STATISTICS IN MEDICINE
Statist. Med. 2008; **27**:5235–5251
Published online 9 July 2008 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sim.3345

Predicting impacts of mass-screening policy changes on breast cancer mortality

Johanna Seppänen[1,*,†], Sirpa Heinävaara[1,2] and Timo Hakulinen[1]

[1]*Finnish Cancer Registry, Helsinki, Finland*
[2]*Radiation and Nuclear Safety Authority, Helsinki, Finland*

SUMMARY

The aim of this study is to present a methodology for taking into account the mass-screening invitation data in breast cancer mortality predictions, particularly in assessing impacts of screening policy changes on the short-term predictions. The methodology is applied to a database that includes observed year- and age-specific screening invitation schemes in Finnish municipalities from the time period 1987–2001. The target year for predictions is 2012.

To predict mortality, breast cancer incidence and patients' survival from breast cancer are modelled with the screening data included. The knowledge of breast cancer survival together with the other cause survival is then used to calculate the number of breast cancer deaths caused by observed (1987–2001) and predicted (2002–2012) incident cases in Finland.

Survival from breast cancer was estimated with a parametric mixture model where the patient population is assumed to be a combination of cured and uncured patients. This approach provides a way of modelling the hazard of fatal cases and the proportion of cured cases simultaneously. In other cause survival, the patients' hazard was allowed to differ from that of the general population.

Breast cancer mortality predictions are presented according to three alternative future scenarios of screening policy. The results show no major differences between predictions yielded by alternative scenarios: Any policy change would have at the most a 3.0 per cent impact on breast cancer mortality in the near future. Copyright © 2008 John Wiley & Sons, Ltd.

KEY WORDS: mortality; survival; breast cancer; mass screening; prediction; cure model

*Correspondence to: Johanna Seppänen, Finnish Cancer Registry, Institute for Statistical and Epidemiological Cancer Research, Liisankatu 21 B, FIN-00170 Helsinki, Finland.
†E-mail: johanna.seppanen@cancer.fi

Contract/grant sponsor: Cancer Society of Finland
Contract/grant sponsor: Academy of Finland (MaDaMe project)
Contract/grant sponsor: Doctoral Programs in Public Health/Department of Public Health of the University of Helsinki

Received 30 May 2007
Accepted 29 April 2008

INTRODUCTION

The most important indicator for the effectiveness of a mass-screening programme is the cause-specific mortality. Screening activities without a decreasing impact on mortality are poorly justifiable. The Finnish invitational screening programme for breast cancer that has been going on since 1987 is reported to be effective [1–4]. Recently, the screening service was decided to be extended from age group 50–59 to 50–69. However, the scope of the programme is still frequently discussed in public, and information about potential impacts of mass-screening practice changes on future mortality is required. This paper is motivated by the lack of proper methodology for such a policy change assessment.

To facilitate alternative future predictions, it is useful to take the programme into pieces which can then be combined to correspond to the desired screening policy. For that we need detailed screening data from a sufficiently long time period. A database that includes actual observed year- and age-specific screening invitation schemes from Finnish municipalities from the time period 1987–2001 has previously become available at the Finnish Cancer Registry. This has made it possible to develop a methodology for taking into account historical municipality-specific schemes of mass screening when modelling and constructing predictions for breast cancer incidence [5, 6]. In this study we want to extend this methodology to breast cancer survival and mortality.

To predict mortality, we need to know the survival in addition to the predicted incidence. To be able to die from breast cancer at a certain time point t, a patient must first stay alive until that. That is, not to die either from breast cancer or other causes before time t. For this reason we need to know the estimates for both survival from breast cancer and other causes. This cause-specific approach requires reliable cancer registry data on the causes of death. Following Heinävaara and Hakulinen [7] and De Angelis *et al.* [8], the cause-specific survival is modelled by applying a parametric mixture model where the patient population is allowed to be a mixture of two subpopulations with distinct risk of dying: those patients who are cured and thus not at risk of dying from breast cancer, and those who are bound to die of the disease. This approach provides a way of modelling the hazard of fatal cases and the proportion of cured cases simultaneously. The role of prognostic factors such as age and state of screening programme at diagnosis can be evaluated separately for these two patient groups. In addition, the proportion of cured patients can also be allowed to vary by covariates. This is the first time a cure model is fitted to Finnish breast cancer data and separate cure fractions are identified for invited and non-invited women in different age categories.

It would be straightforward to assume that breast cancer patients are a random sample and their risk to die from other causes than breast cancer is the same as in the corresponding general population. However, results from stomach [9] and lung cancer [7] suggest that the risk could be significantly elevated and may also depend on covariates. Phillips *et al.* [9] proposed a model where the hazard due to other causes is not assumed to be the same but proportional to that of the general population. Furthermore, our model following Heinävaara and Hakulinen [7] allows the other cause hazard to vary by the stage of cancer, time since diagnosis, and age.

The aim of this study is to present a methodology for taking into account the mass-screening invitation information in breast cancer mortality predictions. The focus is not on assessing the absolute efficacy of mass screening on breast cancer mortality, but, by utilizing the data on invitations, on showing the estimated impacts of programme extensions on the short-term predictions. This involves modelling and predicting incidence, survival from breast cancer and survival from other causes. Making use of the earlier results presented for incidence [5], this study mainly concentrates on estimating the survival and predicting the mortality. Based on different future scenarios of

DOI: 10.1002/sim

mass-screening practice, alternative incidence-based mortality predictions for year 2012 will be calculated. In connection with survival estimation, results for cure fraction and life expectancy (LE) for the uncured will be presented. Additionally, patient's risk to die from other causes than breast cancer compared with the corresponding general population will be quantified. Since we know that screening has a different impact on the incidence of localized than non-localized cases [5, 6], all analyses will be done separately for localized and non-localized breast cancers.

