

RESEARCH

Overdiagnosis from non-progressive cancer detected by screening mammography: stochastic simulation study with calibration to population based registry data



OPEN ACCESS

Arnaud Seigneurin *doctor*^{1,2}, Olivier François *professor*², José Labarère *assistant professor*^{2,3}, Pierre Oudeville *student*², Jean Monlong *student*², Marc Colonna *scientific director*¹

¹Registre du Cancer de l'Isère, Centre Hospitalier Universitaire de Grenoble, BP 217, Pavillon E, 38043 Grenoble Cedex 9, France; ²Université Joseph Fourier Grenoble 1, CNRS, TIMC-IMAG UMR 5525, 38041 Grenoble; ³Unité d'évaluation Médicale, Centre Hospitalier Universitaire de Grenoble, Grenoble

Abstract

Objective To quantify the magnitude of overdiagnosis from non-progressive disease detected by screening mammography, after adjustment for the potential for lead time bias, secular trend in the underlying risk of breast cancer, and opportunistic screening.

Design Approximate bayesian computation analysis with a stochastic simulation model designed to replicate standardised incidence rates of breast cancer. The model components included the lifetime probability of breast cancer, the natural course of breast cancer, and participation in organised and opportunistic mammography screening.

Setting Isère, a French administrative region with nearly 1.2 million inhabitants.

Participants All women living in Isère and aged 50-69 during 1991-2006.

Main outcome measures Overdiagnosis, defined as the proportion of non-progressive cancers among all cases of invasive cancer and carcinoma in situ detected 1991-2006.

Results In 1991-2006, overdiagnosis from non-progressive disease accounted for 1.5% of all cases of invasive cancer (95% credibility interval 0.3% to 2.9%) and 28.0% of all cases of carcinoma in situ (2.2% to 59.8%) detected either clinically or by screening mammography in Isère. When analysis was restricted to the cancers detected by screening mammography only, the estimates of overdiagnosis were 3.3% (0.7% to 6.5%) and 31.9% (2.9% to 62.3%) for invasive cancer and carcinomas in situ, respectively.

Conclusion Overdiagnosis from the detection of non-progressive disease by screening mammography was limited in 1991-2006 in Isère. Because carcinoma in situ accounted for less than 15% of all incident breast cancer cases, its contribution to overdiagnosis was relatively limited and imprecise.

Introduction

The net benefit of cancer screening programmes reflects the extent to which the benefits outweigh the harms.¹ Although controversial,^{2,3} evidence derived from randomised controlled trials suggests that mammography screening reduces mortality rates from breast cancer in women aged 50-70.^{4,5} Mammography screening, however, also exposes women to harm, including false positive results, low dose radiation, and overdiagnosis.^{6,7}

Overdiagnosis refers to the detection of histologically confirmed invasive cancers or carcinoma in situ that would never have clinically surfaced in the absence of screening.^{8,9} Overdiagnosis can result from either the detection of non-progressive cancers or competing causes of death, such that a woman will die from another cause before the cancer becomes symptomatic.^{9,10} There is anecdotal evidence of non-progressive or even spontaneously regressive breast cancers,^{11,12} though the underlying physiopathological mechanism deserves further investigation.^{13,14} Because it is not possible to distinguish between progressive and non-progressive cancers, clinicians treat all detected breast cancers, making overtreatment inevitable.¹⁵ For the same reason, overdiagnosis is mainly an epidemiological concept¹⁰ that remains challenging to quantify.⁹

Ideally, overdiagnosis could be estimated by comparing the cumulative incidence of breast cancers between screened and unscreened women enrolled in a randomised controlled trial with lifelong follow-up.^{15,16} An excess of cases in the screened group would reflect the magnitude of overdiagnosis at the end of follow-up. Published trials have yielded potentially flawed estimates of overdiagnosis because of an inappropriate duration of follow-up or provision of mammography screening to the control group at the end of the intervention period.¹⁶ Apart from the controlled conditions of randomised trials, estimating overdiagnosis associated with population based screening

programmes relies on the comparison of the incidence rate after implementation with that expected in the absence of screening, which is often estimated from trends before screening.^{16 17} Yet an increase in the incidence rate of breast cancer could reflect overdiagnosis as well as a change in the background incidence trend or lead time from early detection of both prevalent and incident cancers.

Because of the inherent limitations of these approaches, various innovative modelling and simulation techniques have been developed for integrating relevant information from separate sources to evaluate the potential benefits and harms of mammography screening.^{18 19} Although few studies have been based on bayesian updating for elucidating the relative contribution of mammography screening to the decline in mortality,^{20 21} none was primarily designed for estimating overdiagnosis.

We quantified the magnitude of overdiagnosis resulting from the detection of non-progressive cancers by screening mammography. For this purpose, we used an approximate bayesian computation approach²² that relied on extensive computer simulations to infer this unobserved quantity.

Methods

Study design and model overview

We developed a stochastic simulation model designed to replicate standardised incidence rates of breast cancer in 1991-2006 in Isère, France. This model was populated by 245 000 women divided into birth cohorts making up a female population aged 50-69 from 1991 to 2006. For each woman, we simulated the occurrence of breast cancer, its natural course, and the participation in mammography screening (fig 1⇓). We hypothesised that both invasive cancers and carcinoma in situ could be progressive or non-progressive.

This model was suited for quantifying overdiagnosis, defined by the proportion of cases that would not have clinically surfaced during 1991-2006, because of the lack of progressive potential, among all cases arising in women from the age of 50 to 69. We focused on non-progressive diseases; overdiagnosis resulting from progressive cancers censored because of competing causes of death was not in the scope of the present study. The proportion of non-progressive cancers was handled as an unknown parameter in the model and estimated with an approximate bayesian computation approach. Approximate bayesian computation has its roots in rejection sampling, a method that is suited for solving complex problems with mathematically or computationally intractable likelihood.^{23 24} In this application of approximate bayesian computation, we first simulated large numbers of female population datasets using model parameters drawn from prior distributions. For each simulated dataset, we computed a set of summary statistics (annual standardised incidence rates of invasive cancer and carcinoma in situ, 1991-2006) and compared them with the values observed in Isère. Then we obtained an approximate sample from the posterior distribution by selecting the model parameters that had generated the summary statistics closest to the observed values. Eventually, the proportion of non-progressive cancers, like the other unknown parameters of the model, was derived from this approximate sample.

