

Keywords: breast cancer; service screening; mammography; mortality; cohort study; New Zealand

Mammography service screening and breast cancer mortality in New Zealand: a National Cohort Study 1999–2011

Stephen Morrell¹, Richard Taylor^{*1}, David Roder², Bridget Robson³, Marli Gregory⁴ and Kirsty Craig⁴

¹School of Public Health and Community Medicine, University of NSW, Level 2, Samuels Building, Randwick, NSW, Australia;

²Sansom Institute for Health Research, University of South Australia, Adelaide, SA, Australia; ³University of Otago, Wellington, New Zealand and ⁴Breast Screen Aotearoa, National Screening Unit, Ministry of Health, Wellington, New Zealand

Background: This breast cancer mortality evaluation of service screening mammography in New Zealand, the first since commencement of screening in 1999, applies to the 1999–2011 diagnostic period. Individual-level linked information on mammography screening, breast cancer diagnosis and breast cancer mortality is used to analyse differences in breast cancer mortality according to participation in organised screening mammography, as provided by BreastScreen Aotearoa (BSA).

Methods: Women were followed from the time they became eligible for screening, from age 50 years (1999–2004) and 45 years (≥ 2004). Breast cancer mortality from cancers diagnosed during the screening period from 1999 to 2011 ($n = 4384$) is examined in relation to individual screening participation or non-participation during preceding person-years of follow-up from the time of screening eligibility. To account for changes from never- to ever-screened status, breast cancer mortality is calculated for each year in relation to prior accumulated time of participation and non-participation in screening. Breast cancer mortality is also examined in regularly screened women (screened ≥ 3 times and mean screening interval ≤ 30 months), and irregularly screened women compared with never-screened women. Statistical analyses are by negative binomial and Poisson regression with adjustment for age and ethnic group (Māori, Pacific women) in a repeated-measures analysis. Relative risks for breast cancer mortality compared with never-screened women, are adjusted also for screening selection bias, to indicate the extent of breast cancer mortality reduction in a population offered and not offered mammography screening. Prognostic indicators at diagnosis of breast cancer are also compared between different screening participation groups, including by grade of tumour, extent of disease (spread), multiple tumour status and maximum tumour size using χ^2 statistics, t -tests and two-sample median tests.

Results: For 1999–2011, after adjusting for age and ethnicity, breast cancer mortality in ever-screened women is estimated to be 62% (95% CI: 51–70) lower than in never-screened women. After further adjustment for screening selection bias, the mortality reduction in NZ is estimated to be 29% (95% CI: 20–38) at an average screening coverage of 64% for 2001–2011, and 34% (95% CI: 25–43) for recent screening coverage (2012–13, 71%). For irregularly screened women, the mortality reduction is estimated to be 31% (95% CI: 21–40), and 39% (95% CI: 22–52) in regularly screened women compared with never-screened women, after adjusting for age, ethnicity and screening selection bias (using recent 2012–2013 screening coverage of 71%). Ever-screened women diagnosed with breast cancer have more favourable prognostic indicators than never-screened women, with a higher proportion of localised cancer (63 compared with 46%), a higher proportion with a well-differentiated tumour (30 compared with 18%), lower risk of multiple tumours (RR = 0.48) and smaller median tumour size (15 mm compared with 20 mm)—all differences are statistically significant ($P < 0.0001$).

Conclusions: This is the first total population cohort study of an established nation-wide screening mammography programme using individual-level information on screening participation and mortality outcomes from breast cancer. The findings are in accord with other mammography screening service evaluations and with randomised trials of mammography screening.

