Methodological approaches for the estimate of overdiagnosis in mammography screening: the cumulative incidence method and the Florence city (Italy) experience

Marco Zappa, Donella Puliti and Eugenio Paci
Overdiagnosis and breast cancer

“Detection of in situ or invasive breast cancers that would have never clinically surfaced in the absence of screening”

It’s the combination of two causes:

1. the natural history of the disease (detection of not progressive cancers)

2. the presence of competing causes of death (detection of potentially progressive cancer in a woman who is going to die from other causes in the near future)

Paci and Duffy, Breast Cancer Research, 2005
A clear distinction must be done between:

**Excess of incidence due to lead time**, during the screening period **and** **overdiagnosis**, i.e. the detection of cancers at screening that would never have clinically surfaced in the absence of screening.
In the service screening context the possible reasons for an observed excess of incidence are:

1) increasing incidence trend before the start of service screening;
2) peak in incidence due to prevalence screening;
3) peak of incidence in subsequent screening
4) a small peak due to women at first screening (new entry for age or migration) in the subsequent rounds;
5) a shift in the age-incidence curve due to lead time;
6) overdiagnosis.

Can we disantangle overdiagnosis from excess of incidence?

Paci and Duffy, Br Cancer Res, 2005
Measures of overdiagnosis

- Overdiagnosis in a population invited to screening
- Increase risk of having a BC diagnosis for women undergoing screening
- % of screen detected cases that are overdiagnosed
.. IN ORDER TO ESTIMATE OVERDIAGNOSIS WE NEED:

**Analysis of a fixed cohort:** Invited to screening or actually screened.

**Comparison group:** Women with the same age who are not screened (not invited) over the same time period with a similar underlying risk of breast cancer (randomised trial). If figures are derived from observational data, adjustments for different breast cancer risk are needed (e.g. time trends).

**Long follow up:** Sufficient follow-up after the last screen (5 years or more) - cumulative-incidence method; otherwise lead-time bias should be adjusted for with statistical methods.
METHODS TO ESTIMATE OVERDETECTION:

A) “The theoretically most robust method to estimate overdetection is the cumulative-incidence approach with data from a randomised controlled trial, in which there is more than several years of follow-up after screening stops, and the control group is never screened.”

B) “If there is little or no follow-up after the last screen, there will be lead-time bias that should be adjusted for statistical methods, otherwise the estimate of overdetection will be too high.” (adjusted for lead-time method)
Figure 1.
Effect of biennial screening of women 50-68 years on incidence of invasive breast cancer in the absence of overdiagnosis.

Several years after the screening end, if there is no overdiagnosis, the cumulative incidence will be identical in the two groups.

Biesheuvel et al, Lancet Oncology, 2007
Figure 2. Effect of biennial screening of women 50-68 years on incidence of invasive breast cancer in the presence of overdiagnosis.

Cumulative incidence method: The comparison of cumulative incidence in the two groups several years after screening stops is a valid estimate of overdiagnosis.

Biesheuvel et al, Lancet Oncology, 2007
APPLICATION OF THE CUMULATIVE-INCIDENCE METHOD TO FLORENTINA DATA:

**Objective:** To evaluate the degree of overdiagnosis of breast cancer 15 years after the introduction of mammographic service screening in Florence in the year 1991.
**Method:** Cumulative incidence method.

**Measure:** The measure of overdiagnosis is the ratio of cumulative incidence of breast cancers in the invited group at least 5 years after the last screen to that in the non-invited group.

**Invited group (observed):**
Cohort of women aged 50-69 at the beginning of service screening (61,568 women) and follow up for breast cancer incidence between 1991 and 2004.

**Non-invited group (expected):**
A Poisson regression model (with age and calendar year) was fitted to Florentine pre-screening incidence data (1986-1990) and the annual trend was forced to 1.2% (pooled estimate in North-Central Italy).
FIGURE 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  

d) 65-69 years
### TABella 1.
Incidence excess and estimate of overdiagnosis by birth cohort.

<table>
<thead>
<tr>
<th>Age at the start of service screening</th>
<th>Years of screening</th>
<th>Incidence excess (95%CI) in the last year of screening</th>
<th>Years after screening stopped</th>
<th>Estimate of overdiagnosis (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50-54</td>
<td>15</td>
<td>1.15 (1.06 to 1.24)</td>
<td>0</td>
<td>n.e.</td>
</tr>
<tr>
<td>55-59</td>
<td>15</td>
<td>1.15 (1.06 to 1.25)</td>
<td>0</td>
<td>n.e.</td>
</tr>
<tr>
<td>60-64</td>
<td>10</td>
<td>1.15 (1.04 to 1.27)</td>
<td>5</td>
<td>1.00 (0.92 to 1.08)</td>
</tr>
<tr>
<td>65-69</td>
<td>5</td>
<td>1.36 (1.18 to 1.57)</td>
<td>10</td>
<td>1.02 (0.94 to 1.10)</td>
</tr>
</tbody>
</table>

1.01 (0.95 - 1.07) for in situ and invasive cases
Sensitivity analysis

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.

