Effect of mammographic breast density on breast cancer screening performance: a study in Nijmegen, the Netherlands

Carla H van Gils, Johannes D M Otten, André L M Verbeek, Jan H C L Hendriks, Roland Holland

Abstract

Study objective—To study the implications of breast density on mammographic screening performance.

Design—Screening outcomes of women with dense breast patterns were compared with those of women with lucent breast patterns (dense > 25% densities, lucent ≤ 25% densities); the women were screened in different periods (before/after improvement of the mammographic technique in 1982).

Setting—Nijmegen, the Netherlands, 1977–1994.

Participants—Between 1977 and 1994, 73 525 repeat screenings were performed in 19 152 participants (aged 50–69 years) in the Nijmegen breast cancer screening programme (repeat screenings were defined as mammographic examinations that were preceded by an examination in the previous screening round). Participants were screened biennially with mammography. There were 258 screen detected and 145 interval cancers.

Main results—Before 1982 (rounds 2–4) the predictive value of a positive screening test (PV+) was lower in women with dense breasts than in those with lucent breasts (dense 29% *v* lucent 52%, p=0.003). Also, the ratio of screen detected cancers to the total number of screen detected plus interval cancers (as a proxy for sensitivity) was lower in this group (based on a one year interval: dense 63% *v* lucent 92%, p=0.001 and based on a two year interval: dense 41% *v* lucent 68%, p=0.002). Moreover, the survival rate was less favourable for those with dense breasts (p=0.07). In rounds 5–10, there were no important differences with respect to PV+ (dense 66% *v* lucent 62%, p=0.57) or survival (p=0.48). Moreover, sensitivity based on a one year interval was nearly as high in women with dense breasts as in those with lucent breasts (85% *v* 86%, p=0.75). However, based on a two year interval sensitivity was lower (dense 59% *v* lucent 72%, p=0.04).

Conclusions—In the early screening years (rounds 2–4) high breast density had an unfavourable effect on screening performance. Nowadays, the situation has improved with respect to PV+, survival and detecting tumours in dense breasts with a lead time of up to one year, but little improvement has occurred in the detection of tumours with a lead time greater than one year.

(*J Epidemiol Community Health* 1998;**52**:267–271)

Department of Epidemiology, University of Nijmegen, the Netherlands
C H van Gils
J D M Otten
A L M Verbeek

National Expert and Training Centre for Breast Cancer Screening, University Hospital, Nijmegen, the Netherlands
J H C L Hendriks
R Holland

Correspondence to:
Carla H van Gils, Department of Epidemiology, University of Nijmegen, PO Box 9101, 6500 HB Nijmegen, the Netherlands.

Accepted for publication 5 June 1997

Mammographic screening for breast cancer in women aged 50–69 years has been shown to lead to a better disease stage distribution at diagnosis and a subsequent reduction in breast cancer mortality.[1–5] However, many cancers still escape detection. High mammographic breast density may partly account for these "missed" carcinomas,[6–9] because dense fibro-glandular tissue has *x* ray attenuation properties similar to those of breast lesions. Uncertainty about the presence of breast cancer because of high breast density[10] may also lead to more women being unnecessarily referred and given a biopsy.[11–13]

Both potential effects of high breast density (low sensitivity and low predictive value of referral) are highly undesirable. Dependent on the magnitude of the problem, efforts should be made to increase the effectiveness of screening in women with dense breast patterns. Possible solutions could vary from "simply" taking additional mammographic views, to the use of other techniques such as digital mammography.[14] Also, more frequent screening may be advantageous for women with dense breasts, because it increases the probability of screen detection if a tumour is present.[15]

The full extent of the problem in screening practice, however, and the consequences for the women concerned are not yet clear. In this study, early indicators of the effectiveness of screening (sensitivity, positive predictive value, stage distribution of cancers), as well as breast cancer specific survival rates were evaluated in a longstanding screening programme in women with dense breast patterns and in women with lucent breast patterns. This analysis was based on the results of the biennial screening programme in Nijmegen, the Netherlands (1975–1994). A distinction was made between the early (1975–1982) and later screening rounds (1983–1994), to see whether the problem of detecting tumours in dense breasts was merely inherent to mid-1970s mammography or whether it also applies to current high-quality mammography.

