ELSEVIER

European Journal of Cancer 41 (2005) 2728–2734

European Journal of Cancer

www.ejconline.com

Early diagnosis, not differential treatment, explains better survival in service screening

Eugenio Paci [a,*], Antonio Ponti [b], Marco Zappa [a], Silvia Patriarca [c], Patrizia Falini [a], Gabriella Delmastro [b], Simonetta Bianchi [d], Anna Sapino [e], Vania Vezzosi [d], Carlo Senore [b], Emanuele Crocetti [a], Alfonso Frigerio [f], Roberto Zanetti [c], Marco Rosselli del Turco [g], Nereo Segnan [b]

[a] *Unit of Clinical and Descriptive Epidemiology, Centre for Study and Prevention of Cancer, Research Institute of Tuscany Region, Via di San Salvi 12, 50125 Firenze, Florence, Italy*

[b] *Unit of Epidemiology, Centre for Prevention of Cancer, Turin, Italy*

[c] *Piedmont Cancer Registry, Centre for Prevention of Cancer, Turin, Italy*

[d] *Department of Pathology, University of Florence, Florence, Italy*

[e] *Department of Pathology, University of Turin, Italy*

[f] *First Breast Screening Unit, S. Giovanni Battista Hospital, Turin, Italy*

[g] *Screening Unit and Scientific Director, Centre for Study and Prevention of Cancer, Research Institute of Tuscany Region, Florence, Italy*

Received 9 May 2005; received in revised form 16 June 2005; accepted 17 June 2005
Available online 18 October 2005

Abstract

Italian population-based breast cancer screening programmes with 2-year, high-quality mammography started in the cities of Florence and Turin in the early 1990s. Breast cancer cases from the local Tumour Registry were classified by method of detection and tumour characteristics (size, nodal-status and grade). Follow-up was at December 2001.

In total, 4444 breast cancer cases were analysed. The Hazard Ratio comparing before and after-invitation breast cancer cases indicated a 27% reduction (HR = 0.73; 95%CI: 0.61–0.87) in the risk of dying for the disease. After adjustment for tumour characteristics, survival rate was comparable by invitation status, whereas the proportion of early cancer was 33.7% and 46.6% in the before and after-invitation group. Survival rates by tumour characteristic subgroups was comparable by invitation status. Late stage and grade 3 were indicators of poor prognosis. Adjustment for tumour characteristics confirmed screening and not differential treatment as the most important reason for the observed survival benefit. The survival analysis by specific subgroups did not support the hypothesis that the difference in prognosis was attributable to differential treatment.

Keywords: Mammography screening; Breast cancer care; Breast cancer treatment; Survival rate; Breast cancer mortality

1. Introduction

The efficacy of breast cancer screening, with or without mammography, was evaluated by means of randomised clinical trials that were carried out in the 1970s and 1980s in the US and Europe. Results from these studies have been reassessed in several metanalyses [1]. The metanalysis by Gotzsche and Olsen [2] raised doubts about the efficacy of screening for reducing breast cancer mortality and the balance between financial and human cost and screening effectiveness [3,4]. Long-term results from randomised studies have recently been re-evaluated and new follow-up data have been published [5]. The International Agency for Research on Cancer (IARC) working group for the evaluation of breast cancer screening

* Corresponding author. Tel.: +39 0556268329; fax: +39 055679954.
E-mail address: e.paci@cspo.it (E. Paci).

0959-8049/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejca.2005.06.026

efficacy stated that there is sufficient evidence from mammography for women aged 50–69 [6].

Breast cancer population-based screening programmes were started at the end of the eighties in several European countries on a national or regional basis [7]. The 50-year and over female target population was invited to receive high-quality mammography screening every 2 or 3 years. There is wide agreement that the challenge for the future is the evaluation of the impact of service screening on breast cancer mortality [8].

A District of Florence breast cancer screening programme was already in operation in 1975, and contributed to the evaluation of screening efficacy by means of a case control study [9,10]. The programme was then also implemented in the City of Florence in 1990. The results from this programme are presented in this paper. Service screening started in Turin in 1992. Limited resource allocation meant that women aged 50–59 were invited up until 1998. After that year, women aged 60–69 were also enrolled.