In this case, the estimate of overdiagnosis for women 60-69 years at the enrolment was 1.13 (1.07 – 1.19).
BENEFIT AND HARM OF BREAST CANCER SERVICE SCREENING: Invited versus Not invited

Harm: overdiagnosis = 13% (no incidence trend)

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%, inviting 1000 women:
- may prevent about 6 BC deaths (6 lives saved) out of 25 expected
- but could lead to an overdiagnosis, in the worst and most unlikely scenario, of up to 8 cases out of 65 expected in situ and invasive breast cancer cases.

1 Saved life : 1 overdiagnosed cancer
Our results show that the degree of overdiagnosis estimated in service screening (1-13%) is within the range estimated in other studies, including those based on RCTs and those which use the statistical adjustment for lead time method.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Method</th>
<th>Data</th>
<th>Estimate of overdiagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Puliti, 2009</td>
<td>Cumulative incidence method</td>
<td>Florence</td>
<td>1-13%</td>
</tr>
<tr>
<td>Zackrisson, 2006</td>
<td>Cumulative incidence method</td>
<td>Malmo Trial</td>
<td>10%</td>
</tr>
<tr>
<td>Moss, 2005</td>
<td>Cumulative incidence method</td>
<td>Canadian trial I</td>
<td>11%</td>
</tr>
<tr>
<td>Moss, 2005</td>
<td>Cumulative incidence method</td>
<td>Canadian trial II</td>
<td>14%</td>
</tr>
<tr>
<td>Morrell, 2010</td>
<td>Adjusted for lead-time method</td>
<td>New South Wales</td>
<td>30-42%</td>
</tr>
<tr>
<td>Paci, 2006</td>
<td>Adjusted for lead-time method</td>
<td>IMPACT study</td>
<td>4.6%</td>
</tr>
<tr>
<td>Paci, 2004</td>
<td>Adjusted for lead-time method</td>
<td>Florence</td>
<td>5%</td>
</tr>
<tr>
<td>Jonsson, 2005</td>
<td>Adjusted for lead-time method</td>
<td>Sweden</td>
<td>21-52%</td>
</tr>
<tr>
<td>Jorgensen, 2009</td>
<td>Analysis of incidence trends</td>
<td>Meta-analysis</td>
<td>52%</td>
</tr>
</tbody>
</table>

(Cohort: Puliti, Zackrisson, Moss, Moss)  
(Dynamic p.: Morrell, Paci, Paci, Jonsson, Jorgensen)
Study *in progress*

- Cohort study with individual definition of screening exposure:
  
  a) estimate of BC mortality reduction
  b) estimate of overdiagnosis of breast cancer
STUDY DESIGN

Definition of the cohort

Follow-up for vital status and cause of death
All women were followed-up for vital status and cause of death until 31 December 2008 through the linkage with the regional mortality registry and with the list of residence.

Follow-up for breast cancer incidence
All women were followed-up for breast cancer incidence until 31 December 2007 through the linkage with the Tuscan Cancer Registry and Pathology Reports.
Definition of screening exposure

Screening exposure was defined on the basis of attendance at the firsts two rounds and the women were classified in:

1) frequent attenders, if they responded to both invitations;
2) irregular attenders, if they only responded to one invitation;
3) never attenders, if they not responded to any invitation.

For the women invited only at the first round, screening exposure was defined on the basis of the attendance at the first round.

We excluded BC occurred in the first 6 monts only from never attenders group
RESULTS: BC mortality reduction

We selected 51,063 women aged 50-69 years invited at the first screening round in Florence.

The women were classified in:

<table>
<thead>
<tr>
<th>Exposure</th>
<th>N° (%)</th>
<th>Mean n° of test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequent</td>
<td>24,580 (48%)</td>
<td>5.1 - 7.6</td>
</tr>
<tr>
<td>Irregular</td>
<td>7,965 (16%)</td>
<td>3.1 - 5.5</td>
</tr>
<tr>
<td>Never</td>
<td>18,518 (36%)</td>
<td>0.4 - 1.8</td>
</tr>
</tbody>
</table>

84% of “never” did not perform any test in the study period.