Table 1   Number of screened and referred women and number of screen detected and interval cancers

	Screening rounds 2–4		Screening rounds 5–10	
	Dense pattern	Lucent pattern	Dense pattern	Lucent pattern
Referred	69	116	91	189
Cancer at screening*	20 (2)	60 (12)	60 (14)	118 (15)
Interval cancer†				
≤1 y	12 (0)	5 (0)	11 (0)	19 (1)
>1 and ≤2 y	17 (3)	23 (1)	30 (2)	28 (1)
PV+ Referral %‡	29.0	51.7	65.9	62.4
Screen/(Screen+Inter) %§				
≤1 y	62.5	92.3	84.5	86.1
total (≤2 y)	40.8	68.2	59.4	71.5

*In parentheses, the number of ductal carcinomas in situ out of the total number of cancers. †Separate data for: diagnosis within one year after a negative screening examination and for: diagnosis after one year, but within two years after a negative screening examination. ‡PV+=positive predictive value. §Ratio of the number of screen detected cancers to the number of screen detected plus interval cancers. Separate data for: only interval cancers diagnosed within one year after a negative screening examination included and for: all interval cancers included.

Methods

In Nijmegen, the Netherlands, a biennial mammographic screening programme for breast cancer was started in 1975. Since then, more than 40 000 women aged 35 years and older have been invited to participate. By the end of 1994, 10 screening rounds had been carried out. More details of the programme have been published elsewhere.[16]

For this study, we used data on women aged 50–69 years at the time of examination, because mammographic screening is widely accepted for this age group. We focused on "repeat screening examinations", defined as mammographic examinations that were preceded by an examination in the previous screening round. This was done to prevent the results being influenced by tumours with a relatively long preclinical phase, which are overrepresented at the initial screening examination, as well as at other screening examinations that were not preceded by an examination in the previous screening round.

Between 1977 and 1994, 73 525 repeat screening examinations were performed on 19 152 women aged 50–69 years. This study includes all cancers detected at repeat screening (screen detected cancers, n=258) or diagnosed in the interval between a negative repeat screening examination and the subsequent scheduled examination (interval cancers, n=145). Patients with lobular carcinoma in situ were not included.

The standard examination included single view mammography. Initially, the direction of the view was lateral, but halfway through the fourth screening round (1982) this was changed to the mediolateral-oblique view.[17] At about the same time, the mammographic technique was improved considerably by the implementation of a General Electric (CGR) 500T and the use of an anti-scatter grid. This resulted in better contrast, especially for the mammograms of women with dense breast patterns. As these changes could have had a considerable influence on the detectability of tumours,[11 18] all the analyses were performed separately for the screening rounds 2–4 and 5–10.

Breast density was assessed on the screening mammograms. The radiologist (JH) classified
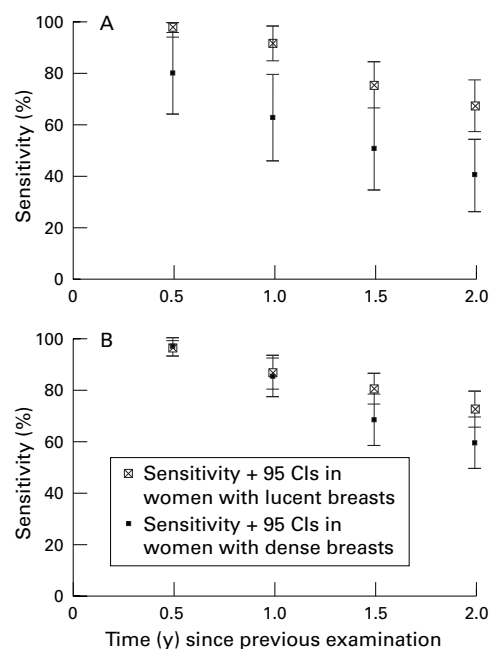

Figure 1   (A) Breast pattern specific sensitivity in screening rounds 2–4 for various time periods since the previous examination. (B) Breast pattern specific sensitivity in screening rounds 5–10 for various time periods since the previous examination.

breast patterns on a two category scale, depending on the relative amount of parenchymal density that was visible on the mammogram; ≤ 25% was defined as lucent, while > 25% was defined as dense.

Several indicators of screening effectiveness were investigated. The positive predictive value of referral was assessed by dividing the number of screen detected cancers by the number of referrals. As a proxy for sensitivity, the proportion of screen detected cancers in the total number of screen detected cancers plus interval cancers was used. A one year interval is often chosen to estimate sensitivity, but this is a quite arbitrary choice. Therefore, we estimated sensitivity by using various definitions of interval period—that is, 0.5, 1, 1.5, and 2 years after a negative screening examination.

