisano ED, Fajardo LL, Tsimikas J, Sneige N, Frable WJ, Gatsonis CA, Evans WP, Tocino I and McNeil BJ and the RDOG5 Investigators (1998): Rate of insufficient samples for fine-needle aspiration for nonpalpable breast lesions in a multicenter clinical trial. Cancer 82:679–688.

Pisano ED, Cole EB, Major S, Zong S, Hemminger BM, Muller KE, Johnston RE, Walsh R, Conant E, Fajaro LL, Feig SA, Nishikawa RM, Yaffe MJ, Williams MB, Aylward SR and the International Digital Mammography Development Group (2000): Radiologists' preferences for digital mammographic display. Radiology 216:820–830.

Pisano ED, Fajardo LL, Caudry DJ, Sneige N, Frable WJ, Berg WA, Tocino I, Schnitt SJ, Connolly JL, Gatsonis CA, McNeil BJ and the Radiologist Investigators of the Radiologic Diagnostic Oncology Group V (2001): Fine-Needle Aspiration Biopsy of Nonpalpable Breast Lesions in a Multicenter Clinical Trial: Results from the Radiologic Diagnostic Oncology Group V. Radiology 219:785–792.

Pointon KS and Cunningham DA (1999): Ultrasound findings in pure invasive lobular carcinoma of the breast: comparison with matched cases of invasive ductal carcinoma of the breast. Breast 8:188–190.

Poplack SP, Tosteson AN, Grove MR, Wells WA and Carney PA (2000): Mammography in 53,803 women from the New Hampshire mammography network. Radiology 217:832–840.

Rahbar G, Sie AC, Hansen GC, Prince JS, Melany ML, Reynolds HE, Jackson VP, Sayre JW and Bassett LW (1999): Benign versus malignant solid breast masses: US differentiation. Radiology 213:889–894.

Reinikainen HT, Rissanen TJ, Piippo UK and Päivänsalo MJ (1999): Contribution of ultrasonography and fine needle aspiration cytology to the differential diagnosis of palpable solid breast lesions. Acta Radiol 40:383–389.

Reinikainen H, Rissanen T, Päivänsalo M, Pääkkö E, Jauhiainen J and Suramo I (2001): B-mode, power doppler and contrast-enhanced power doppler ultrasonography in the diagnosis of breast tumours. Acta Radiol 42:106–113.

Rissanen T (1994): Mammography and breast ultrasound. Acta Universitatis Ouluensis, series D medica 291, Oulu.

Rissanen TJ, Apaja-Sarkkinen MA, Mäkäräinen HP and Heikkinen MI (1997): Ultrasound-guided fine needle aspiration biopsy in the diagnosis of breast cancer recurrence after mastectomy. Acta Radiol 38:232–239.

Rissanen T, Pamilo M and Suramo I (1998a): Ultrasonography as guidance method in the evaluation of mammographically detected nonpalpable breast lesions of suspected malignancy. Acta Radiol 39:292–297.

Rissanen T, Tikkakoski T, Autio A-L and Apaja-Sarkkinen M (1998b): Ultrasonography of invasive lobular breast carcinoma. Acta Radiol 39:285–291.

Roberts MM, Alexander FE, Andersson TJ, Chetty U, Donnan PT, Forrest P, Hepburn W, Huggins A, Kirkpatrick AE, Lamb J, Muir BB and Prescott RJ (1990): Edinburgh trial of screening for breast cancer: mortality at seven years. Lancet 335:241–246.

Robidoux MA, Helvie MA, Wilson TE, Lai NE and Paramagul C (1997): Women with breast cancer: histologic findings in the contralateral breast. Radiology 203.691–694.

Roche NA, Given-Wilson RM, Thomas VA and Sacks NPM (1998): Assessment of a scoring system for breast imaging. Br J Surg 85:669–672.

Romagna Cancer Registry and Collaborators (1997): Breast carcinoma stage in relation to time interval since last mammography. Cancer 80:1432–1437.