DATABASE

During 1987–2001, there was a bylaw on public health in Finland regulating the invitation procedure by stating that organized, free of charge mammography screening should be offered every second year to women aged 50–59 years, whereas screening of other age groups remained optional. Information on annual screening invitation schemes in 267 (60.1 per cent) Finnish municipalities during 1987–2001 were assembled from the files of the Mass Screening Registry, a part of the Finnish Cancer Registry [5, 6]. A total of 67 (25.1 per cent) municipalities in our database offered voluntarily screening to women aged 40–49 and 19 (7.1 per cent) to women aged 70–74 [5]. The screening invitation data were linked with cancer data including all female breast cancer cases from the Finnish Cancer Registry database diagnosed in those 267 municipalities during 1987–2001. All patients were followed up for death until 2002. Observed (1987–2002) and predicted (2003–2012) population counts together with the age-specific and calendar-year-specific expected survival probabilities for the Finnish general population were received from Statistics Finland. The analysis was restricted to the age group of women between 40 and 74 years because that was the age range where screening invitation data were available. The invitation information can be assumed to reflect the actual screening activity very well since the mean compliance among 50–64 year old women was 90 per cent at the first screen and 93 per cent at subsequent screens during the 1990s in Finland [10]. According to stage information, the cases were classified to localized (54.9 per cent of all the cases/25.0 per cent of deaths), non-localized (36.3/70.3 per cent), and cases with unreported stage (8.7/4.8 per cent). Ductal carcinoma *in situ* tumours (3.8 per cent), and cases obtained only from death certificates or autopsy reports (0.1 per cent) were excluded from the data set. Women with multiple breast cancers (2.0 per cent) were included only once according to their first breast cancer.

Components of the screening programme

The Finnish national mass-screening programme involves a 2-year screening interval; hence, one screening round lasts for 2 years. For incidence modelling and predictions, the screening programme was divided into seven components [5]. Every cohort was coded according to the municipality's invitation scheme. All women living in the same municipality and born in the same year had the same invitation pattern.

The first year of the first screening round is the year of the first invitation (component 1). Separation of the first and second years of the screening rounds (components 1 and 2, respectively) was motivated by the idea that, assuming screening is effective, incidence will not be constant during the round; it will increase during the first year when actual screening takes place, and decrease below the baseline during the second year. The first round (components 1 and 2) was separated from the subsequent ones (components 3 and 4). The post-screening period was divided

DOI: 10.1002/sim

into two components: the first 5 years after the programme (component 5) and the rest of the time until the age of 74 or the end of follow-up (component 6). Component 7 was defined as a break in a cohort's screening programme. These are irregular years in between the programme; the main point is that the cohort has been screened at least once and will be screened again in the future. Component 0, the not invited, includes all the women up to the point of first invitation. This component served as a baseline. The coding of the components for an example cohort born in 1939 and screened for the first time at the age of 50 is given in Table I.

When estimating the breast cancer survival some of these components were combined to provide stable estimates for each subgroup: components 1 and 3 were combined to A, as well as 2, 4, and 5 to B, to end up with four components of the screening programme (component $7=C$ and component $6=D$) plus the non-invited ($=N$). The breast cancer deaths occurring during the follow-up were connected with the screening programme component that took place during the year of diagnosis, see Table I. Table II shows the numbers of incident breast cancer cases and subsequent breast cancer deaths observed during each component of the screening programme by age group at diagnosis.

METHODS

Screening scenarios for future

For future incidence predictions up to 2012, alternative scenarios of mass-screening practices for breast cancer were programmed by allowing the value of screening programme component indicator in a given municipality, diagnosis year, and age to alter from scenario to another [5]. In all scenarios, it was assumed that the proportion of cases with unreported stage remains stable and that municipalities will strictly follow the guidelines, that is, there will be no deviations or breaks in the programme. The scenarios were (I) continuing the current practice of inviting 50–59 year old women every second year, (II) extending screening service from 50–59 to 50–69 year old, and (III) extending screening service from 50–59 to 40–65 year old women.

Modelling and predicting the breast cancer incidence

Modelling of breast cancer incidence during the observation period 1987–2001 was performed using the general age–period approach presented by Clayton and Schifflers [11]. The analysis was carried out using the Poisson regression with a logarithmic link function, where the expected incidence EI_{mya} in municipality m ($m=1,\ldots,267$) in calendar year y ($y=1987,\ldots,2001$) for women of age a ($a=40,\ldots,74$) is expressed as

$$\log(\mathrm{EI}_{mya})=\alpha^{(I)}+\beta^{(I)}y+\gamma_1^{(I)}a+\gamma_2^{(I)}a^2+\gamma_3^{(I)}a^3+\gamma_4^{(I)}a^4+\rho_{r(m)}^{(I)}+\delta_{s(m,y,a),c(a)}^{(I)} \quad (1)$$

where $r(m)$ is the university hospital region r ($r=1,\ldots,5$) to which municipality m belongs, $s(m,y,a)$ is the screening invitation component s ($s=0,\ldots,7$) in municipality m in calendar year y for women of age a, and $c(a)$ is the age category c ($c=$40–49, 50–59, 60–74) that includes age a. The term $\delta_{s(m,y,a),c(a)}^{(I)}$ is an interaction between screening invitation status and categorical age. Calendar year and age were treated as numerical variables. The superscript I stands for incidence.

Maximum likelihood estimates together with different scenarios of future screening policy were then used in extrapolating the model into the future and calculating alternative predictions for

DOI: 10.1002/sim

Table I. The connection between screening programme component at diagnosis and subsequent model-based breast cancer deaths occurring during the follow-up in an example cohort born in 1939.

Diagnosis year	Diagnosis age	Screening component	Calendar year of follow-up																									
			1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
2012	73	6																										*D*
2011	72	6																									*D*	*D*
2010	71	5																								*B*	*B*	*B*
2009	70	5																							*B*	*B*	*B*	*B*
2008	69	5																						*B*	*B*	*B*	*B*	*B*
2007	68	5																					*B*	*B*	*B*	*B*	*B*	*B*
2006	67	5																				*B*	*B*	*B*	*B*	*B*	*B*	*B*
2005	66	4																			*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*
2004	65	3																		*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*
2003	64	4																	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*
2002	63	3																*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*
2001	62	4															*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*
2000	61	3														*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*
1999	60	7													*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*
1998	59	7												*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*
1997	58	7											*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*
1996	57	4										*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*
1995	56	3									*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*
1994	55	4								*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*
1993	54	3							*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*
1992	53	4						*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*
1991	52	3					*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*
1990	51	2				*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*
1989	50	1			*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*
1988	49	0		*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*
1987	48	0	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*

This cohort had a 3-year break in screening during 1997–1999, and the service was extended up to 65 years (scenario III). The components at diagnosis are as follows: 0, not invited; 1, first year of the first screening round; 2, second year of the first screening round; 3, first years of subsequent screening rounds; 4, second years of subsequent screening rounds; 5, up to 5 years after the last screening round; 6, more than 5 years after the last round; 7, break. The uppercase letters illustrate the corresponding combined components used in survival estimation.