Breast patterns were only routinely available in women who were referred for further examination (irrespective of the final result) and in women with an interval cancer. Therefore, we were unable to compute the specificity of the screening test.

Other early indicators of screening effectiveness were the distributions of tumour diameter, axillary lymph node status, and disease stage at diagnosis. We studied these distributions in breast cancer patients who were screen detected during a repeat screening examination. Diameters of invasive tumours were determined histologically. There appeared to be a tendency to round measurements off to the nearest 0.5 cm. We therefore classified any tumours with a diameter of ≤ 12 mm as "≤ 1 cm", tumours with a diameter of 13–22 mm as "1.5–2 cm", and tumours with a diameter of ≥ 23 mm as "≥ 2.5 cm".[19]

Axillary lymph nodes have only been examined routinely since 1981. Therefore, results

Table 2   Cancer staging characteristics

	Screening rounds 2–4		Screening rounds 5–10	
	Dense pattern	Lucent pattern	Dense pattern	Lucent pattern
No of cancers★				
In situ	2 (10)	12 (20)	14 (23)	15 (13)
Invasive	18 (90)	48 (80)	46 (77)	103 (87)
Total	20	60	60	118
Invasive tumour diameter				
≤1 cm	7 (41)	26 (57)	13 (28)	45 (44)
1.5–2 cm	7 (41)	14 (30)	22 (48)	37 (36)
≥2.5 cm	3 (18)	6 (13)	11 (24)	20 (20)
Total†	17[1]	46[2]	46	102[1]
Median diameter (mm)	14	12	16	14
(interquartile range)	(6–20)	(9–19)	(12–22)	(10–21)
Axillary lymph node involvement				
Negative node‡	—	—	50 (86)	84 (72)
Positive node	—	—	8 (14)	33 (28)
Total†			58[2]	117[1]
Stage				
0	—	—	14 (24)	15 (13)
I	—	—	29 (50)	59 (59)
II+	—	—	15 (26)	43 (43)
Total†			58[2]	117[1]

Values in columns=number of cases (%). ★Only screen detected patients. †Superscript numbers indicate the number of invasive cancers with an unknown size and/or unknown axillary lymph node status. ‡Ductal carcinomas in situ have been included as node negative.

concerning axillary lymph node status and stage have been restricted to the later screening rounds (5–10). The stage of the disease was determined by the tumour size and the axillary lymph node status (TNM). Stage 0 comprised ductal carcinoma in situ. Stage I tumours were no larger than 2 cm, with negative lymph nodes. Tumours of larger than 2 cm, or those with lymph node involvement were classified as stage II+.

Breast cancer specific survival curves, stratified by breast pattern, were computed for breast cancer patients who where screen detected during a repeat screening examination. The vital status of the patients was obtained from the local council registry. All clinical information of deceased patients was collected to determine the cause of death.

Differences in proportions between women with dense patterns and lucent patterns were tested by χ^2 tests for contingency tables. To indicate the variability in estimates of sensitivity we computed 95% confidence intervals for proportions. Differences in survival curves, which were obtained with the Kaplan-Meier method, were tested by the log rank χ^2 test.

Results
In the period 1977–1994, a total of 73 525 repeat screening examinations were performed on 19 152 women aged 50–69 years. Table 1 shows the screening outcomes in two screening periods (rounds 2–4 and rounds 5–10) and the effect measures calculated from these outcomes, classified by breast density.

PREDICTIVE VALUE AND SENSITIVITY
In the early screening period (rounds 2–4) the positive predictive value of referral was much higher in the women with lucent breasts than in those with dense breasts (lucent 51.7% *v* dense 29.0%, p=0.003). However, during the course of the programme, a sharp increase was seen in the positive predictive value, which ended in equal values in the women with lucent breasts
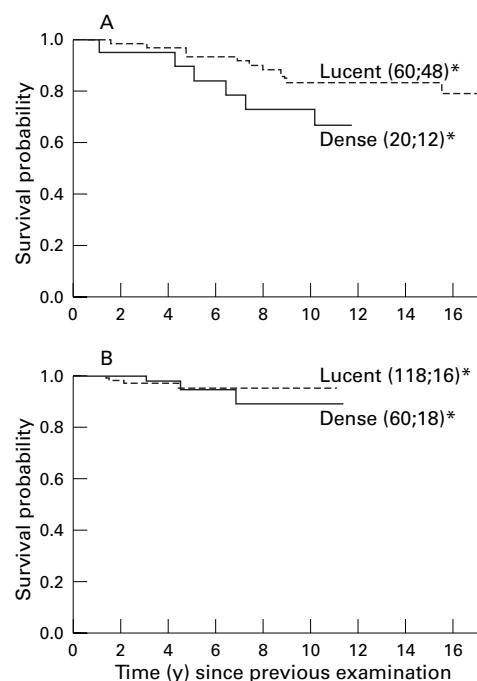