In the 1990s, the European Commission's Europe Against Cancer programme initiated and funded pilot projects in European countries with the aim of developing guidelines for quality assurance and methodology for the epidemiologic evaluation of screening programmes. Pilot projects were set up in Italy in the cities of Florence and Turin. Both cities have been covered by cancer registry since the eighties [11] and were able to collect all population-based data on breast cancer characteristics for the entire target population and document survival rates [12]. Performance indicators for both programmes, including diagnosis and treatment [13], are collected annually within the breast cancer screening programme national survey carried out by the Italian Group for Breast Cancer Screening (www.senologia.it/gisma).

The aim of this study was to evaluate the impact of service screening on breast cancer occurrence by tumour characteristics and survival from the disease. While the improvement in survival rates should not be considered, *per se*, as a breast cancer mortality reduction surrogate, changes in these parameters are essential components along the path of the intended effect [14]. The major problem in using survival as indicator of screening efficacy is length bias, *i.e.*, the tendency of screening to detect more indolent tumours. Higher detection rates are not necessarily worth, because screening could miss more aggressive, fast growing breast cancers. Furthermore the effect of screening might be confounded by the opportunity each woman has to receive effective therapy. If screened women are compared with historical or currently unscreened breast cancer cases, separating the screening contribution from the contribution related to differential treatment is challenging.

In this paper, we have compared breast cancer cases that were diagnosed before *versus* after the invitation to participate in service screening. We tested the statistical difference in the risk of dying within specific subgroups by tumour characteristics. This approach has the characteristic of an intention to treat analysis (non-randomised), thus overcoming the problem of selection bias due to non-attendance.

2. Patients and methods

Breast cancer cases diagnosed in women resident in the two cities were entered into the Tumour Registry according to the IARC rules for cancer registration [11]. *In situ* carcinomas are included in this project, while death certificate only (DCO) and multiple primaries were excluded.

The characteristics of both the breast cancer screening programmes and main performance indicators have been described before in detail elsewhere [15]. National recommendations for service screening in Italy were adopted in both programmes. A list of resident women was issued by the Municipality and a letter of invitation stating appointment details for a mammographic test was sent to eligible women. The participation rate was about the 60%. Two-view, double-read mammographies were performed under strict quality control. The Florence study population (about 60 000) was resident women aged 50–69 invited over the 1990–1998 period to have a screening mammography every two years. The first round of screening started in 1990 and was completed in 1993. The Turin programme started issuing invitations in 1992 and only included women aged 50–59 (about 75 000) for the first invitation during the initial period. Invitations were randomised by clusters of general practitioners in Turin during the building up phase (1992–1998).

All breast cancer tumours were classified by size and nodal-status according to the TNM-UICC classification. The number of examined and positive nodes was registered. Grading was defined at three levels by pathologists in Florence and Turin. Follow-up for live status and cause of death was as at 31 December 2001.

All registry-based breast cancer cases were linked to the screening file and divided by detection method. Breast cancer cases were divided into four main categories as follows:

(1) Cases having a tumour detected at the first screening test (first round or subsequent) are prevalent screen-detected cases; cases screen detected at a repeated screening test are incident breast cancer cases.
(2) Cases diagnosed clinically outside the screening process following a negative screening test (includes interval cancer cases).
(3) Cases within non-respondent group.

(4) Cases diagnosed by service screening before the invitation (as it took several years to achieve full coverage of the population with an invitation to screening).

Follow-up duration was different according to detection method. The design of the study was such that prior to the commencement of screening, the Not-yet invited breast cancer cases had a longer follow-up than the cases diagnosed through invitations, *i.e.*, following the commencement of screening. Median follow-up time (excluding dead women) was 6.2 years for the Not-yet invited group and 4.3 years for non-respondents. It was 3.8 years for screened women and 3.7 and 6.2 years, respectively, for screen detected at incident and prevalent breast cancer cases. The median number of lymph nodes examined at axillary dissection was 17.