Model components and parameters

We assumed that the observed incidence rate of breast cancer was driven by four major components including all cause mortality, lifetime probability of breast cancer, the natural course

of breast cancer, and the detection of breast cancer either clinically or by screening mammography. Each of these components was modelled as a stochastic process constrained by general knowledge and published data.^{20 21}

All cause mortality

Starting the model in 1922, we simulated 7000 women for each single year birth cohort, resulting in a population size of 245 000 women. For each woman, we generated survival time and death from any cause using a Cox-Gompertz model with a mean life expectancy of 80.²⁵

Lifetime probability of breast cancer

For each woman, we simulated the occurrence of breast cancer based on a lifetime probability distribution. In accordance with previous studies,²¹ we modelled only the incidence of first breast cancer. Although the simulations involved 1922-56 birth cohorts, the lifetime risk of breast cancer was calculated for women born from 1900 to 1956. Indeed, the model was also designed for other analyses that required simulations of birth cohorts from 1900. The lifetime probability of cancer incorporated the possibility of a baseline incidence for women born in 1900 as well as an increasing trend in incidence for women born from 1900 to 1950. Because the most recently published risk estimates of developing an invasive breast cancer before age 75 were limited to women born before 1950,²⁶ we assumed that women born in 1950-6 had a lifetime risk of cancer similar to those born in 1950. To account for uncertainty, we used prior uniform distributions for parameters describing the lifetime probability distribution of breast cancer (table 1⇓).

As an increasing secular trend and similar changes in risk factors have been consistently reported for carcinomas in situ and invasive cancers,²⁷ we hypothesised that the underlying risk for both types of cancer followed an increasing trend. Yet, we allowed the trend in lifetime probability to differ between invasive cancers and carcinomas in situ. For this purpose, we sampled a value from a uniform (0.0, 1.0) prior distribution, corresponding to the relative contribution of carcinoma in situ to the overall trend in incidence of breast cancer. A value of 0.0 implied that invasive cancers and carcinomas in situ had the same overall trend in lifetime probability of breast cancer, whereas higher values implied a higher trend for carcinomas in situ than for invasive cancers.

Natural course of breast cancer

The natural course of invasive cancers and carcinomas in situ was simulated, with both having a progressive versus non-progressive potential. Consistent with previous studies,²⁸ we assumed three different types of carcinoma in situ, including non-progressive carcinoma in situ, progressive carcinoma in situ that was clinically detected, and carcinoma in situ that progressed to invasive cancer during the preclinical phase. Invasive cancer could be either progressive or non-progressive. For each simulation, the model's parameters corresponding to the proportion of non-progressive invasive cancers and carcinomas in situ were sampled from two independent uniform (0.0%, 50.0%) prior distributions. Basically, there were several potential reasons for investigating the magnitude of overdiagnosis for invasive cancers and carcinomas in situ separately. Firstly, the natural course of this disease remains partly unknown because most detected cancers are treated. Secondly, the magnitude of overdiagnosis from non-progressive disease might differ for invasive cancers and carcinomas in situ. Thirdly, carcinomas in situ account for less than 15% of all

incident breast cancer cases. Thus the estimate of overdiagnosis would be driven by non-progressive invasive cancers if the model did not distinguish between invasive cancers and carcinomas in situ.

We computed age at onset of preclinical cancer using a scaled β distribution with a range of 100 years. This is the same approach used by other authors for determining age at diagnosis.²⁹ Because the age at onset of preclinical cancer cannot be directly observed, we incorporated uncertainty by using uniform prior distributions for the two parameters of the scaled β distribution.

We used a γ distribution for modelling sojourn times (time spent in the preclinical detectable phase) in the base case analysis, with values that were consistent with published estimates of mean sojourn time (range 2-4 years).³⁰⁻³³

Cancer detection

Breast cancer might be detected either clinically or by screening mammography. In accordance with previous studies,²¹ we did not differentiate between the various pathways leading to the clinical detection of breast cancer, which included diagnosis after symptomatic presentation, self breast examination, clinical breast examination, or incidental presymptomatic detection.

In Isère, mammography screening was opportunistic before 1991, while both opportunistic and organised mammography screening coexisted from 1991 to 2006. For each woman, we simulated the probability of undergoing a screening mammography during a two year period, on either an opportunistic or organised basis. Although accurate information exists on participation in organised mammography screening in Isère,³⁴ data on opportunistic screening are scarce.

In this base case analysis, we assumed that opportunistic screening accounted for 0-40% of all screening mammographies. The upper limit (that is, 40%) was extracted from published surveys.³⁵ Based on this assumption, the lowest simulated percentage of women undergoing a mammography for the 2005-6 period was 46%, corresponding to the participation rate in organised screening (46%) with no opportunistic screening (0%). The upper simulated percentage of women undergoing a mammography for the same period was 74%, corresponding to a 46% participation rate in organised screening with an additional 28% rate of opportunistic screening. In the latter simulation, opportunistic screening accounted for almost 40% (28%/74%) of all women who underwent screening mammography. We sampled the probability of undergoing a screening mammography over a two year period from two separate and independent uniform prior distributions, corresponding to the participation rates in mammography screening in 1991-2 and 2005-6, respectively. Finally, we hypothesised that screening mammography dissemination followed a linear trend from 1991 to 2006.

We also had to consider opportunistic screening before the age of 50 and at the end of the 1980s to analyse the incidence rates among women aged 50-69 during the 1991-2006 period. Indeed, participation in screening before the age of 50 or before the 1991-2006 period could have modified the incidence rates among women aged 50-69 by anticipating the date of diagnosis. For example, a cancer detected by opportunistic screening mammography at age 49 would have surfaced clinically at age 51, assuming a lead time of two years.