Rosen EL, Sickles E and Keating D (1999): Ability of mammography to reveal nonpalpable breast cancer in women with palpable breast masses. Am J Roentgenol 172:309–312.

Roth W-D, von Smitten K, Heikkilä P, Edgren J and Laasonen L (1999): Automated streotactic core needle biopsy of microcalcifications with correlation to surgical biopsy. Acta Radiol 40:390–393.

Rothman LP, Harvald TB and Overgaard H (1996): Impalpable breast lesions diagnosed by mammography: incidence of histologically confirmed malignancy. Eur J Surg 162:691–694.

Rubin E and Farber JL (1988) Pathology. J.B. Lippincott Company, Philadelphia.

Rumack CM, Wilson SR and Charboneau JW (1991): Diagnostic ultrasound. Mosby-Year Book, Inc. St:Louis, Missouri.

Sala E, Warren R, McCann J, Duffy S, Day N and Luben R (1998): Mammographic parenchymal patterns and mode of detection: implications for the breast screening programme. J Med Screen 5: 207–212.

Salminen TM, Saarenmaa IE, Heikkilä MM and Hakama M (1998a): Risk of breast cancer and changes in mammographic parenchymal patterns over time. Acta Oncol 37:547–551.

Salminen T, Hakama M, Heikkilä M and Saarenmaa I. (1998b): Favourable change in mammographic parenchymal patterns and breast cancer risk factors. Int J Cancer 78:410–414.

Salminen TM, Saarenmaa IE, Heikkilä MM and Hakama M. (1999): Unfavourable change in mammographic patterns and the breast cancer risk factors. Breast Cancer Res Treat 57:165–173.

Salminen TM, Saarenmaa IE, Heikkilä MM and Hakama M. (2000): Is dense mammographic parenchymal pattern a contraindication to hormonal replacement therapy? Acta Oncol 39:969–972.

Scaf-Klompf W, van Sonderen FLP, Stewart R, van Dijck JAAM and van den Heuvel WJA (1995): Compliance after 17 years of breast cancer screening. J Med Screen 2:195–199.

Schairer C, Lubin J, Troisi R, Sturgeon S, Brinton L and Hoover R (2000): Menopausal estrogen and estrogen-progesterin replacement therapy and breast cancer risk. JAMA 283: 485–491

Screening mammography re-evaluated (2000): Correpondece. Lancet 355:747–752.

ShapiroS, Venet W, Strax P, Venet L and Roeser R (1982): Ten- to fourteen-year effect of screening on breast cancer mortality. J Natl Cancer Inst 69:349–355.

Shapiro S, Coleman EA, Broeders M, Codd M, Codd M, de Koning H, Fracheboyd J, Moss S, Paci E, Stachenko S, Ballard-Barbash R, the International Breast Cancer Screening Network (IBSN) and the European Network of Pilot Projects for Breast Cancer Screening (1998): Breast Cancer Screening programmes in 22 countries: current policies, administration and guidelines. Int J Epidemiol:27:735–742.

Sickles EA (1991): Periodic mammographic follow-up of probably benign lesions: results in 3,184 consecutive cases. Radiology 179:463–468.

Sickles EA (2000): Breast imaging: from 1965 to the present. Radiology 215:1–16.

Silverstein MJ (1998): Ductal carcinoma in situ of the breast. Brit Med J 317:734–739.

Simpson W, Neilson F,Young JR and the Northern Region Breast Screening Radiology Audit Group (1995): The identification of false negatives in a population of interval cancers: a method for audit of screening mammography. Breast 4:183–188.

Skaane P and Skjorten F (1999): Ultrasonographic evaluation of invasive lobular carcinoma. Acta Radiol 40:369–375.

Skaane P, Engedal K and Skjennald A (1997): Interobserver variation in the interpretation of breast imaging. Comparison of mammography, ultrasonography, and both combined in the interpretation of palpable noncalcified breast masses. Acta Radiol 38:497–502.