Survival rates were estimated using the Kaplan–Meier method and Cox proportional hazard models were fitted using Stata [16] to estimate Hazard Ratios (HR) and a 95% confidence interval (95%CI). Size, measured as pT, nodal-status and grading were used in the model as main effects. Screening centres and age (continuous) were used as adjusting factors. Models 1 and 3 were fitted and compared the survival of invited with the Not-yet invited women (intention to treat analysis) with or without adjustment for main factors. Models 2 and 4 were fitted and compared the detection method (screen detected, screened negative, non-respondents), with the results from Not-yet invited breast cancer cases used as reference category.

3. Results

Four thousand and four hundred and forty four breast cancer cases were registered in total in the 50–69 years-of-age target population: 1689 in Florence and 2755 in Turin. 24.5% of breast cancer cases were diagnosed before the programme invitation in Florence and 69.9% in Turin, whilst 19.2% of cases were screen detected in Turin and 42.7% in Florence.

Prior to the commencement of screening, there were 19 and 98 *in situ* carcinomas: 5.1% and 4.6% of all Not-yet invited breast cancer cases in Florence and Turin, respectively. Breast cancer cases (53.0%) before-screening invitation were Stage II+ in Turin and 49.3% in Florence. Five-year survival rates of Not-yet invited breast cancer cases were 83% and 84% in Turin and Florence, respectively. Given the high comparability of the state distribution and survival rates, the following analysis is presented for the joint Florence and Turin dataset.

Table 1 shows invasive breast cancer cases by detection mode, pathological T, nodal-status and grade. 36.6% and 23.9% of breast cancer cases classified as T2+ were in the before *versus* after-invitation periods, respectively. Positive nodes occurred in 38.6% of before-invitation breast cancer cases and 31.4% in the after-invitation period. Screen-detected and breast cancer cases diagnosed in screened women showed a higher number of early tumours. Grade 3 tumours in before-invitation breast cancer cases represented 15.9% of invasive Stage I *versus* 35.2% of Stage II+ breast cancer cases. The corresponding values in Invited breast cancer cases were 15.3% and 35.6%, respectively. Screen-detected breast cancer cases were analysed separately by a prevalent or incident screening test. The proportion of *in situ* at the first and repeated screening test was 13.6% and 11.4%.

Fig. 1 shows survivor functions by invitation. Detection method was a strong survival determinant with a 95% and 96% 5-year survival for screen-detected (first and repeated) breast cancer cases; the 5-year survival was 80%, 75% and 83% for clinical detected in screened, never responder and Not-yet invited breast cancer cases, respectively. The Hazard Ratio for invasive breast cancer cases was adjusted by age and centre (Table 2). The before-invitation breast cancer cases were used as a reference category and both the Invited (intention to treat analysis) (Model 1) and the detection method (Model 2) were used as index category. The Hazard Ratio in Model 1 was 0.73 (95%CI: 0.61–0.87) with a reduction of 27% for invited women. A comparison of survival rates for breast cancer cases in screened *versus* unscreened women indicates that there was a 56% reduction in the risk of dying for screened women (HR = 0.44; 95%CI: 0.35–0.54).

Pathological T, nodal-status and grade were included as main factors in multivariate Models 3 and 4. As expected, the HR showed increased risk in pathological T, node-positive status and grade 3. The HR for invitation state increased to unity (HR = 1.03; 95%CI: 0.95–1.29) after adjustment for tumour characteristics (Model 3). The screen-detected cases HR increased to 0.63 and 0.69 in Model 4 for the prevalent and incident screening test, respectively, after adjustment for tumour characteristics.

In order to separate the effect of better treatment and access to treatment from the screening effect due to the stage shift of breast cancer screen-detected cases, we calculated the specific survival rates at 5 years by size, nodal-status and grading for specific subgroups before and after invitation (Table 3). Breast cancer cases with missing characteristics were excluded, the specific survival rate was 31.1% and 22.0% of Invited and Not-yet invited, respectively. *P*-values for the statistical difference in survival rates were never statistically significant with regard to before–after-invitation status. Within each category, survival rates were closely comparable by invitation status, whereas the proportion of T1, node-negative tumours was 41.8% and 27.3% in the after and

Table 1
Breast cancer cases by pathological T, nodal-status and grade and by method of detection