We first considered that about 20% of women underwent an opportunistic screening mammography every two years between 1987 and 1990³⁶ and drew the corresponding parameter from a uniform (15.0%, 25.0%) prior distribution. Secondly, the

probability of undergoing an opportunistic screening mammography over a two year period in the age range 40-49 was assumed to be half of the participation rate in organised mammography screening among women aged 50-69. Indeed, women aged 40-49 were not eligible for organised mammography screening, but a substantial proportion underwent opportunistic screening.

Finally, the mean sensitivity of mammography was assumed to be 90%, with the value drawn from a β (31.5, 3.5) distribution.

Approximate bayesian computation analysis

In standard bayesian inference,³⁷ the posterior probabilities of parameters θ are given by the following formula:

$$P(\theta|y) \propto P(y|\theta) \pi(\theta)$$

with y denoting the values observed in an empirical dataset, $P(\theta|y)$ denoting the posterior distribution for each parameter θ , $\pi(\theta)$ denoting the prior distribution for each parameter θ , and $P(y|\theta)$ denoting the likelihood.

For complex problems with mathematically or computationally intractable likelihoods, approximate bayesian computation (ABC) approaches bypass exact likelihood calculations by using simulations and rejection sampling based on summary statistics.²³ Summary statistics (s) are values (standardised annual incidence rates in the present application) that represent the information available in the study.

For this purpose, a basic rejection algorithm performs the following steps:

- 1) Sample a candidate value for each parameter θ from its prior distribution $\pi(\theta)$
- 2) Simulate a dataset made of y_i values from a generating mechanism based on the parameters θ , and compute the corresponding set of summary statistics $s_i = s(y_i)$
- 3) Accept the value of θ if the distance between the summary statistics derived from the simulated dataset s_i and the summary statistics derived from the observed dataset s_{obs} is less than ϵ , a prespecified error value
- 4) If not, reject the parameters θ and go to 1).

Using this basic rejection algorithm, the accepted values (θ_i) form a random sample from an approximation of the posterior distribution.

The posterior probabilities are given by the following formula:

$$p_{\epsilon}(\theta | y) \propto \Pr(|s_i - s_{\text{obs}}| < \epsilon | \theta) p(\theta)$$

Recent improvements in the approximation of the posterior distribution include non-linear transformation of the accepted values of the parameters (θ_i).²⁴ Non-linear transformation consists of weighting the accepted values of the parameters, θ_i , by a quantity that depends on the distance between s_i and s_{obs} and then deriving corrected values of the parameters (θ_i^*) from a non-linear regression:

$$\theta_i^* = \theta_i + g(W, s_{\text{obs}}) - g(W, s_i)$$

where $g(W, s)$ denotes a feed-forward neural network regression function, with weights W adjusted from the simulated data.²⁴

In practice, the proportion of non-progressive cancer was determined via model calibration to the standardised annual incidence rates of invasive cancers and carcinoma in situ observed between 1991 and 2006 in Isère.³⁴ We first simulated 100 000 datasets with model parameters drawn from their prior distributions. For each simulated dataset, we computed a set of summary statistics, including 16 standardised annual incidence rates (from 1991 through 2006) for invasive cancers and carcinoma in situ, respectively. Then, we retained 500 (0.5%)

datasets with the smallest Euclidean distance between the simulated and observed values of standardised incidence rates. The values of the parameters were then corrected to form an approximate sample from the posterior distribution.²⁴

Datasets were re-simulated with the simulation model by using the corrected values of the parameters obtained previously. The posterior predictive distributions of overdiagnosis (%) were calculated among all cases of breast cancers detected clinically or by screening mammography in women aged 50-69 between 1991 and 2006 in Isère.

To assess the robustness of these estimates, we performed supplementary analyses. Firstly, we estimated the proportion of overdiagnosis using various sojourn time distributions and parameters (models 1-7, table 2). Secondly, we repeated the analyses after excluding the 1991-5 study period, which was dominated more by prevalent screens at the beginning of the screening programme. Thirdly, we checked the influence of the prior distribution for non-progressive carcinoma in situ on overdiagnosis estimates by drawing this parameter from a uniform (0.0-0.7) distribution. Finally, we checked for the robustness of our estimates by simulating 200 000 datasets instead of 100 000 and using artificial observations simulated on the basis of the parameter values derived from the best fitting simulations.

All datasets were generated with a C language code and approximate bayesian inference was performed with the ABC package³⁸ in the R statistical software, version 2.12 (R foundation for statistical computing, Vienna, Austria).

Results

We simulated 100 000 datasets, each comprising 245 000 women. Besides the proportion of non-progressive disease, two parameters drawn from prior distributions—the participation rate in mammography screening and the sensitivity of mammography—most influenced the magnitude of overdiagnosis from non-progressive invasive cancers. Overdiagnosis increased with increasing participation rates in organised or opportunistic screening, but the change in overdiagnosis was moderate for participation rates higher than 40%. It also increased with increasing values of mammography sensitivity, with a slower trend for sensitivity values higher than 80%.

Figure 2 depicts the best fitting predicted incidence rates of breast cancer between 1991 and 2006 and the mean standardised incidence rates for the 500 incidence curves predicted from the posterior distributions. Although calibration was satisfactory for invasive cancers, the simulated incidence rates were slightly higher than the observed incidence rates for carcinoma in situ. Compared with the observed incidence rates, the mean predicted incidence rates for 1991-2006 were 30.6% and 4.6% higher for in situ and invasive cancers, respectively. The incidence rate curves obtained from the best fitting dataset, however, were close to the observed incidence rates for both in situ and invasive cancers: the mean predicted incidence rates for 1991-2006 were 8.7% and 1.3% lower for in situ and invasive cancers, respectively, than the observed incidence rates.

Posterior distribution of parameters

The mean proportion of non-progressive cancers was 3% and 6% for invasive cancers and carcinomas in situ, respectively (table 1). The posterior estimate of the proportion of non-progressive cancer was less precise for carcinoma in situ

(95% credibility interval 0% to 17%) than for invasive cancer (3% to 4%).

Posterior estimates of overdiagnosis

In the base case analysis, overdiagnosis accounted for 1.5% of all cases of invasive cancers and for 28.0% of all cases of carcinomas in situ detected either clinically or by screening mammography in Isère for 1991-2006 (table 1). The estimates were more precise for invasive than for in situ cancers (fig 3). The estimates of overdiagnosis obtained from the best fitting dataset were 0.6% and 15.7% for invasive cancer and carcinoma in situ, respectively.