DOI: 10.1002/sim

Table II. Number of breast cancer cases diagnosed during each component (0–7) of the screening programme (time period 1987–2001, 267 municipalities) and the number of subsequent breast cancer deaths in period 1987–2002.

		Localized		Non-localized		Unreported	
Component	Age	Cases	Deaths	Cases	Deaths	Cases	Deaths
(1) First round/first year	40–49	63	3	31	10	6	2
	50–59	948	71	520	156	141	6
	60–74	71	7	22	7	13	1
(2) First round/second year	40–49	15	2	14	5	0	0
	50–59	251	28	167	60	37	2
	60–74	20	1	14	8	4	0
(3) Subsequent rounds/first years	40–49	121	12	48	11	13	2
	50–59	1876	84	928	199	318	18
	60–74	524	28	190	55	71	4
(4) Subsequent rounds/second years	40–49	15	0	7	1	1	0
	50–59	510	18	382	102	96	9
	60–74	151	12	101	50	36	3
(5) Up to 5 years after	40–49	—	—	—	—	—	—
	50–59	28	3	12	1	4	1
	60–74	847	35	530	151	177	13
(6) More than 5 years after	40–49	—	—	—	—	—	—
	50–59	—	—	—	—	—	—
	60–74	259	5	183	26	88	2
(7) Break	40–49	32	2	32	10	3	0
	50–59	101	7	41	8	7	0
	60–74	48	1	23	2	7	1
(0) Not invited	40–49	2068	176	1725	548	300	38
	50–59	333	47	244	126	36	7
	60–74	2461	304	1893	845	352	52
Total		10742	846	7107	2381	1710	161

breast cancer incidence rates up to 2012. In predictions, the exponential relationship between incidence and calendar time was replaced with a linear one. Further details about model selection are presented in Seppänen *et al.* [5].

Survival from breast cancer

Let t be a positive random variable denoting the survival time from breast cancer for a patient living in municipality m ($m=1,\ldots,267$), diagnosed with breast cancer in calendar year y ($y=1987,\ldots,2001$) at age a ($a=40,\ldots,74$). Following Heinävaara and Hakulinen [7] and De Angelis *et al.* [8] we used a parametric mixture model to estimate the cause-specific survival. In our model we assumed that a proportion P of patients is statistically cured and hence has a survival from breast cancer equal to 1; the remaining proportion $1-P$ represents those uncured. The proportion P, as well as the shape and scale parameters of the underlying distribution, is expressed as functions of covariates.

For the cure fraction P and for the shape k and scale λ parameters (see below), we classified the screening invitation status into a new 3-class variable $d=d(m,y,a)$ ($d=$ first years (combined component A), all the other components of the screening programme (B, C, and D), not invited

DOI: 10.1002/sim

(N)) to ensure the convergence of the model and stable estimates. In addition to d, the proportion of cured P is allowed to depend on age category $c=c(a)$, that is

$$P_{cd}=1/[1+\exp(\alpha^{(P)}+\sigma_c^{(P)}+\delta_d^{(P)})] \tag{2}$$

The mixture model for the overall breast cancer survival can then be expressed as

$$S_{mya}^{(C)}(t)=P_{cd}+(1-P_{cd})S_{mya}^{(C,1-P)}(t) \tag{3}$$

where $S_{mya}^{(C,1-P)}$is the breast cancer survival function for the uncured population.

We express the likelihood for the data contributed by each patient i, $i=1,\ldots,N$

$$L=\prod_{i=1}^{N}(P_i+(1-P_i)S_i^{(C,1-P)}(t_i))^{1-\upsilon_i}((1-P_i)S_i^{(C,1-P)}(t_i)h_i^{(C,1-P)}(t_i))^{\upsilon_i} \tag{4}$$

where $P_i=P_{c(a_i),d(m_i,y_i,a_i)}$ according to (2), $S_i^{(C,1-P)}(t)=S_{m_i y_i a_i}^{(C,1-P)}(t)$, and the hazard of dying at time point t for the uncured $h_i^{(C,1-P)}(t)=\log(S_i^{(C,1-P)}(t-1)/S_i^{(C,1-P)}(t))$. It is assumed that only uncured patients can die from breast cancer. The patient's i status when she exits the follow-up is denoted with υ_i (υ_i=1 if died from breast cancer and υ_i=0 otherwise).

The log of the likelihood was maximized using the iterative Gauss–Newton method. Computing was performed using the NLIN procedure in SAS Release 9 [12].

To clarify the notations, the subscript i is suppressed from now on. The expected survival times for the uncured patients were modelled using an accelerated failure time model

$$\log(E(t))=\alpha^{(S)}+\beta^{(S)}y+\gamma^{(S)}a+\sigma_{c(a)}^{(S)}+\delta_{q(m,y,a)}^{(S)} \tag{5}$$

where $q(m,y,a)$ is the combined screening invitation programme component q ($q=A,B,C,D,N$) in municipality m in calendar year y for women of age a, and $c(a)$ is age category as in (1). Calendar year and age are treated as numerical variables.

Exponential, Weibull, lognormal and two-parameter gamma distributions were considered as potential candidates to be used to model the survival time. Discrimination between different probability distributions and selection of the covariates were based on both likelihood ratio tests and graphical diagnostics including examination of (log)negative–log plots and Cox–Snell residuals. The general principle was to make the model as simple and robust as possible. To provide reliable extrapolations, the primary criterion was the fit of model-based number of deaths curve to the observed one, especially towards the end of the observation period [13]. As a result, the two-parameter gamma distribution was chosen both on grounds of the best fit in both stage categories, and the flexibility it has in the shape of the function compared with other candidates (that are all nested within the gamma). We also wanted to allow the shape and scale of the underlying distribution to vary with subgroups defined by covariates. Furthermore, according to Yu *et al.* [14], the cure fraction estimates from the model with gamma distribution are found to be quite robust. Figure 1 shows the fit of the survival model by displaying model-based vs observed numbers of breast cancer deaths. We have also looked at fit curves by other covariates, i.e. screening component and continuous age, and none of them revealed any lack of fit.