*Correspondence: Professor R Taylor; E-mail: r.taylor@unsw.edu.au

The evidence for reduction in breast cancer mortality associated with screening mammography is based on a number of randomised trials conducted in the 1960s, 1970s and 1980s of individuals or populations invited and not invited to screening. Screening recommendations have been based on the results of such trials from Sweden (Tabár *et al*, 1992; Frisell *et al*, 1997), Edinburgh (Alexander *et al*, 1994), New York (Shapiro, 1997) and Canada (Miller *et al*, 1992). Meta-analyses of these trials have suggested reductions in breast cancer mortality in 24–31% of those screened (Kerlikowske *et al*, 1995; Nystrom *et al*, 1996; Nelson *et al*, 2016) and 20–30% in those invited to screening (Tabár *et al*, 2001, 2003). Meta-analyses that exclude several studies because of possible randomisation bias have failed to show an effect of mammography screening on breast cancer mortality (Gøtzsche and Olsen, 2000; Gøtzsche and Nielsen, 2006, 2009, 2011; Gøtzsche and Jørgensen, 2013). This result is due almost entirely to a negative finding from the Canadian breast screening trial (Freedman *et al*, 2004). A review by the International Agency for Research on Cancer (IARC) produced a meta-analysis that indicated a pooled relative risk (RR) of 0.75 (25% breast cancer mortality reduction) for invitation to mammography screening in women aged 50–69 years (International Agency for Research on Cancer, 2002, 2016).

Randomised trials based on invitation to screening (as an ‘intention-to-treat’ analysis) may underestimate the benefit of screening participation because of, *inter alia*, non-adherence in the intervention group and screening in the control group. The effect of actual screening on breast cancer mortality among screening participants has been estimated to be a 35% reduction compared with those not screening (International Agency for Research on Cancer, 2016). However, this effect may be biased by self-selection into screening, in that those who screen may also have lower breast cancer mortality for non-screening reasons (Duffy *et al*, 2002a).

Service studies of mammography screening using a variety of methodologies in the United Kingdom (Quinn and Allen, 1995; UK Trial of Early Detection of Breast Cancer group, 1999; Blanks *et al*, 2000; Threlfall *et al*, 2003), Holland (Broeders *et al*, 2001; Otto *et al*, 2003), Finland (Anttila *et al*, 2002), Sweden (Jonsson *et al*, 2001; Tabár *et al*, 2001, 2003; Duffy *et al*, 2003) and Australia (Taylor *et al*, 2004, 2009; Roder *et al*, 2008; Morrell *et al*, 2012; Nickson *et al*, 2012) have indicated lower mortality associated with screening compared with non-screened populations, although not all results reached statistical significance. Australian studies have shown significant breast cancer mortality reductions associated with screening mammography using a variety of study designs and analytical approaches (Taylor *et al*, 2004, 2009; Roder *et al*, 2008; Morrell *et al*, 2012).

The purpose of population-based mammography screening programmes is to reduce mortality from breast cancer, and such benefits associated with established screening programmes, including that of New Zealand, need to be evaluated to determine whether this purpose is being achieved in a real-world setting. Other benefits include less extensive surgical treatment, radiotherapy and cytotoxic and other pharmaceutical treatment for cancers that are diagnosed at an earlier stage. This needs to be balanced against the potential harms associated with screening mammography, including false positives and subsequent unnecessary investigations, and possible over-diagnosis and accompanying overtreatment (Marmot *et al*, 2013). It is thus important that the extent of breast cancer mortality benefit from established screening be quantified. In the context of an established evidenced-based programme, it is not possible or appropriate to conduct a randomised controlled trial (RCT) of mammography screening to assess its impact on breast cancer mortality. Numerous observational study designs are available to assess breast cancer mortality in relation to participation in the BreastScreen Aotearoa (BSA) programme (Aotearoa being the Indigenous Māori name for New Zealand).

The BSA program commenced operations in December 1998 targeting women aged 50–64 years, with age extended to 45–69 years in 2004. Full geographic coverage was achieved in 1999. The screening interval is two-yearly and two-view bilateral mammograms are performed and read independently by two radiologists, with arbitration by a third or consensus group for conflicting findings. Recruitment to screening is by general health promotion activities, and retention to screening is by personalised reminders to screened women. Digital mammography was first implemented in February 2006 and mammography screening was completely digital by 2013. Cancer detection rates have been within expected (40 per 10 000 women screened), and recall-to-assessment rates also within the quality assurance guidelines (Page *et al*, 2014). Biennial screening attendance rose from 54% for all women (35% for Māori and Pacific women) in the early years of the program, and by 2011 it had risen to 71% for all women and Pacific women, and 63% for Māori women (Page *et al*, 2014).