Figure 2   (A) Breast cancer survival by breast pattern in patients diagnosed in screening rounds 2–4. (B) Breast cancer survival by breast pattern in patients diagnosed in screening rounds 5–10. ★ Number of cases at the start and after 10 years of follow up.

and in those with dense breasts (lucent 62.4% *v* dense 65.9%, p=0.57).

In screening rounds 2–4, the ratio of the number of screen detected cancers to the total number of screen detected plus interval cancers (as a proxy for sensitivity, with interval cancers diagnosed within two years of a negative examination regarded as false-negative) was much higher in the women with lucent breasts than in those with dense breasts (dense 40.8% *v* lucent 68.2%, p=0.002). In screening rounds 5–10, the ratios increased, both in the women with dense breasts and in those with lucent breasts, but there was still a difference (dense 59.4% *v* lucent 71.5%, p=0.04).

Figure 1A and 1B shows the estimates of sensitivity with 95% confidence intervals for various definitions of interval period in rounds 2–4 and 5–10, respectively. In rounds 2–4 (fig 1A) the sensitivity in the women with dense breasts was lower from the beginning of the interval. On the basis of a one year interval in rounds 2–4, the sensitivity in the women with dense breasts was 62.5%, while in the women with lucent breasts it was 92.3% (also see table 1), p=0.001. In rounds 5–10 (fig 1B), however, there was hardly any difference in the sensitivity in the women with dense and lucent breast patterns in the first year of the screening interval (dense 84.5% *v* lucent 86.1%, p=0.75).

BREAST CANCER STAGING AND SURVIVAL
Table 2 shows disease staging characteristics of the tumours that were screen detected during a repeat screening examination.

The proportion of small invasive tumours was somewhat larger in the patients with lucent

breasts than in those with dense breasts (rounds 2–4: p=0.56, rounds 5–10: p=0.19).

Axillary lymph node involvement from round 5 onwards is also shown in table 2. Ductal carcinomas in situ have been included as node negative. The proportion of tumours with positive nodes was nearly twice as high in the patients with lucent breasts as in those with dense breasts (28% *v* 14%, p=0.03). Similar striking results were also observed with regard to stage distribution from round 5 onwards: tumours were more advanced in the patients with lucent breasts (p=0.11).

Figure 2A and 2B shows separate Kaplan-Meier survival curves for patients with dense and lucent breast patterns diagnosed in rounds 2–4 and rounds 5–10, respectively. Only tumours that were screen detected during a repeat screening examination are included. At the end of 1995, 53 of the 258 patients had died, 26 of breast cancer and 26 of other causes. In one patient the cause of death was unknown. In rounds 2–4, the survival curve in the patients with dense breasts was lower than that in the patients with lucent breasts (p=0.07), which could also be demonstrated by 10 year survival probabilities: 73% in the patients with dense breasts *v* 83% in the patients with lucent breasts. In rounds 5–10, however, survival improved in all the patients and the difference in survival curves between the patients with dense and lucent patterns decreased (p=0.48). Ten year survival probabilities were 89% and 96%, respectively.

Discussion

Our results for the screening rounds in which mid-1970s mammography was used (rounds 2–4) show that high breast density hampered the detection of tumours. This is in accordance with the results of other studies on this subject.[6][15] Since the introduction of high quality mammography (1982), however, the situation has greatly improved with respect to positive predictive value of the screening test, survival and the detection of tumours with a lead time of up to one year. There has been little improvement in the detection of tumours with a lead time greater than one year.

A few limitations of this study have to be considered in interpreting the results. Firstly, the number of women with dense breast patterns (and consequently the number of patients with dense breast patterns) was quite small, which led to decreased precision of the results. Despite this, our main finding stands, as we found pronouced differences in screening performance in the rounds 2–4, when numbers were particularly small, while in the rounds 5–10, when the power to detect differences was larger, differences in screening performance diminished instead.