Pathological size	Screen detected (prevalent)	Screen detected (incident)	Clinical detected in screened	Never responders	All invited	Non-yet invited	Total
Tis	101	58	17	18	*194*	*117*	*311*
%	13.63	11.39	4.70	3.65	*9.22*	*5.00*	*7.00*
T1a	80	63	17	16	*176*	*96*	*272*
%	10.80	12.38	4.70	3.25	*8.36*	*4.10*	*6.12*
T1b	183	126	45	32	*386*	*274*	*660*
%	24.70	24.75	12.43	6.49	*18.34*	*11.71*	*14.85*
T1c	254	194	145	158	*751*	*823*	*1574*
%	34.28	38.11	40.06	32.05	*35.68*	*35.19*	*35.42*
T2+	111	62	114	216	*503*	*856*	*1359*
%	14.98	12.18	31.49	43.81	*23.90*	*36.60*	*30.58*
Missing	12	6	24	53	*95*	*173*	*268*
%	1.62	1.18	6.63	10.75	*4.51*	*7.40*	*6.03*
Total	741	509	362	493	*2105*	*2339*	*4444*
%	100.00	100.00	100.00	100.00	*100.00*	*100.00*	*100.00*
Pearson chi square (5) (invited *vs.* non-invited) = 167.4 Pr = 0.0000							
Pathological nodes (invasive only)							
Negative	420	276	183	206	*1085*	*1056*	*2141*
%	65.63	61.20	53.04	43.37	*56.78*	*47.52*	*51.80*
Positive	159	117	131	193	*600*	*857*	*1457*
%	24.84	25.94	37.97	40.63	*31.40*	*38.57*	*35.25*
Missing	61	58	31	76	*226*	*309*	*535*
%	9.53	12.86	8.99	16.00	*11.83*	*13.91*	*12.94*
Pearson chi square(2) (invited *vs.* non-invited) = 15.5 Pr = 0.0000							
Pathological grade (invasive only)							
1	270	144	57	69	*540*	*352*	*892*
%	42.19	31.93	16.52	14.53	*28.26*	*15.84*	*21.58*
2	207	173	121	151	*652*	*752*	*1404*
%	32.34	38.36	35.07	31.79	*34.12*	*33.84*	*33.97*
3	104	88	102	150	*444*	*571*	*1015*
%	16.25	19.51	29.57	31.58	*23.23*	*25.70*	*24.56*
Missing	59	46	65	105	*275*	*547*	*822*
%	9.22	10.20	18.84	22.11	*14.39*	*24.62*	*19.89*
Pearson chi square(3) (invited *vs.* non-invited) = 32.7 Pr = 0.0000							

before-invitation groups, respectively. Tumours with pathological T and other characteristics missed or unknown, showed a non-significant difference in 5-year survival in favour of the before-invitation group (60.7% *vs.* 47%, $P = 0.661$).

Late stage and grade 3 breast cancer cases were the main indicators for poor prognosis. In Not-yet invited women, we observed 432 breast cancer deaths within 8 years of diagnosis. Of these, 77.8% occurred in Stage II+ at diagnosis breast cancer cases and 33.3% with grade 3 tumours. There were 230 deaths from breast cancer in invited women, 69.6% in Stage II+ at diagnosis and 36.5% with grade 3 tumours. 118 screened women died of breast cancer within 8 years of diagnosis, 72.9% were Stage II+ at diagnosis and 34.7% were classified as grade 3.

4. Discussion

Breast cancer screening programmes started in the cities of Florence and Turin at the beginning of the 1990s: the first large-scale population-based screening programmes in Italy. These programmes belong to the European Breast Screening Network, a European funded research group against cancer. The proportion of *in situ* carcinomas was comparable between cities in women diagnosed before the commencement of screening. Invasive breast cancer cases were also comparable in terms of tumour size, nodal-status and grading. Prior to the commencement of screening, breast cancer cases showed similar survival rates by centre, and confirmed the expected relationship between tumour characteristics (size, nodal-status and grade) and survival rates. The proportion of breast cancer cases diagnosed post-commencement of screening is different in the two programmes due to enrolment phase timing. An intention to treat analysis on the invited group showed statistically significant improved survival rates. Selection bias did not affect this result, as invitations were cluster randomised in Turin and proceeded by neighbourhood with no obvious relation by determinants of breast cancer survival in Florence. However, this comparison is affected by non-compliance and contamination, as the