Given the quality and universality of the National Health Index (NHI) linkage key in New Zealand health data, an historical population cohort study of individuals is feasible and was considered to provide the strongest observational study design and largest numbers to assess the effectiveness of the BSA programme. Individual-level screening participation and breast cancer incidence and mortality outcome data were available. A population cohort study offers advantages over a case-control design, in that it circumvents questions concerning selection and appropriateness of controls. The main disadvantage of an historic population cohort design is bias from loss to follow-up, or attrition, especially from outmigration, and inaccuracies in data linkage. However, the procedure in New Zealand for health data linkage is well established and likely to provide reliable results with only a small proportion of mismatches (C Lewis, personal communication).

The aim of this study is to provide an indicator of screening performance in New Zealand based on real-life monitoring data; it is not intended to constitute the underpinning scientific justification for screening, which was established by the RCTs. The hypotheses investigated are that breast cancer mortality is lower in ever-screened women compared with never-screened women, and in women with more compared with less regularity of screening, and that ever-screened women will have prognostic factors at diagnosis of breast cancer indicative of a more favourable outcome than never-screened women.

MATERIALS AND METHODS

Study design and data. This is a retrospective cohort study of breast cancer mortality in New Zealand women in relation to screening mammography. The New Zealand NHI is used to link individual data from the BSA screening service and cancer and death registries, to assemble a cohort comprising all New Zealand women aged 45–69 years during 1999–2011 who were ever screened or diagnosed with breast cancer. Never-screened women not otherwise linkable via the NHI in a given year are inferred by subtraction from ethnic- and age-specific census-derived populations for that year, as provided by Statistics New Zealand (Table 1).

Analytic approach. From 2000 to 2011, breast cancer mortality in each year was calculated in relation to screening participation or non-participation measured in person-years cumulated individually up to the beginning of each year. Breast cancer mortality occurring originates from cumulated prior breast cancer incidence from 1999 onwards. Women were followed from the time they first became eligible to screen, as a ‘screening-inception’ cohort with differing person-years based on participation and non-participation in BSA

Table 1. Cohort populations, annual breast cancer mortality from breast cancers diagnosed in 1999–2011, and person-years of exposure in ever- and never-screened^a women aged ≥ 45 years, by year of death

Year	Ever-screened women			Never-screened women		
	Number at the beginning of each year	Cumulated person-years exposure ^b	Breast cancer deaths	Number at the beginning of each year	Cumulated person-years exposure ^b	Breast cancer deaths
2000	75 562	75 562	5	580 494	580 494	112
2001	154 117	229 612	17	516 246	1 032 492	193
2002	189 075	418 150	27	503 245	1 376 937	244
2003	216 368	633 114	45	492 249	1 696 190	264
2004	240 501	871 106	64	485 611	2 006 453	279
2005	264 539	1 132 195	69	479 766	2 301 812	336
2006	306 023	1 432 983	79	456 778	2 560 596	315
2007	359 385	1 785 670	97	430 703	2 750 229	331
2008	408 896	2 185 630	95	401 855	2 877 408	350
2009	453 086	2 627 404	120	378 058	2 982 745	377
2010	497 697	3 110 437	132	352 890	3 062 181	340
2011	542 234	3 635 261	123	327 623	3 113 051	370
2000–11		3 707 483	873		5 405 518	3511
Unadjusted 2000–2011 ever-screened : never-screened mortality proportional difference = – 63.7%						
^a Ever or never screened with BreastScreen Aotearoa.						
^b As at the beginning of the year allowing for entry to and exit from cohort.						

mammography. In ever-screened women, person-years of participation in screening are calculated from the time of first screen to the beginning of each successive year, or to the year of diagnosis for women diagnosed with breast cancer—since screening participation postdiagnosis is not relevant to breast cancer mortality in this study. Every woman for whom individual data were available was classified as not screened from the time they first became eligible to screen (i.e., at age 50 years before 2004, and at 45 years from 2004) to the time they first screened. This period contributed to the woman's person-years of not screening with BSA. The time from her first screen to either: (i) the end of the study period; (ii) the woman's death; or (iii) to the first diagnosis of breast cancer; counted as person-years since first participating in screening. Thus, at the beginning of each year, each woman may contribute some person-years of cumulated participation and non-participation in screening depending on their age, when they first became eligible to screen, when organised screening commenced and the screening target age range over the period of interest.