Secondly, our classification of breast patterns was performed by optical review, which is subjective. Although misclassification will be minimal as the reproducibility of binary partition of breast density is quite high with inter- and intraobserver agreement percentages of 80–90%,[20] it cannot be excluded. Random misclassification may have occurred and, theo-

retically, systematic misclassification as well, if the radiologist were prone to overestimation of density when he observed a tumour on the screening mammogram. Both forms of misclassification could have led to underestimation of the differences in screening performance between women with dense breasts and women with lucent breasts.

It was remarkable that the axillary lymph node status and disease stage seemed to be more favourable in screen detected patients with dense breasts than in those with lucent breasts. We examined the possibility that most of the tumours in dense patterns that had a poor prognosis were diagnosed as interval cancers (of which there was an excess in the women with dense patterns, table 1). As the proportions of node positive tumours among the interval tumours were equal for patients with dense and patients with lucent breasts (38% and 37%, respectively), this could not explain the surprising association between breast density and axillary lymph node status. We have no other plausible explanation than that, given the small numbers, the relation observed is a result of chance.

In screening rounds 5–10 there was hardly any difference in survival between the patients with dense breasts and those with lucent breasts. If anything, survival was somewhat poorer in the patients with dense breasts. This finding is rather surprising, because axillary lymph node status and disease stage were more favourable in patients with dense patterns than in those with lucent breasts. This may be the result of the influence of other prognostic indicators for which there was no information available. An example of such an indicator would be histological grade, which relates to the intrinsic or potential behaviour of the tumour, whereas tumour size and axillary lymph node status are merely indicators of how long the cancer has been present.[21]

With respect to present day screening practice, the most important effect of high breast density seems to be a persistently lower sensitivity, if the definition of false-negative is based on an interval period of two years. However, when only interval cancers diagnosed

---

KEY POINTS

- In early screening years (1977–1982) high mammographic breast density had an unfavourable effect on screening performance.
- With the introduction of modern mammography the situation has improved to a large extent.
- Women with dense breasts are no longer at a disadvantage concerning the predictive value of a positive screening test, tumour stage, and survival.
- There has been improvement in the detection of tumours in dense breasts with a lead time of up to one year.
- Little improvement has occurred in the detection of tumours with a lead time greater than one year.

within one year of a negative screening examination were considered to be false-negative, the difference in sensitivity between women with dense breasts and those with lucent breasts largely disappeared in rounds 5–10.

Two other studies that used modern mammography techniques examined the effect of breast density on the proportions of screen detected and interval tumours.[9 22] In accordance with our findings, Kerlikowske *et al*[9] observed (in women aged 50 and over) that the sensitivity was lower for dense breasts than for lucent breasts. In contrast with our study, however, they already found this difference when the definition of false-negative was based on an interval period of only 13 months. Ciatto *et al*[22] on the other hand, who studied interval cancers diagnosed within two years after a negative screening examination, did not find a relation at all between breast pattern and the occurrence of interval cancers.

Because of differences in the use of breast pattern classifications and study design it is difficult to compare the results of these studies with our own. Kerlikowske *et al*[9] used qualitative estimates of breast density ("fatty" and "dense") and they restricted their findings to first screening examinations only. The latter implicates that their radiologists had no previous mammograms at their disposal. We, on the other hand, studied repeat screening examinations only and therefore previous mammograms were available for all women in our study. This could have facilitated the detection of tumours, even in dense breasts, and may explain why in our study differences in sensitivity between dense and lucent breasts were smaller than in Kerlikowske's study.

Ciatto *et al*[22] classified breast patterns according to Wolfe's criteria at the initial screening examinations of women aged 40–70 years. They studied the proportions of screen detected and interval tumours that occurred during the five years after the breast pattern classification. However, breast density decreases with increasing age and particularly in perimenopausal women considerable changes may occur, even in a five year period.[23] Therefore, the initial breast patterns may not be representative of the breast patterns later on. This could be the reason that these authors did not find a relation between breast pattern and the occurrence of interval cancers.

In conclusion, from our study it seemed that with the use of modern mammography techniques the only important difference between dense and lucent breasts lies in the detection of tumours with a lead time greater than one year. To solve this problem, we advocate research into the influence of shortening screening intervals, taking additional mammographic views or using advanced imaging techniques (for example, digital mammography) in women with dense breasts.