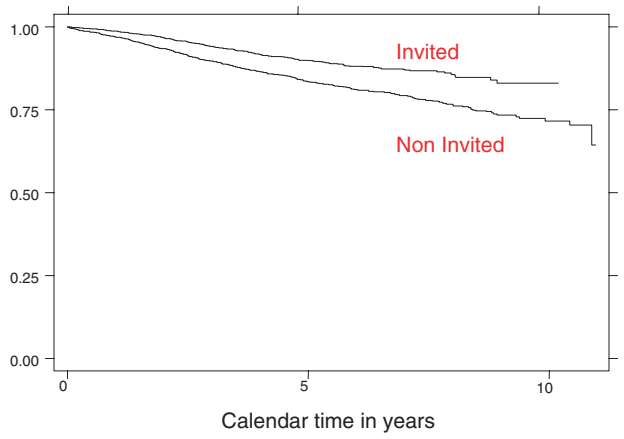

Fig. 1. Survivor functions of invasive breast cancer cases, adjusted for screening centre and age (continuous), by invitation (intention to treat analysis).

participation rate was about 60%. In addition, a portion of Not-yet invited women spontaneously attended mammography outside the screening programme and in the absence of symptoms. This proportion has been estimated in about 15% in the three years before interview, according to a survey conducted in the late eighties in Turin [17] and in Florence before the start of the screening programme [18]. This had an impact on the proportion of Stage II+ breast cancer cases in Turin and Florence prior to the commencement of screening that was comparable to the proportion observed in other cancer registries in Italy [19].

Screen-detected breast cancer cases at the prevalence and repeated screening test showed better survival rates at 5 years. A poor stage distribution and prognosis for non-respondent breast cancer cases has been reported in other screening trials. Breast cancer cases diagnosed in screened women outside the screening process, the majority of which occurred in the two year interval, showed survival rates comparable with the Not-yet invited group.

In multivariate analysis, the adjustment for tumour characteristics explained the survival difference, which is attributable to earlier diagnosis, *i.e.*, earlier stage, in the intention to treat analysis. In the model, the adjustment did not explain the difference in the risk of dying by detection method: survival gain for screen-detected tumours was persistent after adjustment for size, nodal-status and grade. This result is similar to what has been reported by Duffy and co-workers [20] where adjustment for tumour characteristics did not entirely absorb the gain observed for screen-detected cases. Similar estimates were reported by Moss *et al.* [21] in the analysis of survival rates in breast cancer cases diagnosed in the UK Trial. The marked residual survival gain for screen-detected *versus* Not-yet invited cases, after adjustment for tumour characteristics, deserves further investigation.

In the recent debate on breast cancer screening efficacy, some authors have suggested the mortality reduction observed in several countries over recent years is mainly attributable to the change in therapy that

Table 2
Cox model of the risk of dying for invasive breast cancer by method of detection with or without adjustment for tumour size, nodal status and grade

	Model 1	Model 2	Model 3	Model 4
	Haz. ratio	Haz. ratio	Haz. ratio	Haz. ratio
Method of detection				
All invited	0.73 (0.61–0.87)		1.03 (0.85–1.24)	
Screen detected (prevalent)		0.30 (0.21–0.41)		0.63 (0.45–0.89)
Screen detected (repeated)		0.34 (0.22–0.53)		0.69 (0.44–1.09)
Clinical detected in Screened		1.32 (0.99–1.76)		1.40 (1.04–1.86)
Never responders		1.60 (0.27–2.01)		1.23 (0.98–1.55)
City	1.02 (0.86–1.22)	0.98 (0.82–1.18)	1.10 (0.92–1.33)	1.07 (0.89–1.29)
Age	1.02 (1.00–1.03)	1.01 (1.00–1.03)	1.01 (0.99–1.02)	1.00 (0.99–1.02)
Nodal-status				
Negative			0.29 (0.24–0.36)	0.29 (0.24–0.37)
Missing			1.39 (1.09–1.79)	1.40 (1.09–1.80)
Grading				
1			0.37 (0.27–0.50)	0.40 (0.30–0.55)
2			0.53 (0.43–0.65)	0.53 (0.44–0.66)
Missing			0.68 (0.54–0.85)	0.67 (0.54–0.84)
Tumour size				
pT 1a			0.15 (0.08–0.29)	0.18 (0.09–0.34)
pT 1b			0.16 (0.10–0.26)	0.18 (0.12–0.29)
pT 1c			0.35 (0.29–0.43)	0.36 (0.30–0.45)
Missing			1.15 (0.86–1.55)	1.15 (0.86–1.55)

Reference category: non-yet invited, T2+, node positive, grade 3.

