available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

An estimate of overdiagnosis 15 years after the start of mammographic screening in Florence

*Puliti Donella, Zappa Marco, Miccinesi Guido, Falini Patrizia, Crocetti Emanuele, Paci Eugenio**

Clinical and Descriptive Epidemiology Unit, ISPO – Cancer Prevention and Research Institute, via San Salvi 12, 50135 Florence, Italy

ARTICLE INFO

Article history:
Received 20 March 2009
Received in revised form 18 May 2009
Accepted 12 June 2009
Available online 29 October 2009

Keywords:
Breast cancer
Overdiagnosis
Service screening

ABSTRACT

Aim: Several studies have tried to quantify overdiagnosis of breast cancer with mammography screening, but estimates vary widely. The aim of this study is to evaluate the degree of overdiagnosis of breast cancer 15 years after the introduction of service screening in Florence (Italy).

Methods: We selected 61,568 women aged 50–69 years at the beginning of service screening (1990) and we used the cancer registry data to follow up them for breast cancer incidence. The measure of overdiagnosis is the ratio of cumulative incidence of breast cancer in the invited group (observed) at least 5 years after the last screening to that expected in the absence of screening.

Results: Under the assumption of a 1.2% annual trend in pre-screening incidence in women aged 60–69 years at the start of service screening, the ratio of observed to expected cumulative cases was 1.01 (95%CI: 0.95–1.07), but assuming no incidence trend, an unlikely scenario, the estimate of overdiagnosis rose to 1.13 (95%CI: 1.07–1.19).

Conclusion: Overdiagnosis of breast cancer in Florentine service screening can be estimated only for women aged 60–69 years at the start of service screening, for it is only for this group that a sufficient follow-up period is available after the last screening. Although the estimate of overdiagnosis is very sensitive to pre-screening trend estimates, our data show that 15 years after the introduction of mammographic service screening the degree of overdiagnosis was nearly zero and more than likely lower than 13% in this age group.

1. Introduction

The risk of overdiagnosis, i.e. that a cancer would not have clinically presented during a patient's life, is a harm of screening that has been investigated since the results, in the 1970s, of the first randomised trials for lung cancer screening.[1] In his book, Screening in Chronic Disease,[2] Alan Morrison has already presented the difficult issue of the interaction of biological, individual and epidemiological explanations of overdiagnosis (he used the term 'pseudodisease'). More recently, there has been debate, and concern raised, about screening by PSA testing, which has substantially changed the pattern of prostate cancer incidence in western countries.[3]

Today, in many European countries, high numbers of women participate in breast cancer service screening within well-established public health programmes [4]. There is, therefore, a need to estimate the overdiagnosis harm attributable to breast cancer screening. However, studies of service screening are still few and the methodology is uncertain.

* *Corresponding author:* Tel.: +39 0556268329; fax: +39 055679954.
E-mail address: e.paci@ispo.toscana.it (E. Paci).
0959-8049/$ - see front matter
doi:10.1016/j.ejca.2009.06.014

Biesheuvel and co-workers[5] have recently published a systematic review of the effects of study methods and biases on the estimates of overdiagnosis, and have suggested a methodology to reduce the risk of error. They present and discuss methods previously used in the evaluation of overdiagnosis of breast cancer in mammography screening. They suggest that the *cumulative-incidence method* with data from a randomised controlled trial (RCT) is the more valid method when there is a sufficient follow-up (at least 5 years) after the last screening, and, when there is little or no follow-up after the last screening, the lead time bias should be adjusted for with statistical methods, as suggested in other papers.[6,7]

In order to estimate overdiagnosis using the *cumulative-incidence method*, we analysed breast cancer incidence data from the cancer registry in the city of Florence, where a service screening programme has been active for 15 years.

2. Method

The Florentine service screening programme offers high-quality mammography every 2 years to all resident women aged 50–69 years.[8] The overall target population has been about 60,000 with an attendance rate of around 60% at the prevalence round, although this has increased to more than 70% in recent years.[9] Performance indicators, including diagnosis and treatment, are collected annually under a national survey carried out by the Italian Group for Breast Cancer Screening.[10] Breast cancer cases diagnosed in the target population are registered by the Tuscan Tumour Registry, which has been operating in the area since 1985.