1 Andersson I, Aspegren K, Janzon L, *et al*. Mammographic screening and mortality from breast cancer: the Malmö mammographic screening trial. *BMJ* 1988;**297**:943–8.
2 Roberts MM, Alexander FE, Anderson TJ, *et al*. Edinburgh trial of screening for breast cancer: mortality at seven years. *Lancet* 1990;**335**:241–6.
3 Frisell J, Eklund G, Hellström L, Lidbrink E, Rutqvist L-E, Somell A. Randomized study of mammography screening - preliminary report on mortality in the Stockholm trial. *Breast Cancer Res Treat* 1991;**18**:49–56.
4 Tabar L, Fagerberg G, Duffy SW, Day NE, Gad A, Gröntoft O. Update of the swedish two-county program of mammographic screening for breast cancer. *Radiol Clin North Am* 1992;**30**:187–210.
5 Kerlikowske K, Grady D, Rubin SM, Sandrock C, Ernster VL. Efficacy of screening mammograpy. *JAMA* 1995;**273**: 149–54.
6 Feig SA, Shaber GS, Patchefsky A, *et al*. Analysis of clinically occult and mammographically occult breast tumors. *AJR* 1977;**128**:403–8.
7 Holland R, Hendriks JHCL, Mravunac M. Mammographically occult breast cancer. A pathologic and radiologic study. *Cancer* 1983;**52**:1810–19.
8 Bird RE, Wallace TW, Yankaskas BC. Analysis of cancers missed at screening mammography. *Radiology* 1992;**184**: 613–17.
9 Kerlikowske K, Grady D, Barclay J, Sickles EA, Ernster V. Effect of age, breast density and family history on the sensitivity of first screening mammography. *JAMA* 1996;**276**: 33–8.
10 Fajardo LL, Hillman BJ, Frey C. Correlation between breast parenchymal patterns and mammographers' certainty of diagnosis. *Invest Radiol* 1988;**23**:505–8.
11 Peeters PHM, Verbeek ALM, Hendriks JHCL, Holland R, Mravunac M. The predictive value of positive test results in screening for breast cancer by mammography in the Nijmegen programme. *Br J Cancer* 1987;**56**:667–71.
12 Hislop TG, Warren Burhenne LJ, Basco VE, NG VTY. The screening mammography program of British Columbia: Pilot study. *Can J Public Health* 1991;**82**:168–73.
13 Weaver M. Breast cancer in nonpalpable lesions: can mammographic parenchymal pattern improve the predictive value of biopsy? *Am Surg* 1992;**58**:692–4.
14 Jackson VP, Hendrick RE, Feig SA, Kopans DB. Imaging of the radiographically dense breast. *Radiology* 1993;**188**:297–301.
15 Whitehead J, Cooper J. Risk factors for breast cancer by mode of diagnosis: some results from a breast cancer screening study. *J Epidemiol Community Health* 1989;**43**: 115–20.
16 Otten JDM, Van Dijck JAAM, Peer PGM, *et al*. Long term breast cancer screening in Nijmegen, The Netherlands: the nine rounds from 1975–92. *J Epidemiol Community Health* 1996;**50**:353–8.
17 Lundgren B. The oblique view at mammography. *Br J Radiol* 1977;**50**:626–8.
18 Baines CJ, Miller AB, Kopans DB, *et al*. Canadian National Breast Screening Study: Assessment of technical quality by external review. *AJR* 1990;**155**:743–7.
19 Peer PGM, Verbeek ALM, Mravunac M, Hendriks JHCL, Holland R. Prognosis of younger and older patients with early breast cancer. *Br J Cancer* 1996;**73**:382–5.
20 Saftlas AF, Szklo M. Mammographic parenchymal patterns and breast cancer risk. *Epidemiol Rev* 1987;**9**:146–74.
21 Miller WR, Ellis IO, Sainsbury JRC, Dixon JM. ABC of breast diseases: prognostic factors. *BMJ* 1994;**309**:1573–6.
22 Ciatto S, Zappa M. A prospective study of the value of mammographic patterns as indicators of breast cancer risk in a screening experience. *Eur J Radiol* 1993;**17**:122–5.
23 Van Gils CH, Otten JDM, Verbeek ALM, Hendriks JHCL. Short communication: Breast parenchymal patterns and their changes with age. *Br J Radiol* 1995;**68**:1133–5.