Table 3
Invasive breast cancer specific survival rates by size, nodal status, grade and invitation

pT, node, grade	Not-yet invited $N = 1531$		Invited $N = 1491$		P value
	Proportion (%)	5 Year survival (%)	Proportion (%)	5 Year survival (%)	
T1, node negative, grade 1–2	22.0	98.5	34.7	97.7	0.188
T1, node negative, grade 3	5.3	93.3	7.1	94.3	0.884
T1, node positive, grade 1–2	8.4	95.6	10.4	93.2	0.833
T1, node positive, grade 3	4.7	75.0	3.7	79.0	0.644
T2+, node negative, grade 1–2	6.2	86.7	4.9	92.2	0.630
T2+, node negative, grade 3	4.2	82.6	3.5	82.9	0.860
T2+, node positive, grade 1–2	8.9	74.8	6.2	74.7	0.498
T2+, node positive, grade 3	9.1	57.8	7.4	59.0	0.661

Log-rank test P –value between invited and non-yet-invited.
Only cases for which all tumour characteristics were not missed are included. Breast cancer cases without missing characteristics are the: 68.9% in the non-invited and 78.0% of invited cases.

occurred at the end of the eighties [22]. In a previous paper based on incidence-based mortality analysis [23] of the Florence data, the improvement related to new therapies was evident by comparing breast cancer cases diagnosed in 1985–1986 with cases diagnosed in Not-yet invited women in the 1990s. However, there was evidence of an additional screening effect by comparing the before and after-invitation breast cancer cases after the commencement of service screening, which started in Florence in 1990. The results from this larger study confirm that survival rates were highly comparable in the 1990s by before–after invitation period within specific subgroups of size, nodal-status and grade; whereas the proportion of early breast cancer cases was different by before–after period. If there had been differences in treatment regimen by before–after-invitation status, survival rates by state-specific subgroup would have been different.

The aim of breast cancer screening is the reduction in the absolute rates of advanced carcinoma, which has been shown to be a surrogate of mortality reduction [14]. This study shows that earlier diagnosis has been accomplished in the two programmes. Whether the lead-time acquired is sufficient to reduce mortality appreciably in the target population is being investigated at present. These data also show late stage and high-tumour grade were major indicators of the probability of dying both before and after screening: breast cancer deaths in screened women mainly occurred in breast cancer cases classified as advanced at diagnosis. The proportion of deaths occurring within advanced cases should be monitored in the future. This proportion should not decrease substantially in the absence of changes in treatment, in parallel with the reduction in advanced cancer rates, if screening is effective in reducing mortality.

Monitoring changes in prognosis and stage distribution in these two Italian cities shows that the programmes are on the right track. In order to demonstrate the benefit due to screening in terms of mortality, two conditions should be fully met in the near future. First, the absence of over-diagnosis of breast cancer cases: we performed an evaluation of the occurrence of over-diagnosis bias [24] in Florence and we have shown minimal, if any, excess of breast cancer cases in the after-invitation population, when lead-time is taken into account. Second, the survival gain for screen-detected tumours should not be lost as time passes. The follow-up in this study is too short to be conclusive, and is possibly still biased by lead-time.

In conclusion, the results of service screening in Florence and Turin, the leading programmes in Italy, showed an improvement in survival rates by before–after-invitation period in an intention to treat analysis. The proportion of earlier stages explained this gain and supports the conclusion that screening is changing the pattern of tumour characteristics in the target population. Survival analysis showed comparable behaviour within the same tumour characteristic subgroup by before–after commencement of screening, thus this result did not support the hypothesis that the difference in prognosis was due to differential treatment or access to treatment by the before *versus* after-invitation group.