DOI: 10.1002/sim

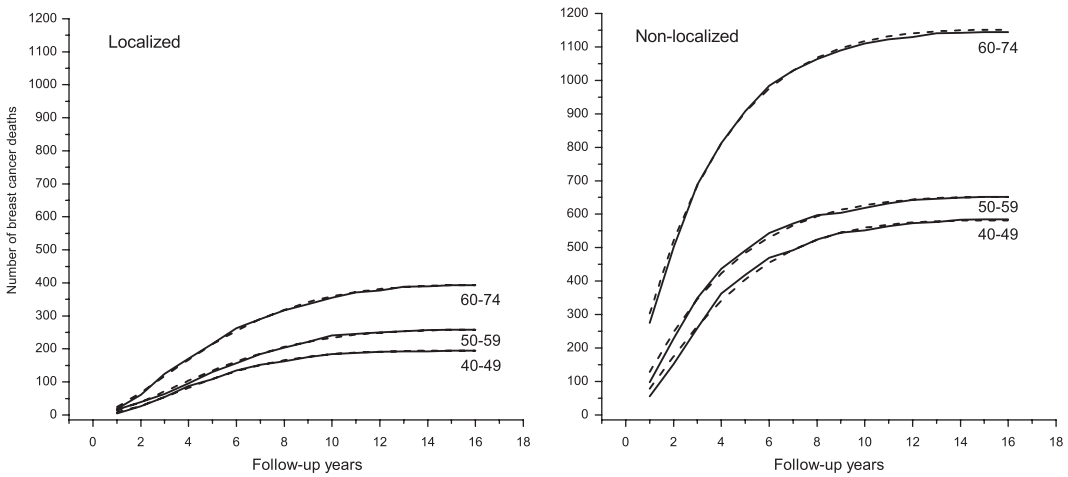

Figure 1. Model-based (dashed line) and observed (solid line) numbers of breast cancer deaths plotted by age group.

The survival functions for the uncured patient group $S_{mya}^{(C,1-P)}$ are different for each m, y, and a as expressed in (5). Their general form is

$$S^{(C,1-P)}(t)=1-F(t)=1-\left[\frac{1}{\lambda^k\Gamma(k)}\int_0^t x^{(k-1)}\exp(-x/\lambda)\,\mathrm{d}x\right] \tag{6}$$

where Γ is the gamma function and $F(t)$ is the cumulative distribution function for the two-parameter gamma distribution that cannot be presented in a closed form. The shape parameter $k>0$ is allowed to depend on age category $c=c(a)$ and screening invitation status $d=d(m,y,a)$ as follows:

$$k=k_{cd}=\exp(\alpha^{(K)}+\sigma_c^{(K)}+\delta_d^{(K)}) \tag{7}$$

and the scale parameter $\lambda>0$ is defined as

$$\lambda=E(t)/k \tag{8}$$

where $E(t)$ is the expected survival time for the uncured patients according to (5).

Variation in breast cancer survival

To estimate the uncertainty related with breast cancer survival, and especially with its components LE for uncured (5) and proportion of cured patients (2), we constructed a likelihood-based 95 per cent confidence region

$$2(L_{\max}-L)\leqslant\chi^2_{(\mathrm{df}=2,\alpha=0.05)} \tag{9}$$

where $L_{\max}$ is the maximum of the log likelihood as in Equation (4) and L is a log likelihood obtained by altering the values of expected survival time and proportion of cured. When calculating L, the shape and scale parameters were not changed but held fixed as their maximum likelihood estimates.

DOI: 10.1002/sim

After constructing the confidence region, different points were systematically chosen from the border of the region, and mortality predictions were re-calculated using corresponding values of LE and P. The 95 per cent confidence bounds for breast cancer mortality predictions related with uncertainty in breast cancer survival estimates were defined as the annual minimum and maximum of these predictions.

Survival from other causes than breast cancer

The cancer patients' other cause survival is not necessarily the same as in the corresponding general population [7, 9]. The relationship between patients' and general population's hazards might also not be constant but depend on covariates such as stage of the tumour, age, and time passed from diagnosis. To allow the patients' hazard to die from other causes than breast cancer $h^{(O)}$ to differ from that of the general population h^*, we used the ratio between patients' and general population's hazards $r = h^{(O)}/h^*$ as a correction coefficient. The general (Finnish) population was matched to the patient population by calendar year and numerical age. To assess the proportionality between hazards, ratio r was modelled as a function of follow-up period stratified by stage and age group using Poisson regression with a logarithmic link function. No intercept was included in the model.

Survival from other causes than breast cancer t at follow-up time t (in years) for women diagnosed in calendar year y at age a can be expressed as

$$S_{ya}^{(O)}(t) = \exp\left[-\sum_{j=1}^{t} h^*((a-1)+j, (y-1)+j) \cdot r(c(a), j) \right] \tag{10}$$

where $h^*(u, v)$ is the hazard to die at age u in calendar year v in the general population, and $r(c, p)$ is the model-based ratio during follow-up period p ($p = 1, 2$, 3–5, 6–9, and 10 or more years after diagnosis) for women whose age at diagnosis is included in age group c ($c =$ 40–49, 50–59, 60–74).

Predicting mortality

The survival was defined on a patient level, but the breast cancer incidence predictions were grouped to correspond to the accuracy of the screening invitation data: by municipality, calendar year, and age at the end of the year of diagnosis. To get down to mortality, incidence and survival were combined. This was done by assuming that survival is the same among a group of women diagnosed with a breast cancer of a certain stage, age and year, and during the same component of screening.

Since the log likelihoods of survival from breast cancer and survival from other causes are additive, the overall survival can be expressed as a product of these two components [7]. If we denote the survival from breast cancer in a certain stage with $S^{(C)}$, survival from other causes with $S^{(O)}$, and the hazard to die from breast cancer with $h^{(C)}$ the expected number of breast cancer deaths D at follow-up time t (in months) in municipality m among women diagnosed in calendar year y in month z ($z = 1, \ldots, 12$) at age a can be calculated as

$$D(t; m, y, z, a) = \tfrac{1}{12} B(m, y, a) \cdot S_{mya}^{(C)}(t; m, y, z, a) \cdot S_{ya}^{(O)}(t; y, z, a) \cdot h_{mya}^{(C)}(t; m, y, z, a) \tag{11}$$

where $B(m, y, a)$ is the number of cancer cases of the same stage observed (years 1987–2001) or predicted to occur (2002–2012) in municipality m in calendar year y for women of age a. The predicted number of cases as well as the screening component in $S^{(C)}$ alters depending on the

DOI: 10.1002/sim

future scenario used as the basis of incidence predictions, as described earlier. The yearly number of new cases is assumed to be evenly distributed between months; hence, the number of cases in each month is $B/12$. The diagnosis was programmed to take place in the middle of the diagnosis month z. The hazard of dying from other causes than breast cancer was assumed to be constant during each calendar year.