When we restricted the analysis to the cases of cancer detected by screening mammography only, the estimates of overdiagnosis were 3.3% and 31.9% for invasive cancer and carcinoma in situ, respectively (fig 4). In the best fitting analysis, the mean time between the detection of a non-progressive breast cancer by screening mammography and a woman's death was 19 years (SD 10 years).

Supplementary analyses

Firstly, varying the distribution of sojourn time yielded mean estimates of the proportion of overdiagnosis that ranged from 0.0% to 3.9% for invasive cancer and from 12.4% to 51.7% for carcinomas in situ (table 2). Secondly, excluding data from 1991-5 resulted in estimates of overdiagnosis that were unchanged for carcinoma in situ (30.8%, 95% credibility interval 8.5% to 55.4%) and slightly lower for invasive cancer (0.9%, 0.6% to 1.2%) compared with the base case analysis. Thirdly, using a (0.0%, 70.0%) uniform prior distribution for non-progressive carcinoma in situ dramatically altered the point estimate of the overdiagnosis proportion (7.5% instead of 28%) but did not improve the precision of the corresponding 95% credibility interval (0.0% to 45.1%). Finally, increasing the number of simulations from 100 000 to 200 000 did not modify the estimates of overdiagnosis, and the values of the parameters used to simulate artificial observations were correctly estimated with 100 000 simulations.

Discussion

Main findings

Using an approximate bayesian computation approach, we found that overdiagnosis from the detection of non-progressive disease by screening mammography ranged from 1.5% for invasive cancers to 28% for carcinoma in situ. Because carcinomas in situ accounted for less than 15% of all incident breast cancers, overdiagnosis was of limited importance between 1991 and 2006 in Isère.

Strengths and weaknesses

The strengths of the present modelling approach included adjustment for potential sources of bias and calibration to observed incidence rates of breast cancers. The model accounted for important biases that can affect estimates of overdiagnosis in randomised controlled trials and population based screening programmes.¹⁶ Firstly, we addressed the issue of opportunistic screening, which could contribute to underestimating overdiagnosis,¹⁶ by simulating the probability of undergoing screening mammography on either an organised or opportunistic basis. Secondly, because estimates of overdiagnosis might be flawed by secular changes in background risk of breast cancer, the model allowed for the possibility of an increasing linear trend in the lifetime probability of breast cancer. Thirdly, the

model was adjusted for lead time by simulating sojourn times with various distributions in the base case and sensitivity analyses.

Moreover, the uncertainty concerning the extent of opportunistic screening and the lifetime risk of breast cancer was taken into account through the bayesian approach instead of making restrictive assumptions. Our assumption that the risk of breast cancer was constant for women in 1950-6 probably had a limited impact on the estimates of overdiagnosis. Indeed, only six birth cohorts were involved, comprising 6.6% of the overall number of person years generated by the simulations. Assuming that the risk of breast cancer for women born in 1950-6 increased at the same rate as for those born before 1950 would yield a 19.4% lifetime risk instead of 18%.

Another important feature of the approach we used was the use of bayesian rejection sampling according to observed standardised incidence rates of breast cancer for determining posterior distribution of unknown parameters, like the proportion of non-progressive disease. That these estimates of overdiagnosis were consistent with the incidence rates of breast cancer observed between 1991 and 2006 in Isère supports the validity of the results. We cannot exclude that the findings might be different in other countries because of the specificities relative to the breast cancer epidemiology, screening procedures, and the rate of participation in the organised mammography screening programme in Isère.

Firstly, with an estimated standardised rate of 99.7 per 100 000 person years in 2008, France ranked among the countries with the highest incidence of breast cancer worldwide.³⁹ In Isère, the standardised incidence rate was 97.8 per 100 000 person years in 2003-6.⁴⁰

Secondly, most Western countries have developed breast cancer screening programmes but the provision of screening mammography varies markedly, with different settings (private practices versus national health services), ages of women screened, intervals between mammographic examinations, and numbers of mammographic views. In Isère, the organised screening programme was launched in 1991 and consisted of a single oblique external view mammography for women aged 50-69. Two view mammography (craniocaudal and oblique external) was introduced in 2000 for both first and subsequent screens. In 2002, the programme was extended to women aged 50-74, included clinical breast examination, and had a screening interval shortened from 30 to 24 months to comply with the requirements for the French nationwide programme.⁴¹

Thirdly, the low participation rate in the early 1990s (25% in 1991-3 and 30% in 1993-5) and the coexistence of opportunistic individual screening constitute another characteristic of the breast cancer screening programme in Isère.

Uncertainty and complexity are the main limitations of this modelling approach. Because the model incorporated uncertainty for various unknown parameters, the estimated proportions of non-progressive disease were uncertain, especially for carcinomas in situ. The relatively low incidence rates of carcinoma in situ also probably contributed to the imprecise estimate of overdiagnosis for this disease, as reflected by the large credibility interval. This could explain the difference observed between the best fitting (15.7%) and the mean (28%) estimates of overdiagnosis for carcinomas in situ. Another potential explanation for this observation concerns the skewed distribution of this parameter. Indeed, the best fitting estimates were close to the mode of the distribution, which was somewhat different from the mean.

Complexity is considered a major drawback of the modelling approach for quantifying the magnitude of overdiagnosis. Indeed complex models are often required to capture true disease processes, making the assessment of their validity problematic.¹⁶ Although less complex models seem more transparent, they might oversimplify the true situation and lead to inaccurate estimates.

Finally, the number of simulations needed by rejection sampling methods might be substantial as the number of independent summary statistics increases, a problem that is called the "curse of dimensionality." In addition, some of the 12 parameters we estimated in our model probably exhibit non-negligible correlations. The choice of using feed-forward neural networks deals with these concerns specifically, and the improvements on approximate bayesian computation with those non-linear transformations has been extensively discussed²⁴ for examples of similar complexity.