References

1. Wald NJ, Chamberlain J, Hackshaw A. European society of mastology consensus conference on breast cancer screening: report of the evaluation committee. *Br J Radiol* 1994, **67**, 925–933.
2. Gotzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet* 2000, **8**, 129–134.
3. Gotzsche PC. On the benefits and harms of screening for breast cancer. *Int J Epidemiol* 2004, **33**, 56–64.
4. Freedman DA, Petitti DB, Robins JM. On the efficacy of screening for breast cancer. *Int J Epidemiol* 2004, **33**, 43–55.
5. Nystrom L, Andersson I, Bjurstam N, *et al*. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet* 2002, **359**, 909–919.
6. *IARC handbooks of cancer prevention. Breast cancer screening* vol. 7. Lyon, France, IARC, 2002.

7. Ballard-Barbash R, Klabunde C, Paci E, *et al*. Breast cancer screening in 21 countries: delivery of services, notification of results and outcomes ascertainment. *Eur J Cancer Prev* 1999, **8**, 417–426.
8. Cuzick J. Breast cancer screening–time to move forward. *Breast* 2002, **11**, 209–210.
9. Palli D, Del Turco MR, Buiatti E, *et al*. A case-control study of the efficacy of a non-randomized breast cancer screening program in Florence (Italy). *Int J Cancer* 1986, **38**, 501–504.
10. Paci E, Ciatto S, Buiatti E, *et al*. Early indicators of efficacy of breast cancer screening programmes. Results of the Florence District Programme. *Int J Cancer* 1990, **46**, 198–202.
11. Zanetti R, Gafà L, Pannelli F, Conti E, Rosso S. *Il Cancro in Italia*, vol. 3. Roma, Il Pensiero Scientifico Editore, 2003.
12. Rosso S, Casella C, Crocetti E, *et al*. Survival of cancer patients in Italy in the nineties: figures from the Cancer Registries. *Epidemiol Prev* 2001, **25**, 1–375.
13. Distante V, Mano MP, Ponti A, *et al*. Monitoring surgical treatment of screen-detected breast lesions in Italy. *Eur J Cancer* 2004, **40**, 1006–10012.
14. Tabar L, Duffy SW, Vitak B, *et al*. The natural history of breast carcinoma: what have we learned from screening. *Cancer* 1999, **86**, 449–462.
15. Giordano L, Giorgi D, Fasolo G, *et al*. Breast cancer screening: characteristics and results of the Italian programmes in the Italian group for planning and evaluating breast cancer screening programmes (GISMa). *Tumori* 1996, **82**, 31–37.
16. Stata software. Available from: www.stata.com.
17. Segnan N, Ronco G, Ponti A. Practice of early diagnosis of breast and uterine cervix cancer in a northern Italian town. *Tumori* 1990, **76**, 227–233.
18. Giorgi D, Paci E, Zappa M, *et al*. Breast cancer early diagnosis experience in Florence: can a self referral policy achieve the results of service screening? *J Epidemiol Community Health* 1994, **48**, 471–475.
19. Buiatti E, Barchielli A, Bartolacci S, *et al*. The impact of organised screening programmes on the stage-specific incidence of breast cancer in some Italian areas. *Eur J Cancer* 2003, **39**, 1776–1782.
20. Duffy SW, Tabar L, Fagerberg G, *et al*. Breast screening, prognostic factors and survival-results from the Swedish two county study. *Br J Cancer* 1991, **64**, 1133–1138.
21. Moss SM, Ellman R, Coleman D, *et al*. Survival of patients with breast cancer diagnosed in the United Kingdom trial of early detection of breast cancer. United Kingdom Trial of Early Detection of Breast Cancer Group. *J Med Screen* 1994, **1**, 193–198.
22. Jatoi I, Miller AB. Why is breast-cancer mortality declining? *Lancet Oncol* 2003, **4**, 251–254.
23. Paci E, Duffy SW, Giorgi D, *et al*. Quantification of the effect of mammographic screening on fatal breast cancers: the Florence programme 1990–96. *Br J Cancer* 2002, **87**, 65–69.
24. Paci E, Warwick J, Falini P, *et al*. Overdiagnosis in screening: is the increase in breast cancer incidence rates a cause for concern? *J Med Screen* 2004, **11**, 23–27.