For each successive year, cumulated person-years of participation and non-participation in screening (while eligible) are recalculated taking into account women newly screened since the previous year, those who become eligible by ageing into the screening target age group in that year and those ageing out of the screening target age group in that year (along with women diagnosed with breast cancer or who have died). For 1999–2003, person-years of participation or non-participation in screening are calculated for the target group of women aged 50–64 years. From extension of the target age range in 2004, the calculation is for women aged 45–69 years. The overall approach is illustrated in Figure 1. For the remainder of this article, we use the term 'screening exposure' synonymously with 'participation in screening'.

For women with no recorded screening participation, their person-years of non-participation in screening were calculated as above. For women without individual data (i.e., the remaining female population with no recorded screening or breast cancer history), the screening commencement age was subtracted from the median age in each 5-year age group for a given year (supplied in 5-year age groups). Person-years formed the denominators for subsequent analyses of those ever- or never-screened up to the given year of interest.

The outcome variable is breast cancer mortality occurring over 2000–2011 originating only from cancers diagnosed during 1999–2011 in women in the screening target age groups. With this approach, lead-time bias does not affect the estimates of breast cancer mortality risk in screened compared with unscreened women, as breast cancer mortality is examined in relation to time since first participating in screening, and time not participating in screening from the time of first screening eligibility, not time from diagnosis as in a clinical cohort from incident cases.

Mortality analyses. Relative risks for breast cancer mortality are determined from comparison of mortality in those ever screened compared with the never screened (RR = 1.00), and adjusted by regression analysis for confounding by 5-year age group at death (45–49 to 75–79 years) and ethnicity: Māori, Pacific and Other women. Māori are the indigenous people of New Zealand and comprise 9.6% of the female population aged 45–69 years (2006 Census). Pacific women comprise migrants or their descendants from Pacific Islands states, especially from Samoa, Tonga, Cook Islands and Niue who comprise 3.9% of the 45–69-year female population (2006 Census). The main ('other') component of the population consists predominantly of New Zealanders of European descent. Ethnicity in New Zealand is based on self-designation, and in this study ever designated as Māori (or Pacific) from any of the ethnicity indicator variables available across data sources ('prioritised' ethnicity) is used to attribute ethnicity. Adjustment for confounding by age group and ethnicity is to account for differences in distribution of such groups with variation in breast cancer mortality (from reasons other than screening) in those never screened, ever screened, irregularly screened or regularly screened.

Analyses are undertaken of effects of screening regularity defined as: screened at least three times with a mean screening interval of 30 months or less, as previously used in a South Australian case-control study (Roder *et al.*, 2008). Irregular screening is defined as ever screened, but not conforming to this definition of regular screening. These categories were compared with breast cancer mortality in the never screened. To test for a screening dose-response relationship, the statistical test of significance involved modelling screening regularity as an ordinal covariate (with a single degree of freedom) in a regression analysis