In this analysis we selected 61,568 women aged 50–69 years at the beginning of service screening (September 1990), distinguishing four birth cohorts:

(a) Born 1937–1941 (aged 50–54 years at the beginning of service screening).
(b) 1932–1936 (55–59 years).
(c) 1927–1931 (60–64 years).
(d) 1922–1926 (65–69 years).

We used the cancer registry data to follow up each of the four cohorts for breast cancer incidence between 1990 and 2004.

In the *cumulative-incidence method* a comparison is made of the cumulative incidence among a group of women who were invited to screening with the cumulative incidence in a non-invited group over the same time period. Assuming similar underlying risks of breast cancer, in the absence of overdiagnosis, the cumulative incidence in the two groups is the same given a sufficient follow-up after the last screening.

In randomised controlled trials, the breast cancer incidences of invited and non-invited populations were obtained, respectively, from the intervention and control arms. In service screening the incidence in the invited population can be obtained from the population incidence rates after the introduction of service screening, but the incidence in the non-invited population over the same time period is not available. Therefore, data are derived from incidence before implementation of the screening programme with adjustment for changes in breast cancer incidence over time and for the different age distribution.

We previously estimated the pooled annual trend of incidence in the pre-screening period in six areas of Northern and Central Italy.[11] The annual percentage change in the pooled incidence trend was 1.2% (95%CI: 0.8–1.6%) for all breast cancers and 0.9% (95%CI: 0.5–1.3%) for invasive breast cancers only. In this analysis, a Poisson regression model was used to fit Florentine incidence data for 1986–1990, the period before the implementation of the screening programme. The model incorporated age and calendar year as continuous variables, but the calendar year parameter was forced to the values previously estimated (1.2% for all cancers and 0.9% for invasive only). To take into account the random fluctuation in the trend estimate, we also performed a sensitivity analysis of the overdiagnosis estimate assuming no temporal trend in breast cancer incidence.

The measure of overdiagnosis is the ratio of cumulative incidence of breast cancer in the invited group (observed) to that in the non-invited group (expected) at least 5 years after the last screening.

3. Results

In Fig. 1a–d the observed and expected breast cancer incidence rates in the target population are presented for each cohort. Each curve shows a 'prevalence peak' during the first round (about 3 years were needed for the enrolment period) and other peaks every 2–3 years. The excess of incidence remains as long as the screening programme is in place. The youngest cohorts (50–54 and 55–59 years old at the beginning of service screening) were still being invited to screening at the end of the study period or there was not enough follow-up period after the last screening. However, a follow-up of 5 and 10 years, respectively, was available for the oldest cohorts: the 60–64-year-old cohort received 3 or 4 screenings and then they were no longer actively invited, whilst women in the oldest cohort attended 1 or 2 rounds and then they, too, were no longer invited. After service screening stopped, incidence rates declined sharply at older ages.

The cumulative breast cancer cases, observed and expected, are presented in Fig. 2a–d by cohort and age at diagnosis over the study period.

For women aged 50–54 years, observed cumulative incidence rates per 10.000 were 72, 197 and 359 and expected cumulative incidence rates were 67, 188 and 316 at 5, 10 and 15 years of screening activity, respectively. For women aged 55–59 years, corresponding cumulative rates were 92, 236, 401 (observed) and 73, 206, 348 (expected); for women aged 60–64 years, corresponding cumulative rates were 106, 261, 377 (observed) and 82, 227, 378 (expected); and for women aged 65–69 years, corresponding cumulative rates were 121, 269, 412 (observed) and 89, 247, 405 (expected).

In Table 1 the total study time (15 years) is divided into two periods – years of screening activity and years after the last screening, for each birth cohort. The cumulative incidence excess at the last available year of screening is reported for all birth cohorts, but an estimate of overdiagnosis can be calculated only for the two cohorts for which there is a follow-up