In predictions for all cases combined, the cancer cases with an unreported stage had to be accounted for (Table II). The unreported stage group consists of localized and non-localized cases in an unknown ratio. To estimate that ratio, relative survival of the cases with unreported stage R_U was defined to be a combination of the relative survivals of localized R_L and non-localized R_N cases: $R_U = xR_L + (1-x)R_N$. As a result, x was estimated to be 0.7535, indicating that approximately 75 per cent of the unreported cases would behave as localized and 25 per cent as non-localized cancer cases in terms of relative survival. Since the proportion of unreported cases was 8.7 per cent of all cases (Table II), a total of 8.7 per cent $\times 0.75 + 54.9$ per cent $= 61.4$ per cent of all cases would be localized and 8.7 per cent $\times 0.25 + 36.3$ per cent $= 38.5$ per cent would be non-localized. By dividing the estimated proportion by the observed, we obtained coefficients 61.4/54.9 per cent $= 1.12$ for localized and 38.5/36.3 per cent $= 1.06$ for non-localized cases. The estimated number of all deaths is then $1.12 \times$ localized deaths $+ 1.06 \times$ non-localized deaths.

The annual breast cancer mortality rate predictions for target year Y can then be derived as the sum of predicted deaths over all municipalities, ages, and diagnosis years up to Y divided by the corresponding number of person years. Similarly, this method can be directly applied to obtain mortality predictions by age (at diagnosis or at death) or screening component.

RESULTS

As an example, Table III displays LEs for uncured patients diagnosed in year 1995 at the age of 45, 55, or 65 during different components of the screening programme. The amount of uncertainty associated with the estimates is relatively high leading to insignificant differences between the point estimates. In the cases diagnosed during the first years of the screening rounds the estimates indicate lower (localized) or similar (non-localized) LE compared with non-invited. These are the years when the invitation and also the mammography test in most cases take place. LEs for cases diagnosed during the programme breaks are the lowest in all subgroups.

In general, the proportion of cured patients (P) was the highest in cases diagnosed during the first years of the screening rounds, but the effects were not significant, see Table III. Estimated proportions also decreased with age and were almost double in localized cases compared with women with a more advanced disease. Standard errors for LE and P were first calculated on log and logit scales, respectively, and then back transformed.

A point of statistical cure for the cured population was defined as the point in follow-up time when $S^{(C)}$ reaches the level of P, that is, the point in time when $S^{(C)}$ approaches its asymptote. The precision used for calculating the point of cure appeared to be very crucial. When using a 1 per cent unit precision to define the equality between P and $S^{(C)}$, the time needed to reach the point of cure after the diagnosis varied from 15 to 45 years in localized, and from 20 to 46 years in non-localized cases diagnosed in year 1995, depending on the age and the screening component. Due to long and flat tails of the survival distributions, an increase in the precision with one decimal would yield results of the order of 10 years longer. As a result, the time to the point of cure for breast cancer patients was very long in relation to LEs.

DOI: 10.1002/sim

Table III. Life expectancies for uncured breast cancer patients diagnosed in year 1995 at the age of 45, 55, or 65 during combined components (*A*, *B*, *C*, *D*, and *N*) of the screening programme, and corresponding proportions of cured patients. Estimates are provided with 95 per cent confidence intervals.

Age at diagnosis	Life expectancy (years)	Proportion of cured (per cent)
Localized		
45 years		
(*A*) First years	7.7 (2.9, 20.2)	89.1 (77.2, 95.2)
(*B*) Second years and up to 5 years after	6.8 (2.3, 20.4)	87.4 (70.1, 95.3)
(*C*) Break	5.1 (1.3, 20.5)	*
(*N*) Not invited	9.2 (5.3, 16.0)	82.5 (74.6, 88.4)
55 years		
(*A*) First years	12.2 (5.4, 27.6)	85.0 (72.2, 92.5)
(*B*) Second years and up to 5 years after	10.9 (3.8, 31.1)	82.8 (62.7, 93.2)
(*C*) Break	8.2 (2.1, 31.9)	*
(*N*) Not invited	14.7 (5.0, 42.7)	76.7 (55.7, 89.6)
65 years		
(*A*) First years	11.9 (4.2, 33.7)	82.0 (61.1, 93.0)
(*B*) Second years and up to 5 years after	10.5 (3.5, 31.8)	79.5 (53.5, 92.9)
(*D*) More than 5 years after	8.9 (2.3, 34.7)	*
(*C*) Break	7.9 (1.9, 32.8)	*
(*N*) Not invited	14.3 (6.3, 32.2)	72.6 (53.9, 85.7)
Non-localized		
45 years		
(*A*) First years	8.3 (3.6, 19.3)	51.8 (31.0, 72.0)
(*B*) Second years and up to 5 years after	4.9 (2.2, 11.1)	51.0 (30.7, 70.9)
(*C*) Break	7.1 (2.5, 20.3)	*
(*N*) Not invited	8.2 (5.8, 11.7)	42.8 (32.3, 53.9)
55 years		
(*A*) First years	8.9 (5.0, 15.7)	48.1 (32.9, 63.6)
(*B*) Second years and up to 5 years after	5.2 (2.8, 9.9)	47.3 (32.0, 63.1)
(*C*) Break	7.5 (2.9, 19.4)	*
(*N*) Not invited	8.7 (4.5, 16.9)	39.2 (24.5, 56.1)
65 years		
(*A*) First years	8.3 (3.7, 18.7)	43.6 (24.2, 65.2)
(*B*) Second years and up to 5 years after	4.9 (2.4, 10.0)	42.8 (25.3, 62.4)
(*D*) More than 5 years after	9.2 (3.0, 28.1)	*
(*C*) Break	7.1 (2.6, 19.3)	*
(*N*) Not invited	8.2 (5.4, 12.4)	35.0 (25.5, 45.8)

*According to the model, the same as in 'second years and up to 5 years after'.