Comparison with other studies

Inconsistencies in published percentage estimates of overdiagnosis result partly from the use of different denominators across studies.⁴² Indeed, some authors reported the percentages of overdiagnosis observed among all cases of cancer diagnosed either clinically or by screening mammography, while others reported the percentages of overdiagnosis observed among the cases of cancer detected by screening mammography only. To ensure comparability, we recalculated the percentage of overdiagnosis among all cases of cancer diagnosed either clinically or by screening mammography, based on published data for each study. Overdiagnosis was defined as the incidence in the screened population minus the incidence in the unscreened population, divided by the incidence in the screened population.

Published estimates of overdiagnosis range from less than 1% to 35% for invasive breast cancers and from 4% at the incident screen to 46% at the prevalent screen for carcinomas in situ. The various approaches and related biases probably explain much of the variability in the estimates of overdiagnosis, although we cannot quantify the contribution of heterogeneity in the sensitivity of the screening programme and baseline characteristics of the population.

These estimates of overdiagnosis for invasive cancers were consistent with those from well designed randomised controlled trials and population based studies that compared the cumulative incidence of breast cancer between screened and unscreened women. The overdiagnosis rate was 6.5% (95% confidence interval -4.2% to 15.3%) in the Malmö mammographic trial in Sweden⁴³ and 1.7% (-10.2% to 11.5%) for women aged 50-59 enrolled in the Canadian National Breast Screening Study.⁴⁴ Comparable estimates were reported in a population based screening programme in Florence, Italy, after adjustment for lead time.⁴⁵ Yet, comparing the cumulative incidence of breast cancer between a screened cohort and an age matched historical control cohort, Zahl et al found a much higher overdiagnosis rate (18.0%, 13.8% to 23.1%), which might reflect the effect of lead time resulting from a limited follow-up period.⁴⁶

Our results also agree with those reported by authors who used multistate models with explicit assumptions regarding the natural course of disease and sojourn times. Overdiagnosis rates reported by Olsen et al were 7.8% (0.3% to 28.5%) and 0.5% (0.0% to 2.1%) at the prevalent and incident screen in the Copenhagen screening programme in Denmark.³³ Using a similar approach, Duffy et al found that overdiagnosis rates at the prevalent and incident screen were 3.1% (0.1% to 10.9%) and

0.3% (0.1% to 1.0%) in the two county trial and 4.2% (0.0% to 28.8%) and 0.3% (0.0 to 2.0%) in the Gothenburg trial in Sweden.⁴⁷ Yen et al quantified the magnitude of overdiagnosis for carcinoma in situ only, with estimates ranging from 23% to 46% for the prevalent screen and from 4% to 21% for the incident screen.³²

In contrast, studies that compared the trend in incidence rates before and after implementation of screening reported much higher estimates of overdiagnosis for invasive cancers. An analysis of five screening programmes organised in Europe, Australia, and Canada estimated the overdiagnosis rate at 35% (29% to 42%).¹⁵ None of the screening programmes, however, systematically adjusted for lead time other than the exclusion of the implementation phase with its prevalence peak. Similar rates were observed in a screening programme in Norway, which could be explained by the lack of precise adjustment for lead time and changes in background incidence.⁴⁸ Studies that compared incidence rates after adjusting for lead time based on explicit, although unverified, assumptions, however, consistently reported lower rates of overdiagnosis.⁴⁹

The magnitude of overdiagnosis for carcinoma in situ might have been overestimated in the previous years. Indeed, our estimate of overdiagnosis for carcinomas in situ was consistent with the 32% recurrence rate after surgical treatment reported recently, with ipsilateral and invasive recurrences accounting for 81% and 40% of all recurrences.⁵⁰ These findings probably correspond to the lower range of progressive disease among treated carcinomas in situ.

Conclusion

Overdiagnosis of invasive breast cancer in a population offered organised and individual screening was 1.5% (95% credibility interval 0.3% to 2.9%) after adjustment for the effect of lead time, the uncertainty around the trend in incidence of breast cancer for successive birth cohorts, and opportunistic screening. Further study is warranted to obtain a more precise estimate of overdiagnosis for carcinoma in situ.

Linda Northrup from English Solutions (Voiron, France) provided assistance in preparing and editing the manuscript.

Contributors: AS, OF, JL, and MC designed and planned the study. OF, AS, PO, and JM analysed the data. AS wrote the paper. All authors contributed to the drafting of the paper. AS is guarantor.

Funding: This study was funded by grants from the Institut National du Cancer, Paris, France, and the Comité de l'Isère de la Ligue Nationale Contre le Cancer, Grenoble, France. The study sponsor had no role in the study design, collection, analysis, and interpretation of the data; or in the writing of the article and decision to submit the article for publication.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: Not required.

Data sharing: Statistical code is available from the corresponding author.

- 1 Quanstrum KH, Hayward RA. Lessons from the mammography wars. *N Engl J Med* 2010;363:1076-9.
- 2 Gotzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet* 2000;355:129-34.
- 3 Jorgensen KJ, Zahl PH, Gotzsche PC. Breast cancer mortality in organised mammography screening in Denmark: comparative study. *BMJ* 2010;340:c1241.