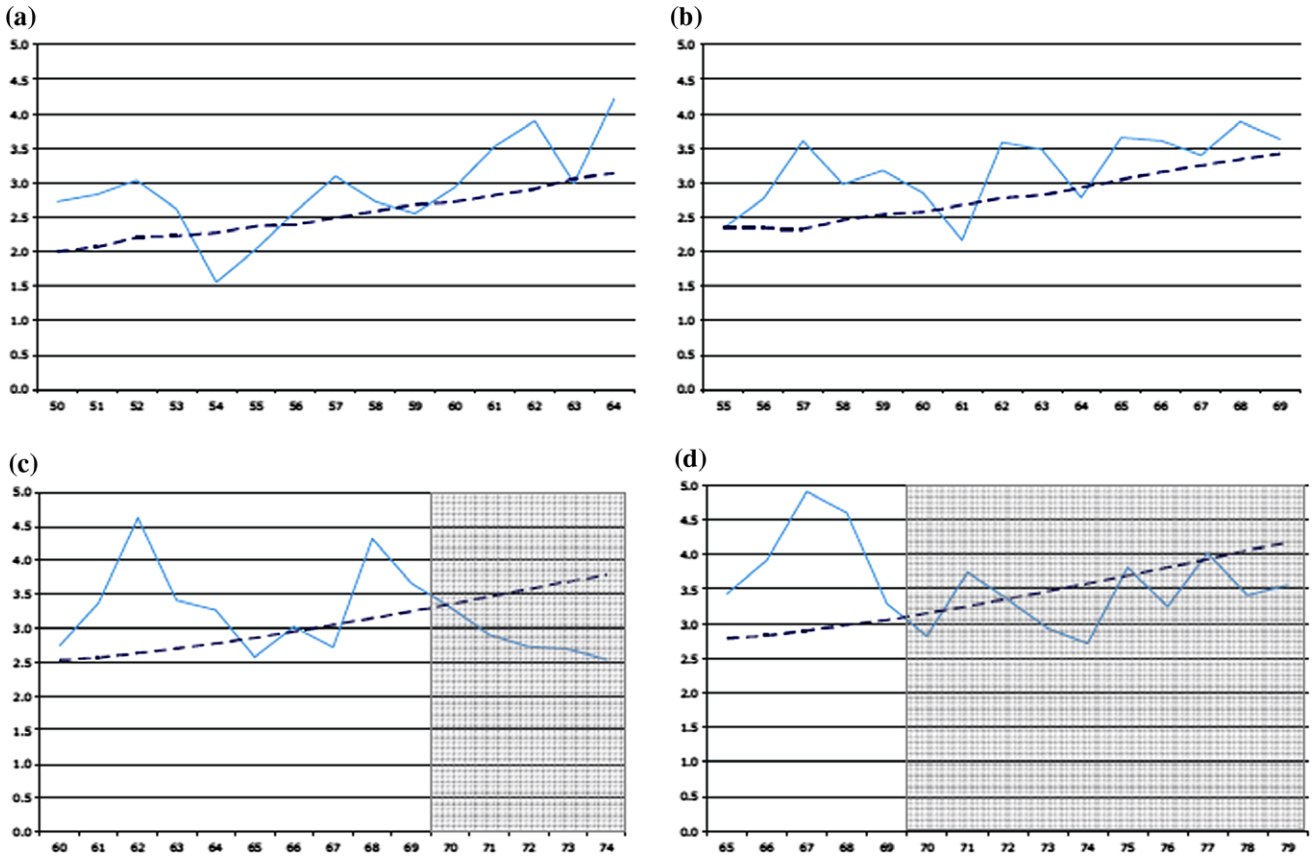

Fig. 1 – Invited (observed) and non-invited (expected) incidence breast cancer rates by age at the beginning of service screening: (a) 50–54 years, (b) 55–59 years, (c) 60–64 years and (d) 65–69 years. Solid line = observed and dotted line = expected.

period after the last screening. The excess of incidence at the end of the study period for the two youngest cohorts, when they were still enrolled in the service screening programme, was, respectively, 1.14 (95%CI: 1.04–1.23) and 1.15 (95%CI: 1.06–1.25). For the two older cohorts, the excess of incidence observed at 69 years was 1.15 (95%CI: 1.04–1.27) for the women who started screening at 60–64 years, and was 1.36 (95%CI: 1.17–1.57) for the oldest cohort, where the weight of prevalence screening is greater. When screening is discontinued, the incidence rate drops to a lower level than expected, and the overwhelming majority of the cases in excess was absorbed again. At the end of follow-up, the estimate of overdiagnosis was 1.00 (95%CI: 0.92–1.08) and 1.02 (95%CI: 0.94–1.10) for 60–64 and 65–69 years at the start of service screening, respectively.