Results for the proportionality of other cause mortality to that of general population (r) are given in Table IV. The estimates were heterogeneous between localized and non-localized cases: In patients with a localized cancer r was significantly decreased during the first year in age groups 50–59 and 60–74, and also during the second year in women aged 60–74. In patients with a non-localized cancer, r was increased in the youngest age group from the third follow-up year onwards, and in age group 50–59 during the follow-up years 6–9.

DOI: 10.1002/sim

Table IV. Estimates for ratio (r) between breast cancer patients' hazard to die from other causes than breast cancer and the hazard to die in comparable general population.

	Follow-up time after diagnosis				
	1 year	2 years	3–5 years	6–9 years	10–15 years
Age at diagnosis	r (95 per cent CI)	r (95 per cent CI)	r (95 per cent CI)	r (95 per cent CI)	r (95 per cent CI)
Localized					
40–49	0.46 (0.12, 1.84)	0.45 (0.11, 1.78)	0.77 (0.41, 1.43)	1.00 (0.59, 1.69)	1.41 (0.83, 2.38)
50–59	**0.20 (0.06, 0.62)**	0.85 (0.49, 1.46)	0.94 (0.69, 1.28)	1.14 (0.87, 1.48)	0.88 (0.63, 1.22)
60–74	**0.63 (0.46, 0.87)**	**0.73 (0.55, 0.98)**	0.92 (0.79, 1.07)	0.99 (0.86, 1.13)	1.03 (0.89, 1.19)
Non-localized					
40–49	0.89 (0.29, 2.75)	1.80 (0.81, 4.01)	**2.31 (1.47, 3.62)**	**2.00 (1.20, 3.31)**	**2.05 (1.10, 3.81)**
50–59	1.10 (0.57, 2.12)	1.40 (0.77, 2.53)	1.17 (0.77, 1.77)	**1.50 (1.02, 2.20)**	1.16 (0.71, 1.89)
60–74	1.17 (0.88, 1.56)	1.13 (0.84, 1.52)	1.10 (0.91, 1.33)	1.03 (0.84, 1.25)	1.05 (0.83, 1.33)

The statistically significant estimates are given in bold.

DOI: 10.1002/sim

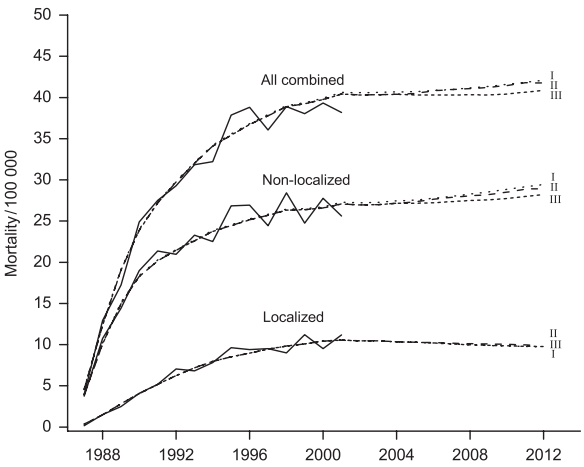

Figure 2. Observed (solid line) and predicted breast cancer mortality rates among women with a localized and a non-localized cancer. All cases combined includes also cases with unreported stage, see section 'Predicting mortality'. Predictions are based on alternative mass-screening scenarios. In scenario I the current practice of inviting 50–59 year old women every second year is continued, in scenario II the screening service is extended from 50–59 to 50–69 and in scenario III to 40–65 year old women. Changes in screening practice have been programmed to take place in year 2005.

Figure 2 displays the overall curves for both observed and predicted breast cancer mortality rates. The predicted rate for the target year 2012 according to scenario I where the current practice is continued is 9.7 breast cancer deaths/100 000 person-years in reported localized cases, 29.4/100 000 in reported non-localized cases, and 42.1/100 000 in all cases combined. There are no major differences between predictions yielded by the alternative scenarios. The short-term effect in breast cancer mortality caused by extending the service from current practice (scenario I) to 50–69 year old women (scenario II) is $+1.8$, -1.7, and -0.8 per cent in localized, non-localized, and all cases combined, respectively. Corresponding results for extending from current practice (scenario I) to 40–65 year old women (scenario III) are $+0.7$, -4.5, and -3.0 per cent, respectively.

Both 95 per cent confidence regions based on likelihood ratio in Figure 3 show negative correlation, especially in the non-localized cases where change in LE is highly negatively correlated with change in P. Therefore, even if there is substantial variation in the estimated values of LE and P (Table III), the estimated breast cancer mortality is fairly stable (Figure 4). Figure 4 shows the 95 per cent confidence bounds for breast cancer mortality predictions related with uncertainty in breast cancer survival estimates. In target year 2012, the confidence bounds for mortality caused by localized breast cancers are 8.9–10.6 cases/100 000, and for mortality due to non-localized cancers are 28.2–30.6 cases/100 000.

DISCUSSION

Detailed screening invitation data were incorporated into modelling of breast cancer incidence and survival by defining a screening variable that gives the component of the screening programme by municipality, year, and age. Cause-specific survival was then modelled by applying a parametric

DOI: 10.1002/sim

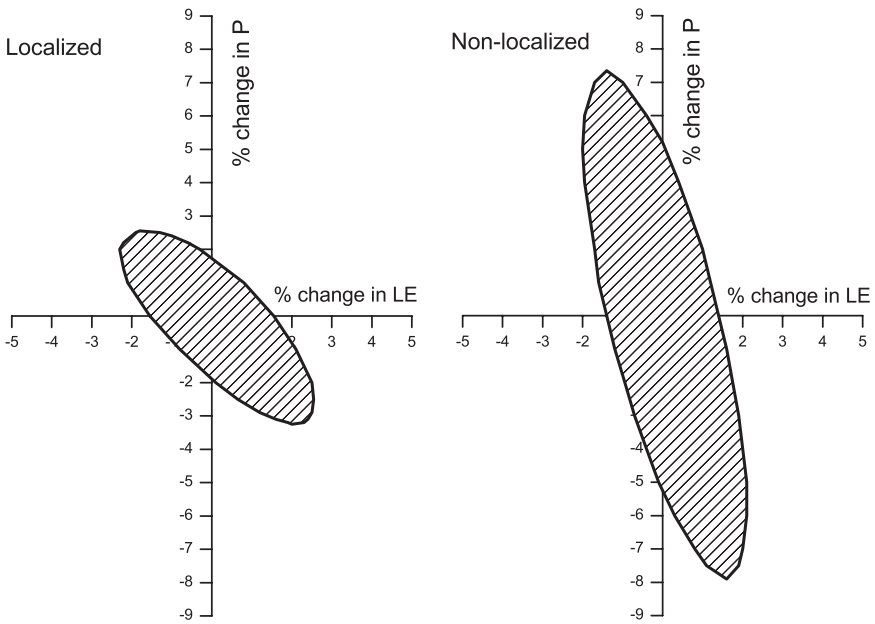

Figure 3. Likelihood-based 95 per cent confidence regions defined by $2(L_{\max} - L) \leqslant \chi^2$ (df $= 2, \alpha = 0.05$), where $L_{\max}$ is the maximum log likelihood and L is the log likelihood obtained by altering the values of expected survival time for the uncured patients (LE) and proportion of cured patients (P).