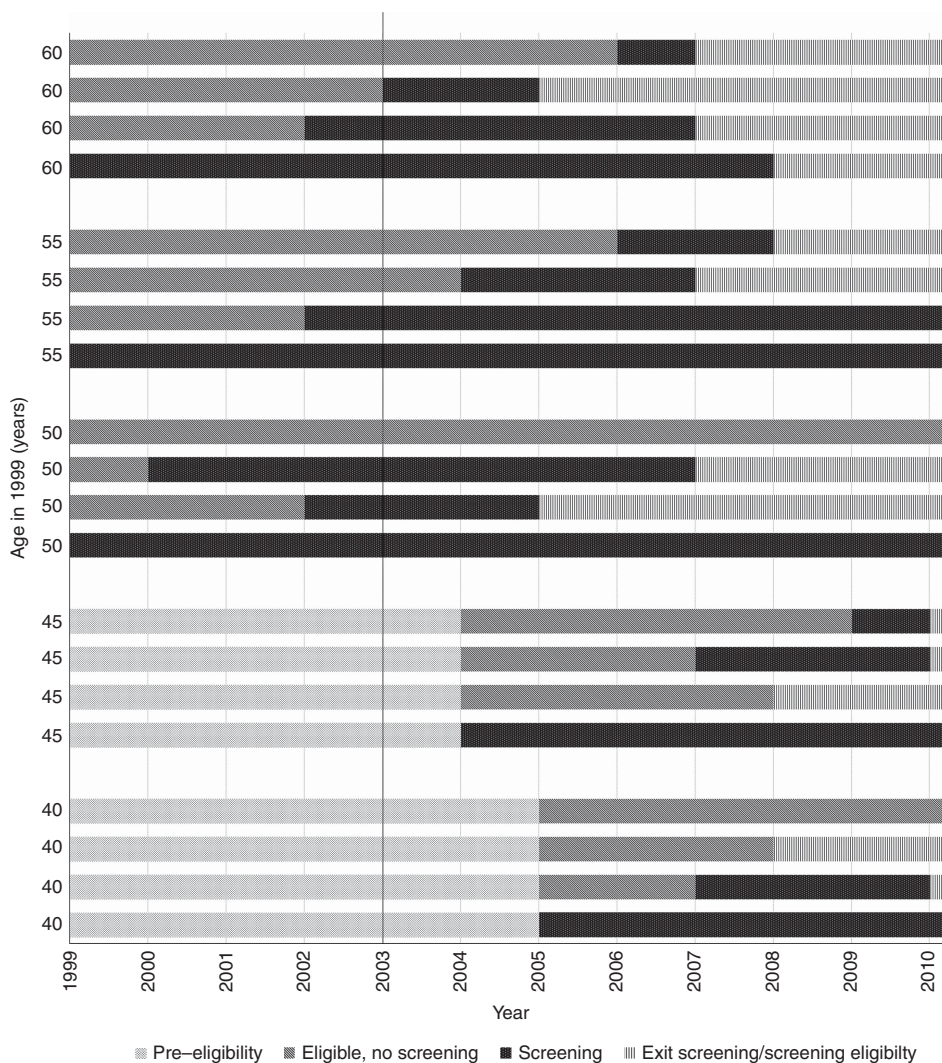


Figure 1. Individual examples of person-year contributions to screening exposure and non-exposure under a variety of screening participation scenarios. The total cumulated person-years of participation and non-participation in screening up to the end of 2003 (vertical line) in a hypothetical cohort of 20 women aged 40–65 years in 1999. Cumulated person-years contributing to participation is 18, and the total cumulated person-years from first eligibility contributing to screening non-participation is 30.

with the significance of the β -estimate serving as the test of trend across never-, irregularly- and regularly-screened women.

Statistical modelling was by negative binomial regression adjusting for repeated measures, as breast cancer mortality in a largely similar population was analysed repeatedly each year to account for changing screening status in the population. Where negative binomial models failed to converge, Poisson regression is used and standard errors corrected for overdispersion. Counts of breast cancer deaths are the outcome and the offset is the log of person-years of exposure or non-exposure to screening. With increasing time, the total person-years of exposure and non-exposure to screening cumulate; correspondingly, annual breast cancer mortality counts derive from a cumulation of incident breast cancer cases diagnosed over the same time period preceding the given year of death, as occurs in all incidence-based analyses of breast cancer mortality. SAS version 9.4 (SAS Institute Inc., Cary, NC, USA) is used for statistical analyses.

Screening selection bias. Relative risks are further adjusted for screening selection bias, which results from comparison of breast cancer mortality in screened women compared with those who do

not participate in screening despite its availability (Duffy *et al*, 2002a). In some populations, the latter have been shown to have higher breast cancer mortality compared with women not offered screening, some of whom would screen if it were available. Such comparisons can thus inflate estimates of mortality reduction from screening compared with a population never offered screening, and corrections have been used frequently in previous studies (Duffy *et al*, 2002b; Gabe and Duffy, 2005; Swedish Organised Service Screening Evaluation Group, 2006) to produce estimates of screening effects that are comparable to results of randomised trials. The correction relies on: (1) empirical estimates from RCTs and screening service studies of breast cancer mortality RR in women not screening when offered compared with women not offered screening (D_r); (2) an estimate of population-based screening participation (p); and (3) the empirical RR estimates of never- vs ever-screening derived from the Poisson or negative binomial modelling (RR_{der}) from the present study of New Zealand women. The resulting adjusted RR estimate is an intention-to-treat estimate and represents the RR in a population where screening is available or offered compared with a population where screening is unavailable or not offered. In detail (Duffy *et al*, 2002a):