In Fig. 3 the cumulative excess breast cancer rates are shown for different ages in the study period for each birth cohort. The excess of incidence was highest during the initial screening round. At the end of follow-up the cumulative excess rate was about 0.40 per 1000 person years for women enrolled in the programme when they were 50–59 years old and still being invited for screening. For women aged 60–69 years at the enrolment, the excess decreased over time and was under 0.05 per 1000 person years at the end of follow-up.

Overall, the estimate of breast cancer overdiagnosis for women who started screening between 60 and 69 years of age was 1.01 (95%CI: 0.95–1.07) for *in situ* and invasive cases and was 0.99 (95%CI: 0.94–1.05) for invasive cases only.

We performed a sensitivity analysis assuming no trend in the pre-screening period. In this case, the estimate of overdiagnosis for women aged 60–69 years at the enrolment was 1.13 (95%CI: 1.07–1.19) for all cases and was 1.08 (95%CI: 1.02–1.15) for invasive cases only.

4. Discussion

Several studies have tried to quantify overdiagnosis of breast cancer due to mammography screening, but estimates vary widely because there are many biases that might affect them.[5]

The large increase in incidence observed at the prevalence screening should not be confused with overdiagnosis for the increase, which is inherent to screening and needed for screening efficacy in reducing breast cancer mortality, is an effect of lead time. This increase extends into the subsequent rounds because of the shift in the age-incidence curve due to lead time.[12] Moreover, the diagnostic anticipation is at the same time an anticipation of the increasing incidence temporal trend.

Therefore, a valid estimate of overdiagnosis can be obtained only from the comparison of cumulative incidence in the invited (i.e. the intervention group in case of RCT) and non-invited groups (i.e. the control group in case of RCT) after the elapse, following the last screening, of a time equivalent to the lead time (*cumulative-incidence method*).