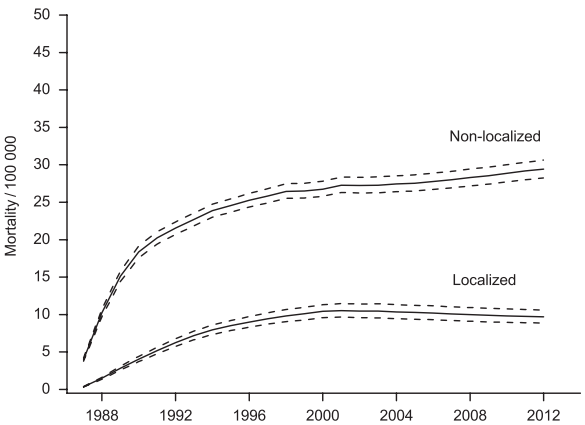

Figure 4. The breast cancer mortality predictions made under scenario I (solid line) and the 95 per cent confidence bounds related with uncertainty in breast cancer survival estimates (dashed line).

mixture cure model, where the patient population was allowed to be a combination of cured and uncured patients. In addition to survival for the uncured, effects of covariates were also incorporated through the proportion of cured patients and shape and scale parameters of the survival time distribution. The patients' risk to die from other causes than breast cancer was allowed to differ from that of a corresponding general population group and to depend on age and follow-up time.

DOI: 10.1002/sim

Finally, extrapolations of incidence and survival models together with hypothetical scenarios of future screening policy were used to calculate breast cancer mortality predictions up to 2012. The precision assessment of the predictions was based on variation in the breast cancer mortality estimates.

Following the cure model fitted for relative survival by Berkson and Gage [15], a cure model for cause-specific survival was applied to the Finnish breast cancer data due to its good fit. In addition, Verdecchia *et al.* [16] have applied a cure model for relative survival to U.S. breast cancer patients. Distinguishing between delayed mortality and cure may, however, be a problem [17], and empirical data so far do not show that any point of cure would actually be reached for breast cancer patients [18]. Thus, the cure model with its interpretation can well be regarded as a tentative solution. If cure is not a reasonable assumption, the interpretation of the cure proportion and the LEs for the uncured patients may not be sensible. In terms of the short-term predictions in this study this issue is not necessarily that important, but a warning may well be in place for the use and interpretation of these parameters in general.

Part of the short LEs for uncured patients diagnosed during the years between invitations and 5-year post-screening period, as well as during irregular breaks in the screening programme (Table III), may be explained by more aggressive tumour types. These are, at least in theory, patients with a former negative result at screening. It should also be notified that the proportions of cured are higher in invited than in non-invited women (Table II). Suppose that as a consequence of screening, more cancers are detected in an early stage before the disease becomes unresponsive to treatment. In this case, screening removes some patients with relatively good prognosis from the group of fatal cases. Since this is not happening in the group of non-invited women, their prognoses are better on average.

We did not actually have any *a priori* expectations that the breast cancer patients' other cause mortality would differ from that of the general population, but since results showing this kind of selection had been reported by others [7, 9], we thought that this issue is worth to consider. The decrease in the other cause mortality compared with the general population during the first year after diagnosis among 50+ women with a localized breast cancer, and the increase starting from the third year until the end of the follow-up period among young (40–49) patients with a non-localized diagnosis are noteworthy results. In localized cases, an explanation may be the 'healthy patient effect', whereby the patients experience lower mortality due to other causes as a result of having greater than average contact with the health system. Among non-localized cases one possible explanation is that patients diagnosed with an advanced cancer are more frail. Even so, the risk of dying from other causes after being diagnosed with a breast cancer is clearly dependent at least on stage, age at diagnosis, and follow-up time, and there could also be other factors not studied here explaining the results. On the other hand, it is the other cause survival, not the death probability that determines the breast cancer mortality; the annual probabilities of death in this age group are small even after correcting them, leading to very minor modifications in overall survival probabilities. We have also conducted sensitivity analyses by fixing $r = 1$, and the obtained overall results for breast cancer mortality were very similar.

The nonexistent differences in breast cancer mortality rates between alternative future scenarios may be surprising. In fact, they all fall within the confidence limits presented in Figure 4, except the prediction based on scenario III in non-localized cases which is at the lower borderline. However, one limitation in our analysis is that the results are restricted to ages below 75. As both the lead and survival times in breast cancer are long [19], it might be that benefits of extending the programme to 69 or 65 years old women would be visible only above the age of 74. The results can

DOI: 10.1002/sim

also partly be explained by invitational mammographic tests paid by women themselves, which are known to be common among Finnish women above 60. Even so, it is unlikely that these limitations would explain the similarities between alternative predictions in full. It is possible that the current programme inviting women between ages 50 and 59 is sufficient, and potential decreases in mortality obtained by extending the programme to older women are also achieved by effective treatments [20]. As medical treatments progress, it may be that the prognoses for cancers detected in elderly women only when they give clinical symptoms are as good as for those preclinical diseases detected by screening.