Skaane P, Olsen JB, Sager EM, Abdelnoor M, Berger A, Kullmann G and Wolff PA (1999): Variability in the interpretation of ultrasonography in patients with palpable noncalcified breast tumours. Acta Radiol 40:169–175.

Soini I and Hakama M (1978): Failure of selective screening for breast cancer by combining risk factors. Int J Cancer 22:275–281.

Solbiati L, Rizzatto G and Charboneau JW (1995): Ultrasound of superficial structures. High frequencies, doppler and interventional procedures. Churchill Livingstone, Edinburgh.

Sonnenfeld MR, Frenna TH, Weidner N and Meyer JE (1991): Lobular carcinoma in situ: mammographic-pathologic correlation of resuts of needle-directed biopsy. Radiology 181:363–367.

Spicer DV, Ursin G, Parisky YR, Pearse JG, Shoupe D, Pike A and Pike MC (1994): Changes in mammographic densities induced by a hormonal contraceptive designed to reduce breast cancer risk.J Natl Cancer Inst 86:431–436.

Stewart RAL, Howlett DC and Hearn FJ (1997): Pictorial review: the imaging features of male breast disease. Clin Radiol 52:739–744.

Stomper PC, Davis SP, Weidner N and Meyer JE (1988). Clinically occult, noncalcified breast cancer: serial radiologic pathologic correlation in 27 cases. Radiology 169:621–626.

Stomper PC, Van Voorhis BJ, Ravnikar VA and Meyer JE (1990): Mammographic changes associated with postmenopausal hormone replacement therapy: a longitudinal study. Radiology 174:487–490.

Stomper PC, D'Souza DJ, DiNitto PA and Arredondo MA. (1996): Analysis of parenchymal density on mammograms in 1353 women 25–79 years old. Am J Roentgenol 167:1261–1265.

Svensson WE, Tohno E, Cosgrove DO, Powles TJ, Murrani BA and Jones AL (1992): Effects of fine-needle aspiration on the US appearance of the breast. Radiology 185:709–711.

Tabar L and Dean PB (1982): Mammographic parenchymal patterns. Risk indicator for breast cancer? JAMA 247:185–198.

Tabar L and Dean PB (1983): Teaching Atlas of Mammography. Georg Thieme Verlag, Stuttgart.

Tabar L, Gad A, Holmberg LH, Ljunquist U, Eklund G, Fagerberg CJG, Baldetorp L, Gröntoft O, Lundström B, Månson JC Day NE and Pettersson F (1985): Reduction in mortality from breast cancer after mass screening with mammography. Lancet I:829–832.

Tabar L, Fagerberg G, Duffy SW, Day NE, Gad A and Gröntoft O (1992): Update of the Swedish two-county program of mammographic screening for breast cancer. Radiol Clin N Am 30:187–210.

Tabar L, Fagerberg G, Chen H-H, Duffy SW, Smart CR, Gad A and Smith RA (1995): Efficacy of breast cancer screening by age. New results from the Swedish two-county trial. Cancer 75:2507–2517.

Tabar L, Chen H-H, Fagerberg G, Duffy SW and Smith TC (1997): Recent results from the Swedish Two-County trial: the effects of age, histologic type, and mode of detection on the efficacy of breast cancer screening. J Natl Cancer Inst 22:43–47.

Tabar L, Vitak B, Chen H-HT, Yen M-F, Duffy SW and Smith RA (2001): Beyond randomized controlled trials. Organized mammographic screening substantially reduces breast carcinoma mortality. Cancer 91:1724–1731.

Teh W and Wilson ARM (1998): The role of ultrasound in breast cancer screening. A consensus statement by the European Group for Breast Cancer Screening. Eur J Cancer 34:449–450.

Teixidor HS, Wojtasek DA, Reiches EM, Santos-Buch CA and Minick CR (1992): Fine-needle aspiration of breast biopsy specimens: correlation of histologic and cytologic findings. Radiology 184:55–58.