The adjusted RR is

$$RR_{adj} = D_r(pRR_{der} + 1 - p)$$

The variance for RR_{adj} is calculated from

$$V\{\ln(RR_{adj})\} = V\{\ln(D_r)\} + \frac{p^2(RR_{der})^2 V\{\ln(RR_{der})\}}{(pRR_{der} + 1 - p)^2}$$

The standard error is then

$$s.e.(\ln(RR_{adj})) = \sqrt{V\{\ln(RR_{adj})\}}$$

The upper and lower 95% confidence intervals of $\ln(RR_{adj})$ are

$$\ln(RR_{adj}) \pm 1.96s.e.(\ln(RR_{adj}))$$

and these are exponentiated to produce the upper and lower 95% confidence intervals of RR_{adj} .

The excess breast cancer mortality of women not screening over women not offered screening, reported in the literature and used in the present study, is an RR of 1.17, from the Swedish screening service studies (Swedish Organised Service Screening Evaluation Group, 2006) considered appropriate for evaluation of the New Zealand screening program because it emanates from a service screening environment similar to New Zealand. The variance of $\ln(D_r)$, ($V\{\ln(D_r)\}$), is estimated as 0.0014995, derived from the 95% confidence interval reported for the overall estimate of D_r by the Swedish Organised Service Screening Evaluation Group (Swedish Organised Service Screening Evaluation Group, 2006) and $V\{\ln(RR_{der})\}$ is provided directly by the regression outputs from the present analyses, as the standard error of the regression estimate squared.

Adjustment for screening selection bias also incorporates screening coverage (Duffy *et al*, 2002a). Accordingly, adjusted estimates for different screening participation rates are derived from the mean recorded participation rate for: 2001–2011 (64%), the most recent period 2012–2013 (71%; Page *et al*, 2014), and the screening participation target of 70%.

Prognostic indicators. Prognostic indicators at diagnosis of breast cancer are compared between never- and ever-, and between

irregularly- and regularly-screened groups. Prognostic indicators included grade of tumour, extent of disease (spread), multiple tumours and maximum tumour size. To assess possible down-shifting of cancer stage from screening, partly attributable to possibly inconsequential cancers, examination of proportions of cancers with regional or distant spread is also undertaken. In particular, the proportions of distant cancer of distant + regional cancer were compared between the screening comparison groups to eliminate differences in proportions of distant cancer being attributable only to a screening-related inflation of localised cancer.

Chi-squared, two-sample median and *t*-tests for differences in proportions, medians and means, respectively, are used for testing ever- and never-screened differences in prognostic indicators.

RESULTS

Breast cancer mortality reduction in relation to mammography screening. Breast cancer mortality for 2000–2011 (from cancers diagnosed from 1999) was 23.5 per 100 000 for ever screened (873 deaths, $n = 3\,707\,484$) and 65.0 per 100 000 for never screened (deaths = 3511, $n = 5\,405\,518$). The unadjusted mortality ratio is thus 0.36, indicating a mortality reduction from screening of 64% (Table 1). Using person-years as the denominator and adjusting for age and ethnicity by negative binomial regression, breast cancer mortality in ever-screened women is estimated to be 62% (95% CI: 51–70) lower compared with that in never-screened women based on the RR adjusted for age and ethnicity (Table 2). After further adjustment for screening selection bias, the mortality reduction is estimated to be 29% (95% CI: 20–38) at average screening participation of 64% for 2001–2011. For recent (2012–2013) screening coverage (71%), the estimated mortality reduction is estimated as 34% (95% CI: 25–43).