Presentation of the prediction intervals would be an appropriate way to measure the uncertainty involved in mortality predictions. However, since the mortality is a combination of different, independently fitted models, a common variance for the parameters in mortality model cannot directly be calculated. As a solution, we present an approach where only the uncertainty related to breast cancer survival is taken into account, leaving out the contributions of incidence and other cause mortality models. As in Hakulinen and Dyba [21], the uncertainty related to incidence analysed using Poisson regression could be approximated using the delta method based on Taylor series expansion. Also bootstrap methods or Bayesian approaches could be useful. The variation in the other cause mortality estimates consists of the error related to population death probabilities and the error due to the estimates of correction coefficient r; the former can be assumed to be small and the latter was shown to be minor in a sensitivity analysis, see above. In conclusion, if the variation of all parameters in the mortality model was taken into account, the true prediction intervals of breast cancer mortality predictions would be wider than the presented ones, the increase in uncertainty originating mainly from the incidence model.

As stated above, even if the confidence intervals for the LE and cure fraction are relatively wide (Table III), the uncertainty in breast cancer survival and further in mortality estimates is acceptable due to the high negative correlation between the estimates (Figures 3 and 4). When the LE in a certain subgroup defined by covariates increases, the proportion of cured decreases, and *vice versa*. It is very unlikely of LE and proportion of cured to have their maximum or minimum extreme values at the same time.

Since this paper deals with an invitational screening programme, the presented methodology is for calculating the incidence-based mortality ignoring the prevalent cases diagnosed before the beginning of mammography screening in Finland. A beneficial way to extend this method would be to incorporate the mortality caused by the cancers diagnosed in the past to obtain the total breast cancer mortality. In this case, the breast cancer survival model for the incident cases from pre-observation (and pre-screening) period should be fitted again without screening as a covariate.

The method outlined shows how the incidence, survival, and both observed and hypothetical patterns of population-based cancer screening may be utilized for prediction of future disease-specific mortality in the population. When appropriate, the cause-specific survival can be exchanged with relative survival, which provides flexibility in using the method. Also by considering alternative parametric distributions, the model can be applied to other sites of cancer subjected to screening, for instance, cervix or prostate.

ACKNOWLEDGEMENTS

This work was financially supported by the Cancer Society of Finland, Academy of Finland (MaDaMe project) and Doctoral Programs in Public Health/Department of Public Health of the University of Helsinki, Finland.

DOI: 10.1002/sim

REFERENCES

1. Hakama M, Pukkala E, Söderman B, Day N. Implementation of screening as a public health policy: issues in design and evaluation. *Journal of Medical Screening* 1999; **6**:209–216.
2. Anttila A, Koskela J, Hakama M. Programme sensitivity and effectiveness of mammography service screening in Helsinki, Finland. *Journal of Medical Screening* 2002; **9**:153–158.
3. Parvinen I, Helenius H, Pylkkänen L, Anttila A, Immonen-Räihä P, Kauhava L, Räsänen O, Klemi PJ. Service screening mammography reduces breast cancer mortality among elderly women in Turku. *Journal of Medical Screening* 2006; **13**:34–40.
4. Sarkeala T, Heinävaara S, Anttila A. Organised mammography screening reduces breast cancer mortality: a cohort study from Finland. *International Journal of Cancer* 2008; **122**:614–619.
5. Seppänen J, Heinävaara S, Hakulinen T. Influence of alternative screening scenarios on breast cancer incidence predictions (Finland). *Cancer Causes and Control* 2006; **17**:1135–1144.
6. Seppänen J, Heinävaara S, Anttila A, Sarkeala T, Virkkunen H, Hakulinen T. Effects of different phases of an invitational screening programme on breast cancer incidence. *International Journal of Cancer* 2006; **119**:920–924.
7. Heinävaara S, Hakulinen T. Predicting the lung cancer burden: accounting for selection of the patients with respect to general population mortality. *Statistics in Medicine* 2006; **25**:2967–2980.
8. De Angelis R, Capocaccia R, Hakulinen T, Söderman B, Verdecchia A. Mixture models for cancer survival analysis: application to population-based data with covariates. *Statistics in Medicine* 1999; **18**:441–454.
9. Phillips N, Coldman A, McBride ML. Estimating cancer prevalence using mixture models for cancer survival. *Statistics in Medicine* 2002; **21**:1257–1270.
10. Sarkeala T, Anttila A, Forsman H, Luostarinen T, Saarenmaa I, Hakama M. Process indicators from ten centres in the Finnish breast cancer screening programme from 1991 to 2000. *European Journal of Cancer* 2004; **40**:2116–2125.
11. Clayton D, Schifflers E. Models for temporal variation in cancer rates. I: age–period and age–cohort models. *Statistics in Medicine* 1987; **6**:449–467.
12. SAS Institute Inc. *SAS/STAT® User's Guide*, *Version 8*. SAS Institute Inc.: Cary, NC, 1999; 2371–2418.
13. Chatfield C. *Time-series Forecasting*. Chapman & Hall: Boca Raton, FL, 2001; 151–180.
14. Yu B, Tiwari RC, Cronin KA, Feuer EJ. Cure fraction estimation from the mixture cure models for grouped survival data. *Statistics in Medicine* 2004; **23**:1733–1747.
15. Berkson J, Gage RP. Survival curves for cancer patients following treatment. *Journal of American Statistical Association* 1952; **47**:501–515.
16. Verdecchia A, De Angelis G, Capocaccia R. Estimation and projections of cancer prevalence from cancer registry data. *Statistics in Medicine* 2002; **21**:3511–3526.
17. Lambert PC, Thompson JR, Weston CL, Dickman PW. Estimating and modelling the cure fraction in population-based cancer survival analysis. *Biostatistics* 2007; **8**(3):576–594.
18. Brenner H, Hakulinen T. Are patients diagnosed with breast cancer before age 50 years ever cured? *Journal of Clinical Oncology* 2004; **22**:432–438.
19. Paci E, Miccinesi G, Puliti D, Baldazzi P, De Lisi V, Falcini F, Cirilli C, Ferretti S, Mangone L, Finarelli AC, Rosso S, Segnan N, Stracci F, Traina A, Tumino R, Zorzi M. Estimate of overdiagnosis of breast cancer due to mammography after adjustment for lead time. A service screening in Italy. *Breast Cancer Research* 2006; **8**:R68.
20. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, Mandelblatt JS, Yakovlev AY, Habbema JDF, Feuer EJ. Effect of screening and adjuvant therapy on mortality from breast cancer. *The New England Journal of Medicine* 2005; **353**:1784–1792.
21. Hakulinen T, Dyba T. Precision of incidence predictions based on Poisson distributed observations. *Statistics in Medicine* 1994; **13**:1513–1523.

DOI: 10.1002/sim