Threatt B, Norbeck JM, Ullman NS, Kummer R and Roselle P (1980): Association between mammographic parenchymal pattern classification and incidence of breast cancer. Cancer 45:2550–2556.

Thurfjell E (1994): Mammography screening. One versus two views and independent double reading. Acta Radiol 35:345–350.

Thurfjell EL and Lindgren JAÅ (1996): Breast cancer survival rates with mammographic screening: similar favorable survival rates for women younger and those older than 50 years. Radiology 201:421–426.

Thurfjell EL, Holmberg LH and Persson IR (1997): Screening mammography: sensitivity and specifity in relation to hormone replacement therapy. Radiology 203:339–341.

Thurfjell E, Thurfjell MG, Egge E and Bjurstam N (1998): Sensitivity and specifity of computer assisted breast cancer detection in mammography screening. Acta Radiol 39:384–388.

UK Trial of Early Detection of Breast Cancer group (1999): 16-year mortality from breast cancer in the UK trial of early detection of breast cancer. Lancet 353:1909–1914.

van Dijck JAAM, Verbeek ALM, Hendriks JHCL and Holland R (1993): The current detectability of breast cancer in a mammographic screening program. A review of the previous mammograms of interval and screen-detected cancers. Cancer 72:1933–1938.

van Dijck JAAM, Verbeek ALM, Beex LVAM, Hendriks JHCL, Holland R, Mravunac M, Straatman H and Werre JM (1996): Mammographic screening after the age of 65 years: evidence for a reduction in breast cancer mortality. Int J Cancer 66:727–731.

van Gils CH, Otten JDM, Verbeek ALM, Hendriks JHCL and Holland R (1998): Effect of mammographic breast density on breast cancer screening performance: a study in Nijmegen, the Netherlands. J Epidemiol Commun H 52:267–271.

Varas X, Leborgne F and Leborgne JH (1992): Nonpalpable, probably benign lesions: role of follow-up mammography. Radiology 184:409–414.

Warner E, Lockwood G, Tritchler D and Boyd NF (1992): The risk of breast cancer associated with mammographic parenchymal patterns: a meta-analysis of the published literature to examine the effect of method of classification. Cancer Detect Prev 16:67–72.

Watson GMT and Given-Wilson RM (1999): Probably benign palpable lumps in women over 50 – is surgery the only option? Breast 8:48–50.

Weinreb JC and Newstead G (1995): MR imaging of the breast. Radiology 196:593–610.

Whitehead J, Carlile T, Kopecky KJ, Thompson DJ, Gilbert FI, Present AJ, Threatt BA, Krook P and Hadaway E (1985): Wolfe mammographic parenchymal patterns. A study of the masking hypothesis of Egan and Mosteller. Cancer 56:1280–1286.

Wolfe JN (1976a): Risk for breast cancer development determined by mammographic parenchymal pattern. Cancer 37:2486–2492.

Wolfe JN (1976b): Breast patterns as an index of risk for developing breast cancer. Am J Roentgenol 126:1130–1139.

Wolfe JN, Saftlas AF and Salane M (1986): Mammographic parenchymal patterns and quantitative evaluation of mammographic densities: a case-control study. Am J Roentgenol 148:1087–1092.

Wolverton DE and Sickles EA (1996): Clinical outcome of doubtful mammographic findings. Am J Roentgenol 167:1041–1045.

Woodman CBJ, Threlfall AG, Boggis CRM and Prior P (1995): Is the three year breast screening interval too long? Occurence of interval cancers in NHS breast screening programme's north western region. Brit Med J 310:224–226.

Zonderland HM, Hermans J, van de Vijver MJ, Postema S, Tollenaar RAEM and van de Velde CJH (1998): Triple diagnostic approach versus ultrasound-guided 18 gauge core biopsy in suspicious breast masses. Breast 7:168–172.

Zonderland HM, Goerkamp EG, Hermans J, van de Vijver MJ and Voorthuisen AE (1999): Diagnosis of breast cancer: contribution of US as an adjuct to mammography. Radiology 213:413–422.

Original publications