Compared with never-screened women, breast cancer mortality is estimated to be 58% (95% CI: 48–66) lower in irregularly screened women, and 67% (95% CI: 46–81) lower in regularly

Table 2. Adjusted relative risk^a of breast cancer mortality in ever- and never-screened^b New Zealand women, 1999–2011

Variable	Regression estimate (s.e.)	P-value	Relative risk (95% CI)	% Mortality difference (95% CI)
Screening status				
Never screened			1.00	0
Ever screened	−0.9685 (0.1273)	<0.0001	0.38 (0.30–0.49)	−62 (−70 to −51)
Ever screened (adj.) ^c	—	—	0.71 (0.62–0.80)	−29 (−38 to −20)
Ever screened (adj.) ^d	—	—	0.66 (0.57–0.75)	−34 (−43 to −25)
Ever screened (adj.) ^e	—	—	0.66 (0.58–0.76)	−34 (−42 to −24)
Age at death (years)				
60–64 (referent)			1.00	
45–49	−0.0628 (0.0512)	0.2198	0.94 (0.85–1.04)	
50–54	−0.2660 (0.0742)	0.0003	0.77 (0.66–0.89)	
55–59	−0.1378 (0.0582)	0.0179	0.87 (0.78–0.98)	
65–69	−0.0608 (0.0336)	0.0704	0.94 (0.88–1.01)	
70–74	−0.1211 (0.0625)	0.0526	0.89 (0.78–1.00)	
75–79	−0.0676 (0.0846)	0.4242	0.93 (0.79–1.10)	
80–84	−0.0072 (0.1587)	0.9637	0.99 (0.73–1.36)	
85+	0.5230 (0.1082)	<0.0001	1.69 (1.36–2.09)	
Ethnicity				
Other			1.00	
Māori	0.5130 (0.0475)	<0.0001	1.67 (1.52–1.83)	
Pacific	0.4668 (0.0606)	<0.0001	1.59 (1.42–1.80)	
Intercept	−8.8737 (0.0784)	<0.0001	—	—

Abbreviation: CI = confidence interval.

^aFrom negative binomial regression model adjusted for repeat measures.

^bEver or never screened with BreastScreen Aotearoa.

^cAdjusted for screening selection bias, assuming relative risk in non-screeners to women not offered screening = 1.17 and mean screening participation rate of 64% for 2001–2011.

^dAdjusted for screening selection bias, assuming relative risk in non-screeners to women not offered screening = 1.17 and recorded screening participation rate of 71% for 2012–2013.

^eAdjusted for screening selection bias, assuming relative risk in non-screeners to women not offered screening = 1.17 and target screening participation rate of 70%.

Table 3. Adjusted relative risk^a of breast cancer in regularly, irregularly and never-screened^b New Zealand women, breast cancers diagnosed 2003–2011

Variable	Regression estimate (s.e.)	P-value	Relative risk (95% CI)	% Mortality difference (95% CI)
Screening status				
Never screened			1.00	0
Not screened regularly	– 0.8685 (0.1117)	<0.0001	0.42 (0.34–0.52)	– 58 (– 66 to – 48)
Not screened regularly (adj.) ^c	—	—	0.74 (0.65–0.83)	– 26 (– 35 to – 17)
Not screened regularly (adj.) ^d	—	—	0.69 (0.60–0.79)	– 31 (– 40 to – 21)
Not screened regularly (adj.) ^e	—	—	0.69 (0.61–0.79)	– 31 (– 39 to – 21)
Screened regularly ^b	1.1231 (0.2622)	<0.0001	0.33 (0.19–0.54)	– 67 (– 81 to – 46)
Screened regularly (adj.) ^c	—	—	0.67 (0.55–0.82)	– 33 (– 45 to – 18)
Screened regularly (adj.) ^d	—	—	0.61 (0.48–0.78)	– 39 (– 52 to – 22)
Screened regularly (adj.) ^e	—	—	0.62 (0.49–0.78)	– 38 (– 51 to – 22)
Screening category trend test ^f :	$\chi^2_{(1)} = 25.6 < 0.0001$			
Age at death (years)				
60–64 (referent)			1.00	
45–49	0.1149 (0.0696)	0.0987	1.12 (0.98–1.29)	
50–54	– 0.2360 (0.0736)	0.0013	0.79 (0.68–0.91)	
55–59	– 0.1400 (0.0328)	<0.0001	0.87 (0.82–0.93)	
65–69	0.0006 (0.0654)	0.9924	1.00 (0.88–1.14)	
70–74	– 0.1312 (0.0750)	0.0802	0.88 (0.76–1.02)	
75–79	– 0.2142 (0.0863)	0.0131	0.81 (0.68–0.96)	
Ethnicity				
Other			1.00	
Māori	0.5925 (0.0409)	<0.0001	1.81 (1.67–1.96)	
Pacific	0.5550 (0.0586)	<0.0001	1.74 (1.55–1.95)	
Intercept	– 8.9849 (0.0932)	<0.0001	—	—

Abbreviation: CI = confidence interval.