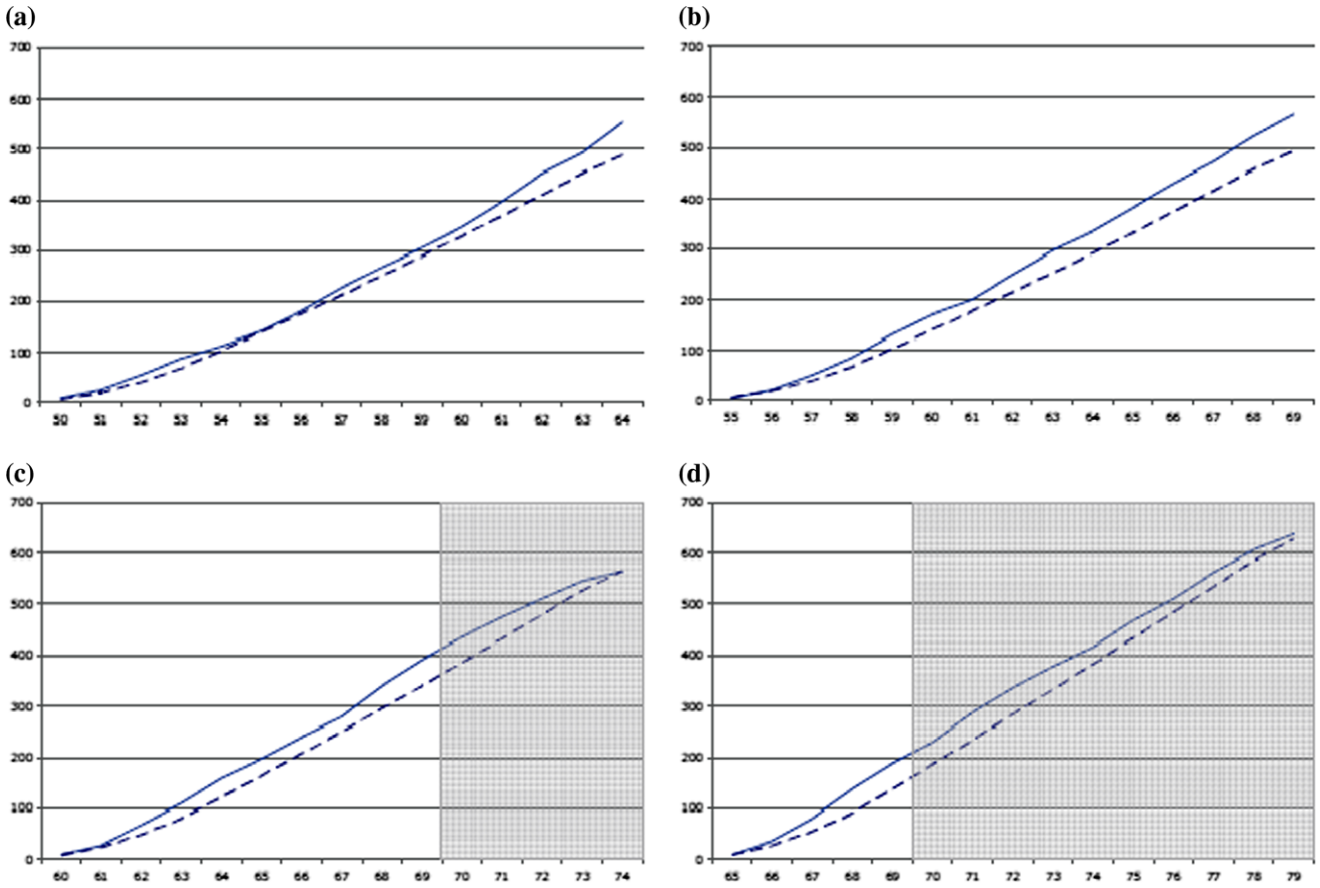

Fig. 2 – Invited (observed) and non-invited (expected) cumulative breast cancer cases by age at the beginning of service screening: (a) 50–54 years, (b) 55–59 years, (c) 60–64 years and (d) 65–69 years. Solid line = observed (O), dotted line = expected (E).

Table 1 – Incidence excess and estimate of overdiagnosis by birth cohorts.

Age at the start of service screening	Years of screening	Incidence excess (95%CI) in the last year of screening	Years after screening stopped	Estimate of overdiagnosis (95%CI)	Cumulative-incidence rates (per 10.000) at 15 years	
					Observed	Expected
50–54	15	1.14 (1.04–1.23)	0	n.e.	359	316
55–59	15	1.15 (1.06–1.25)	0	n.e.	401	348
60–64	10	1.15 (1.04–1.27)	5	1.00 (0.92–1.08)	377	378
65–69	5	1.36 (1.17–1.57)	10	1.02 (0.94–1.10)	412	405

Zackrisson and co-workers[13] and Moss[14] were the only authors to use this method. Zackrisson et al.[13] followed the subjects of the Malmo trial for 15 years after the end of the trial and estimated a rate of overdiagnosis of 10% in women randomised to screening at age of 55–69 years. In Moss[14] a review of estimates of overdiagnosis from RCTs is presented distinguishing between trials in which the control group was or was not screened at the end of the intervention period. In the trials in which the control arm was offered screening there is no evidence of overdiagnosis as a result of incidence screens. In the trials in which the control arm was not offered screening, an excess incidence of breast cancer remains after many years of follow-up: in the two Canadian trials the ratio of the incidence in the intervention arm to that in the control arm was 1.11 and 1.14, respectively. More detailed analysis of the Swedish Two-county and Gothenburg Trials is presented by Duffy et al.[15] showing the catching-up of the two groups and the possible reduction of invasive carcinomas due to the diagnostic anticipation of *in situ* cancers.