- 4 Humphrey LL, Helfand M, Chan BK, Woolf SH. Breast cancer screening: a summary of the evidence for the US Preventive Services Task Force. *Ann Intern Med* 2002;137:347-60.
- 5 Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK, Humphrey L. Screening for breast cancer: an update for the US Preventive Services Task Force. *Ann Intern Med* 2009;151:727-42.
- 6 Jatoi I. Breast cancer screening. *Am J Surg* 1999;177:518-24.
- 7 Seigneurin A, Exbrayat C, Labarere J, Delafosse P, Poncet F, Colonna M. Association of diagnostic work-up with subsequent attendance in a breast cancer screening program for false-positive cases. *Breast Cancer Res Treat* 2010;127:221-8.
- 8 Paci E, Duffy S. Overdiagnosis and overtreatment of breast cancer: overdiagnosis and overtreatment in service screening. *Breast Cancer Res* 2005;7:266-70.
- 9 Welch HG, Black WC. Overdiagnosis in cancer. *J Natl Cancer Inst* 2010;102:605-13.
- 10 Paci E, Duffy S. Overdiagnosis and overtreatment of breast cancer: overdiagnosis and overtreatment in service screening. *Breast Cancer Res* 2005;7:266-70.
- 11 Ross MB, Buzdar AU, Hortobagyi GN, Lukeman JM. Spontaneous regression of breast carcinoma: follow-up report and literature review. *J Surg Oncol* 1982;19:22-4.
- 12 Dussan C, Zubor P, Fernandez M, Yabar A, Szunyogh N, Visnovsky J. Spontaneous regression of a breast carcinoma: a case report. *Gynecol Obstet Invest* 2008;65:206-11.
- 13 Serrano M. Cancer regression by senescence. *N Engl J Med* 2007;356:1996-7.
- 14 Mooi WJ, Peepker DS. Oncogene-induced cell senescence—halting on the road to cancer. *N Engl J Med* 2006;355:1037-46.
- 15 Jorgensen KJ, Gotzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. *BMJ* 2009;339:b2587.
- 16 Biesheuvel C, Barratt A, Howard K, Houssami N, Irwig L. Effects of study methods and biases on estimates of invasive breast cancer overdiagnosis with mammography screening: a systematic review. *Lancet Oncol* 2007;8:1129-38.
- 17 Jorgensen KJ, Gotzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. *BMJ* 2009;339:b2587.
- 18 Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. *N Engl J Med* 2005;353:1784-92.
- 19 Mandelblatt J, Saha S, Teutsch S, Hoerger T, Siu AL, Atkins D, et al. The cost-effectiveness of screening mammography beyond age 65 years: a systematic review for the US Preventive Services Task Force. *Ann Intern Med* 2003;139:835-42.
- 20 Berry DA, Inoue L, Shen Y, Venier J, Cohen D, Bondy M, et al. Modeling the impact of treatment and screening on US breast cancer mortality: a Bayesian approach. *J Natl Cancer Inst Monogr* 2006;36:30-6.
- 21 Fryback DG, Stout NK, Rosenberg MA, Trentham-Dietz A, Kuruchittham V, Remington PL. The Wisconsin breast cancer epidemiology simulation model. *J Natl Cancer Inst Monogr* 2006;36:37-47.
- 22 Csillery K, Blum MG, Gagliotti OE, Francois O. Approximate Bayesian computation (ABC) in practice. *Trends Ecol Evol* 2010;25:410-8.
- 23 Beaumont MA, Zhang W, Balding DJ. Approximate Bayesian computation in population genetics. *Genetics* 2002;162:2025-35.
- 24 Blum MG, Francois O. Non-linear regression models for approximate Bayesian computation. *Stat Comput* 2010;20:63-73.
- 25 Bender R, Augustin T, Blettner M. Generating survival times to simulate Cox proportional hazards models. *Stat Med* 2005;24:1713-23.
- 26 Belot A, Grosclaude P, Bossard N, Jouglu E, Benhamou E, Delafosse P, et al. Cancer incidence and mortality in France over the period 1980-2005. *Rev Epidemiol Sante Publique* 2008;56:159-75.
- 27 Erbas B, Provenzano E, Armes J, Gertig D. The natural history of ductal carcinoma in situ of the breast: a review. *Breast Cancer Res Treat* 2006;97:135-44.
- 28 Tan SY, van Oortmarssen GJ, de Koning HJ, Boer R, Habbema JD. The MISCAN-Fadia continuous tumor growth model for breast cancer. *J Natl Cancer Inst Monogr* 2006;56:65.
- 29 Silcocks PB, Robinson D. Simulation modelling to validate the flow method for estimating completeness of case ascertainment by cancer registries. *J Public Health (Oxf)* 2007;29:455-62.
- 30 Chen HH, Duffy SW. A Markov Chain Method to estimate the tumour progression rate from preclinical to clinical phase, sensitivity and positive predictive value for mammography in breast cancer screening. *The Statistician* 1996;45:307-17.
- 31 Paci E, Duffy SW. Modelling the analysis of breast cancer screening programmes: sensitivity, lead time and predictive value in the Florence District Programme (1975-1986). *Int J Epidemiol* 1991;20:852-8.
- 32 Yen MF, Tabar L, Vitak B, Smith RA, Chen HH, Duffy SW. Quantifying the potential problem of overdiagnosis of ductal carcinoma in situ in breast cancer screening. *Eur J Cancer* 2003;39:1746-54.
- 33 Olsen AH, Agbaje OF, Myles JP, Lynge E, Duffy SW. Overdiagnosis, sojourn time, and sensitivity in the Copenhagen mammography screening program. *Breast J* 2006;12:338-42.
- 34 Seigneurin A, Exbrayat C, Labarere J, Colonna M. Comparison of interval breast cancer rates for two-versus single-view screening mammography: a population-based study. *Breast* 2009;18:284-8.
- 35 Wait S, Schaffer P, Seradour B, Chollet M, Demay M, Dejouhanet S. Le dépistage spontané du cancer du sein en France. *Bull Cancer* 1997;84:619-24.
- 36 Fagnani F, Le Galès C, Lefauve C. Analyse économique du dépistage du cancer du sein par mammographie: comparaison des différentes organisations. *J d'Econ Méd* 1989;7:319-31.
- 37 Bland JM, Altman DG. Bayesians and frequentists. *BMJ* 1998;317:1151-60.
- 38 Csillery K, Francois O, Blum MGB (2011) abc: an R package for approximate Bayesian computation (ABC). *Methods Ecol Evol* (in press).
- 39 International Agency for Research on Cancer. Cancer incidence and mortality worldwide in 2008. 2011. <http://globocan.iarc.fr/>.
- 40 Institut de Veille Sanitaire. Incidence observée dans les départements couverts par les registres Francim entre 1978-1982 et 2003-2006. 2011. www.invs.sante.fr/applications/cancers/francim2010/default.htm.
- 41 Seigneurin A, Exbrayat C, Labarere J, Delafosse P, Poncet F, Colonna M. Association of diagnostic work-up with subsequent attendance in a breast cancer screening program for false-positive cases. *Breast Cancer Res Treat* 2011;127:221-8.
- 42 De Gelder R, Heijnsdijk EA, van Ravesteijn NT, Fracheboud J, Draisma G, de Koning HJ. Interpreting overdiagnosis estimates in population-based mammography screening. *Epidemiol Rev* 2011;33:111-21.
- 43 Zackrisson S, Andersson I, Janzon L, Manjer J, Garne JP. Rate of over-diagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. *BMJ* 2006;332:689-92.
- 44 Moss S. Overdiagnosis and overtreatment of breast cancer: overdiagnosis in randomised controlled trials of breast cancer screening. *Breast Cancer Res* 2005;7:230-4.