Screened (64%)
MORTALITY RATES

The follow-up for vital status was updated at 31 December 2008 with a median follow-up of 16.5 years.
**N° of breast cancer deaths and standardized mortality rates by exposure category**

On the 16-years study period, in total we observed 9,624 deaths for whatever cause and 392 breast cancer deaths.

<table>
<thead>
<tr>
<th>Exposure</th>
<th>BC deaths</th>
<th>Person-years</th>
<th>STD rates (x10.000)</th>
<th>SRR (95%CI) adjusted*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>208</td>
<td>264.636</td>
<td>7.3</td>
<td>1</td>
</tr>
<tr>
<td>Screened</td>
<td>184</td>
<td>503.961</td>
<td>3.6</td>
<td>0.53 (0.43 - 0.66)</td>
</tr>
</tbody>
</table>

* Adjusted for age, deprivation index, marital status, and previous breast examination
RESULTS: estimate of overdiagnosis

We selected 26,514 women aged 60-69 years invited at the first screening round in Florence.
**N° of breast cancer cases and standardized incidence rates by exposure category**

On the 15-years study period, we observed 1182 breast cancer (1110 invasive and 72 in situ)

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Cases</th>
<th>Person-years</th>
<th>STD rates (x1000)</th>
<th>SRR (95%CI) adjusted*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>438</td>
<td>142.542</td>
<td>2.99</td>
<td>1</td>
</tr>
<tr>
<td>Screened</td>
<td>744</td>
<td>216.325</td>
<td>3.44</td>
<td>1.15 (1.02 - 1.30)</td>
</tr>
</tbody>
</table>

* Adjusted for age, deprivation index, marital status, and previous breast MX

**For invasive cases only: 1.10 (0.97-1.24)**
BENEFIT AND HARM FOR A SCREENED WOMAN:  
**Screened versus Never**

**Benefit:** reduction in BC mortality = 47%
**Harm:** overdiagnosis = 15%

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%, **screening 1000 women:**

- may prevent about 12 BC deaths (12 saved lives) out of 25 expected
- could lead to an overdiagnosis of 10 cases out of 65 expected

1 saved life : 1 overdiagnosed cancer
• Thank you!
Adjusted for lead-time cumulative incidence method

If there’s little or no follow-up after the last screen (5 years or more), there will be lead-time bias that should be adjusted for statistical methods.

Assuming an exponential distribution for the *sojourn time*, the probability that a tumour currently detected at screening in the pre-clinical phase would have of progressing to the clinical phase within the next n years is as follows:

\[
\int_{0}^{n} \lambda \cdot e^{-\lambda x} \, dx = 1 - e^{-\lambda n}
\]

thus, it is possible to calculate the probability that each screen-detected case would have been identified clinically each year after detection until the end of the study period.
Number of corrected for lead time cases = \[ \text{n° of observed incidence cases} - \text{n° of screen-detected cases in that year} + \text{estimated n° of SD cases that would have arisen clinically in that year}. \]

Estimate of overdiagnosis

The corrected-for-lead-time cases should be compared with the predicted number in the absence of screening. The percentage excess after correction for lead time is an indicator of overdiagnosis, given the lead time estimate.

Application of the adjusted for lead time method to the IMPACT dataset

Estimate of overdiagnosis of breast cancer due to mammography after adjustment for lead time. A service screening study in Italy
Eugenio Paci, Guido Miccinesi, Donella Puliti, Paola Baldazzi, Vincenzo De Lisi, Fabio Falcini, Claudia Cirilli, Stefano Ferretti, Lucia Mangone, Alba Carola Finarelli, Stefano Rosso, Nereo Segnan, Fabrizio Stracci, Adele Traina, Rosario Tumino and Manuel Zorzi

Available online http://breast-cancer-research.com/content/8/6/R68

Open Access

Research article
Breast cancer incidence rates predicted, observed and observed corrected for lead time.

Age: 50-74 years. Data: Impact study (Firenze, Torino, Parma, Ferrara, Modena e Romagna)

Excess ratio: 4.6% (2% - 7%) all carcinomas (invasive and in situ)
3.2% (1% - 6%) only invasive carinomas
**Person-years**

For each woman we calculated the person-years at risk to experiment the event (breast cancer diagnosis/ breast cancer death):

- Date of first invitation
- Person-years at risk
- Date of the event
- Date of death
- Date of migration
- Date of study end
**FIGURE 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

d) 65-69 years
Growth rates of cancers (IARC, 2002)

The diagnosis of these cancers (very slow and non-progressive), that Morrison (1992) have called “pseudodisease”, is overdiagnosis.