In the evaluation of service screening a control group is not available and the incidence curve in the absence of screening needs to be estimated. There is enough agreement that a certain trend was present in the incidence of breast cancer before the start of service screening programmes in Europe.[16] In the IMPACT project, which is an evaluation of service screening in Italy, the estimate of the annual percent

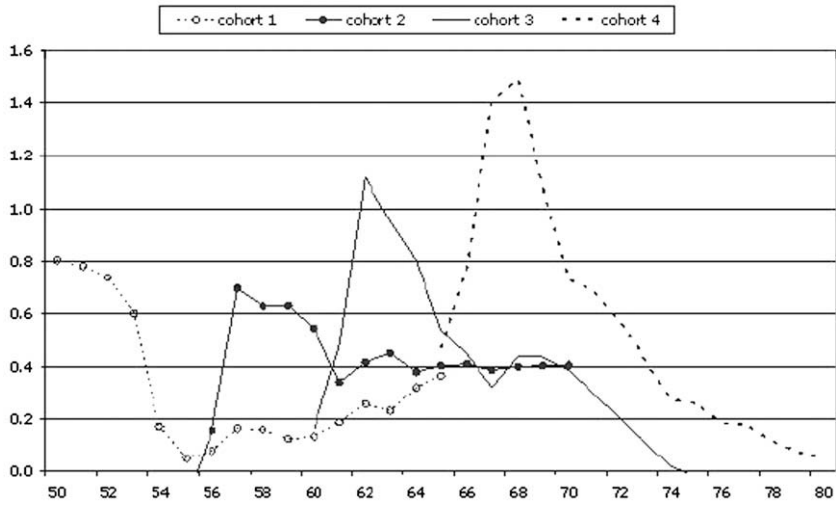

Fig. 3 – Cumulative excess breast cancer rates by birth cohort.

change was 1.2% for all breast cancer cases and was 0.9% for invasive breast cancer only.[11] In the IARC handbook[17] an underlying increase in incidence of 1.5% is reported. Moreover, in some countries a possible change of risk factors for breast cancer, in particular related to the use of hormone replacement therapy, has been suggested. Therefore, the estimation of cancer incidence in the absence of screening is a very complex question. In order to assess how our estimate of overdiagnosis depends upon pre-screening trend estimates, we performed a sensitivity analysis assuming the most extreme scenario (the absence of incidence trend). The results show that the estimate of overdiagnosis is very sensitive to pre-screening trend estimates and, consequently, to the expected incidence estimate. On the other hand, our data show that the degree of overdiagnosis was nearly zero assuming an increasing trend and more than likely lower than 13% (no trend) for all breast cancers (including carcinoma *in situ*).

It should be noted that a deficit in incidence above the age limits for screening can only occur in cohorts that have actually been through the screening programme. One cannot expect, therefore, to observe a significant reduction in incidence at older ages immediately after the first round of screening.[12] The implications are that an evaluation of overdiagnosis in the short term can be misleading, and that when long-term data are not available a statistical correction for lead time is needed.

In Italy we evaluated a model based on the statistical adjustment for lead time of the observed incidence rate, and found risks of overdiagnosis in the Florentine programme of 5% for *in situ* and invasive cases and 2% for invasive case only.[6] More recently in the multicentre Italian IMPACT project,[11] using an improved methodology, we evaluated service screening in several areas in Italy estimating an overdiagnosis of 4.6% for *in situ* and invasive cases and 3.2% for invasive case only.

To take into account both the benefit and the harm of breast cancer service screening, we assumed that the estimate of overdiagnosis ranged from 0% (best evidence) to 13% (no incidence trend) and that the estimated reduction in breast cancer mortality for the invited group was 25%, as we estimated in a case-control study of the Italian IMPACT project.[18] In a population where the risk of breast cancer between 50 and 79 years is 6.5% and the risk of dying from breast cancer in the same age class is 2.5% (both risks were estimated in the area before the start of screening), inviting 1000 women may prevent about six breast cancer deaths of 25 expected (1:4), but would lead to an overdiagnosis, in the worst and most unlikely scenario, of up to 8 cases of 65 expected *in situ* and invasive breast cancer cases (1:8). Under the assumption of 1.2% annual trend, less than one overdiagnosed cancer is expected.

In conclusion, our results show that the degree of overdiagnosis estimated in service screening is within the range estimated in other studies, including those based on RCTs[13,14] and those which use the statistical adjustment for lead time method.[6,11] As many factors may influence incidence trend estimates in the absence of screening, these data suggest that overdiagnosis in breast cancer screening is likely to be a minor phenomenon when compared with other screenings for cancer, as for prostate cancer.[19] The estimate of the benefits and harms in service screening strongly support the population-based screening policy.

Funding source

Institutional funding (no role).

Conflict of interest statement

All authors declare to have neither personal/financial conflict of interest nor relationship with other people or organisations that could inappropriately influence their work.