What is already known on this topic

We do not know the extent of overdiagnosis related to breast cancer screening by mammography
Published estimates vary widely and might be flawed because of inadequate allowance for biases

What this study adds

In a French population offered organised and individual screening, overdiagnosis for invasive cancers was smaller than expected in a stochastic simulation model designed to replicate standardised incidence rates of breast cancer

- 45 Puliti D, Zappa M, Miccinesi G, Falini P, Crocetti E, Paci E. An estimate of overdiagnosis 15 years after the start of mammographic screening in Florence. *Eur J Cancer* 2009;45:3166-71.
- 46 Zahl PH, Maehlen J, Welch HG. The natural history of invasive breast cancers detected by screening mammography. *Arch Intern Med* 2008;168:2311-6.
- 47 Duffy SW, Agbaje O, Tabar L, Vitak B, Bjurstam N, Bjorneld L, et al. Overdiagnosis and overtreatment of breast cancer: estimates of overdiagnosis from two trials of mammographic screening for breast cancer. *Breast Cancer Res* 2005;7:258-65.
- 48 Zahl PH, Strand BH, Maehlen J. Incidence of breast cancer in Norway and Sweden during introduction of nationwide screening: prospective cohort study. *BMJ* 2004;328:921-4.
- 49 Paci E, Miccinesi G, Puliti D, Baldazzi P, De Lisi V, Falcini F, 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.

- 50 Cuzick J, Sestak I, Pinder SE, Ellis IO, Forsyth S, Bundred NJ, et al. Effect of tamoxifen and radiotherapy in women with locally excised ductal carcinoma in situ: long-term results from the UK/ANZ DCIS trial. *Lancet Oncol* 2011;12:21-9.

Accepted: 26 September 2011

Cite this as: *BMJ* 2011;343:d7017

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited, the use is non commercial and is otherwise in compliance with the license. See: <http://creativecommons.org/licenses/by-nc/2.0/> and <http://creativecommons.org/licenses/by-nc/2.0/legalcode>.

Tables

Table 1 | Prior distributions, posterior estimates, and posterior predictive distributions of overdiagnosis for model parameters in base case analysis model

Parameter	Prior distribution*		Mean (95% CI) estimated values†
	Type	Mean	
Lifetime probability of breast cancer for women born in 1900	Uniform (6.0%, 12.0%)	9.0%	7.0% (6.0% to 8.0%)
Increase in lifetime probability for women born 1900-50	Uniform (0.5, 2.0)	1.25	1.06 (0.50 to 1.69)
Proportion of increase in lifetime probability from carcinomas in situ	Uniform (0.0%, 100.0%)	50.0%	41.0% (5.0% to 85.0%)
Age at onset of preclinical cancer: parameter 1 for scaled β distribution	Uniform (7.0, 8.0)	7.50	7.48 (6.93 to 7.99)
Age at onset of preclinical cancer: parameter 2 for scaled β distribution	Uniform (5.5, 7.0)	6.25	5.94 (5.79 to 6.12)
Proportion of invasive cancer at onset of preclinical phase	β (3, 7)	30.0%	52.0% (23.0% to 74.0%)
Proportion of in situ cancer evolving to invasive during preclinical phase	β (8, 2)	80.0%	91.0% (84.0% to 97.0%)
Proportion of non-progressive carcinoma in situ	Uniform (0.0%, 50.0%)	25.0%	6.0% (0.0% to 17.0%)
Proportion of non-progressive invasive cancer	Uniform (0.0%, 50.0%)	25.0%	3.0% (3.0% to 4.0%)
Probability of undergoing screening mammography in 1987-8 and 1989-90	Uniform (15.0%, 25.0%)	20.0%	20.0% (19.0% to 21.0%)
Probability of undergoing screening mammography in 1991-2	Uniform (19.0%, 30.0%)	25.0%	21.0% (18.0% to 24.0%)
Probability of undergoing screening mammography in 2005-6	Uniform (46.0%, 74.0%)	60.0%	57.0% (46.0% to 71.0%)
Overdiagnosis for in situ cancers			
Among cancer diagnosed	—	—	28.0% (2.2% to 59.8%)
Among cancer screened	—	—	31.9% (2.9% to 62.3%)
Overdiagnosis for invasive cancers			
Among cancer diagnosed	—	—	1.5% (0.3% to 2.9%)
Among cancer screened	—	—	3.3% (0.7% to 6.5%)

CI=credibility interval.

*Proportion of overdiagnosis for carcinoma in situ and invasive cancer was not parameter of model and consequently did not have prior distributions. Values derived from proportion of non-progressive disease for each type of cancer.

†Obtained from posterior distributions for 12 parameters and from posterior predictive distributions for proportion of overdiagnosis.

Table 2| Posterior estimates of proportion of overdiagnosis among women aged 50-69 in Isère, 1991-2006, for models 1 to 7