Acknowledgments

Donella Puliti was partially supported by the Master in Epidemiology, University of Turin, and San Paolo Foundation.

REFERENCES

1. Parkin DM, Moss SM. Lung cancer screening: improved survival but no reduction in deaths – the role of "overdiagnosis". *Cancer* 2000;**89**:2369–76.
2. Morrison AS. *Screening in chronic disease. Monographs in epidemiology and biostatistics*, vol. 7. New York (NY): Oxford University Press; 1985.
3. Zappa M, Ciatto S, Bonardi R, Mazzotta A. Overdiagnosis of prostate carcinoma by screening: an estimate based on the results of the Florence Screening Pilot Study. *Ann Oncol* 1998;**9**:1297–300.
4. von Karsa L, Anttila A, Ronco G, et al. Cancer screening in the European Union. Report on the implementation of the Council Recommendation on cancer screening – First Report; 2008. <http://ec.europa.eu/health/ph_determinants/genetics/documents/cancer_screening.pdf>.
5. Biesheuvel C, Barratt A, Howard K, Houssami N, Irwig L. Effects of study methods and biases on estimates of invasive breast cancer overdetection with mammography screening: a systematic review. *Lancet Oncol* 2007;**8**:1129–38.
6. Paci E, Warwick J, Falini P, Duffy SW. Overdiagnosis in screening: is the increase in breast cancer incidence rates a cause for concern? *J Med Screen* 2004;**11**:23–7.
7. Jonsson H, Johansson R, Lenner P. Increased incidence of invasive breast cancer after the introduction of service screening with mammography in Sweden. *Int J Cancer* 2005;**117**:842–7.
8. Giorgi D, Ambrogetti D, Bianchi S, et al. Design and preliminary results of the Florence Breast Cancer Screening Programme (Progetto Firenze Donna). *Eur J Cancer Prev* 1994;**3**(Suppl 1):29–34.
9. Falini P, Giorgi D, Ventura L, Piccini P, Zappa M. I programmi di screening mammografico in Toscana. In: *"I programmi di screening della regione toscana (Ottavo rapporto)". A cura di Paola Mantellini*. CSPO; 2007.
10. Ministero della Salute. Osservatorio Nazionale screening. <http://www.osservatorionazionalescreening.it/>.
11. Paci E, Miccinesi G, Puliti D, et al. Estimate of overdiagnosis of breast cancer due to mammography after adjustment for lead time. A service screening study in Italy. *Breast Cancer Res* 2006;**8**:R68.
12. Paci E, Duffy S. Overdiagnosis and overtreatment of breast cancer: overdiagnosis and overtreatment in service screening. *Breast Cancer Res* 2005;**7**:266–70.
13. Zackrisson S, Andersson I, Janzon L, Manjer J, Garne JP. Rate of overdiagnosis of breast cancer 15 years after end of Malmo mammographic screening trial: follow-up study. *BMJ* 2006;**332**:689–92.
14. Moss S. Overdiagnosis and overtreatment of breast cancer: overdiagnosis in randomised controlled trials of breast cancer screening. *Breast Cancer Res* 2005;**7**:230–4.
15. Duffy SW, Agbaje O, Tabar L, et al. Overdiagnosis and overtreatment of breast cancer: estimates of overdiagnosis from two trials of mammographic screening for breast cancer. *Brit Cancer Res* 2005;**7**:258–65.
16. Waller M, Moss S, Watson J, Møller H. The effect of mammographic screening and hormone replacement therapy use on breast cancer incidence in England and Wales. *Cancer Epidemiol Biomark Prev* 2007;**16**:2257–61.
17. International Agency for Research on Cancer Expert Group. Breast Cancer Screening. In: *IARC handbooks of cancer prevention*, vol. 7. Lyon, France: IARC; 2002. p. 248.
18. Puliti D, Miccinesi G, Collina N, et al. Effectiveness of service screening: a case-control study to assess breast cancer mortality reduction. *Brit J Cancer* 2008;**99**:423–7.
19. Schroder FH, Hugosson J, Roobol MJ, et al. Screening and prostate-cancer mortality in a randomized European study. *New Engl J Med* 2009;**360**:1320–8.