Model	Distribution	Sojourn time				Mean (95% CI) overdiagnosis for carcinomas in situ (%)		Mean (95% CI) overdiagnosis for invasive cancers (%)	
		Mean sojourn time (years)							
		In situ to invasive							
		In situ	Phase in situ	Phase invasive	Invasive	Among all cancers diagnosed*	Among cancers screened†	Among all cancers diagnosed*	Among cancers screened†
Model 1	γ	4.00	3.00	1.00	4.00	28.0 (3.9 to 58.1)	30.3 (4.5 to 60.1)	0.0 (0.0 to 0.1)	0.1 (0.0 to 0.2)
Model 2	γ	4.00	2.00	2.00	4.00	51.7 (13.2 to 80.4)	54.7 (15.4 to 81.8)	0.2 (0.0 to 0.5)	0.4 (0.1 to 1.0)
Model 3	γ	2.00	0.50	1.50	2.00	12.4 (5.3 to 19.7)	17.3 (7.8 to 28.5)	0.9 (0.0 to 2.9)	3.4 (0.0 to 11.4)
Model 4	γ	5.00	1.25	3.75	5.00	39.6 (10.5 to 61.4)	43.5 (13.9 to 63.3)	3.0 (0.3 to 7.6)	5.8 (0.7 to 14.5)
Model 5	Exponential	4.00	3.00	1.00	4.00	33.8 (7.4 to 65.4)	35.4 (8.0 to 65.9)	0.0 (0.0 to 0.0)	0.0 (0.0 to 0.1)
Model 6	γ	2.00	0.50	3.00	4.00	23.3 (0.6 to 65.9)	30.9 (1.0 to 72.9)	3.9 (0.2 to 10.6)	8.9 (0.5 to 24.0)
Model 7	γ	6.00	1.50	3.00	4.00	45.8 (3.6 to 85.7)	49.0 (4.4 to 86.9)	0.3 (0.0 to 1.4)	0.6 (0.0 to 3.2)

CI=credibility interval.

*Cancers diagnosed either clinically or by screening mammography.

†Cancers detected by screening mammography only.

Figures

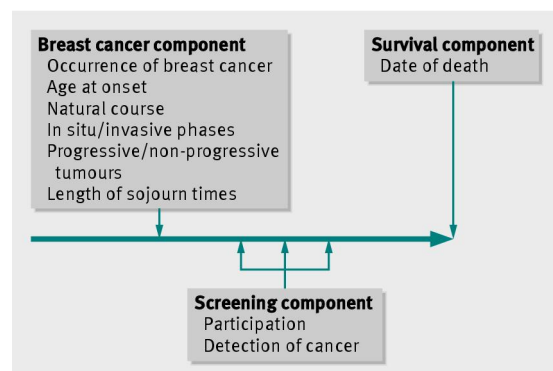


Fig 1 Simulation model for each woman

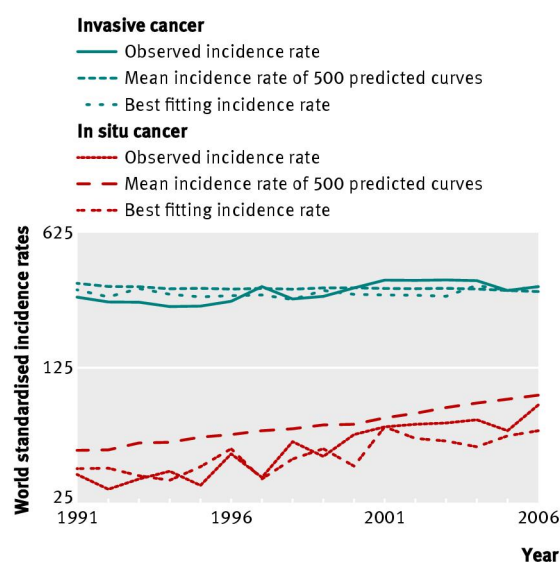


Fig 2 Mean predicted and observed standardised incidence rates of breast cancer among women aged 50-69, Isère, 1991-2006

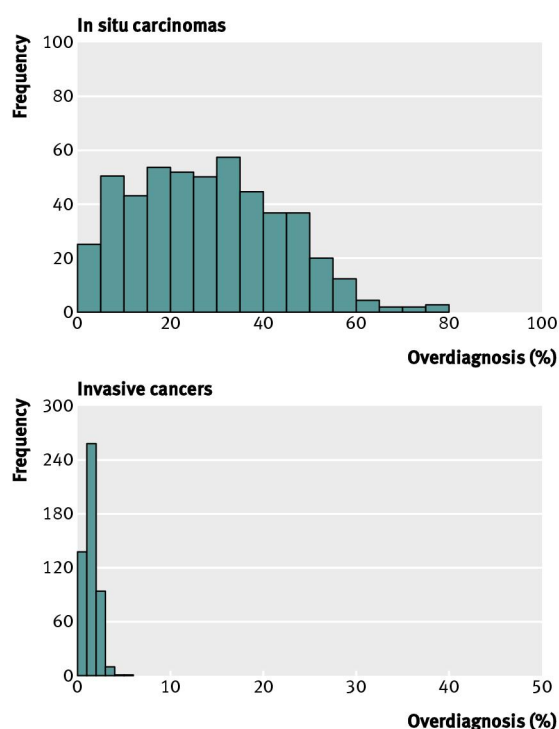


Fig 3 Posterior predictive distribution of overdiagnosis (%) among all cases of breast cancers detected either clinically or by screening mammography in women aged 50-69, Isère, 1991-2006

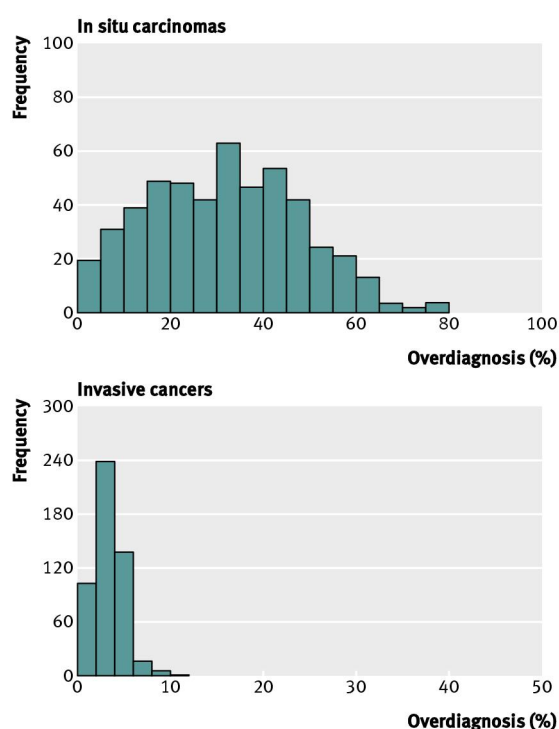


Fig 4 Posterior predictive distribution of overdiagnosis (%) among all cases of breast cancers detected by screening mammography in women aged 50-69, Isère, 1991-2006