^aFrom negative binomial regression model adjusted for repeat measures.

^bScreened ≥ 3 times, ≤ 30 months mean screening interval, as screened by BreastScreen Aotearoa.

^cAdjusted for screening selection bias, assuming relative risk in non-screeners to women not offered screening = 1.17 and mean screening participation rate of 64% for 2001–2011.

^dAdjusted for screening selection bias, assuming relative risk in non-screeners to women not offered screening = 1.17 and recorded screening participation rate of 71% for 2012–2013.

^eAdjusted for screening selection bias, assuming relative risk in non-screeners to women not offered screening = 1.17 and target screening participation rate of 70%.

^fTrend test from regressing screening category as ordinal variable: no screening = 1; irregular screening = 2; regular screening = 3.

screened women compared with never-screened women based on the RRs (Table 3). The indicator of regularity of screening in this analysis is screened at least three times with a mean screening interval of 30 months or less. Age at first screen, to indicate earlier or later age at first exposure to screening, is controlled for, along with ethnicity, to minimise confounding.

After adjustment for screening selection bias, the mortality benefit in women screened less regularly is estimated to be 26% (95% CI: 17–35) for screening participation of 64% for 2001–2011. Based on the most recent (2012–2013) screening rate of 71%, the mortality reduction is estimated to be 31% (95% CI: 21–40), similar to that for the screening participation target of 70%. The mortality benefit attributable to regular screening, adjusted for screening selection bias, is estimated as 33% (95% CI: 18–45), based on screening for 2001–2011, and 39% reduction based on the 2012–2013 screening coverage (71%). The test for trend across screening groups was significant ($P < 0.0001$), confirming a screening dose–response relationship in breast cancer mortality (Table 3 and Figure 2).

Prognostic factors in diagnosed breast cancer in relation to mammography screening. Ever-screened women diagnosed with breast cancer had more favourable prognostic indications than never-screened women, on all indicators in 2000–2011 (Table 4). A significantly higher proportion of ever-screened women diagnosed with breast cancer had well-differentiated tumours (30%) compared with 18% of never-screened diagnosed women ($P < 0.0001$), and 63% of diagnosed ever-screened women had localised cancer compared with 46% in never-screened women ($P < 0.0001$). These proportions and their differences are similar when recalculated excluding unknown or not recorded categories.

The proportion of breast cancers with distant spread, of distant plus regional spread, was 4.9% in ever-screened women and 11.3% in never-screened women, which equates to a corresponding RR for ever-screened women of 0.44 (95% CI: 0.37–0.52). This indicates that the lower proportion of distant cancer in ever-screened women was not entirely attributable to screening-related inflation of localised cancer.

Of ever-screened women diagnosed with breast cancer, 1.8% had multiple tumours compared with 3.8% of never-screened women, or RR = 0.48 (95% CI: 0.41–0.56) compared with never-screened women (RR = 1.00). The median maximum tumour size in ever-screened women with breast cancer was 15 mm compared with 20 mm in corresponding never-screened women ($P < 0.0001$, significantly smaller). Excepting grade of tumour, there were significantly better prognostic indicators evident in regular screening than in irregular screening for the degree of spread (localised 67 compared with 60%), multiple tumour status (RR = 0.57) and maximum tumour size (14 mm vs 15mm median) (Table 5).

DISCUSSION

All initial hypotheses proposed in this study were confirmed. Breast cancer mortality in New Zealand was lower in women ever screened by BSA compared with women never screened by BSA during 1999–2011, and lower in women with higher compared with lower regularity of screening; and ever-screened women show more favourable breast cancer prognostic factors than never-screened women. There is an evident dose–response relationship in breast cancer mortality reduction from never, to irregular, then to regular screening mirrored by differences in prognostic indicators

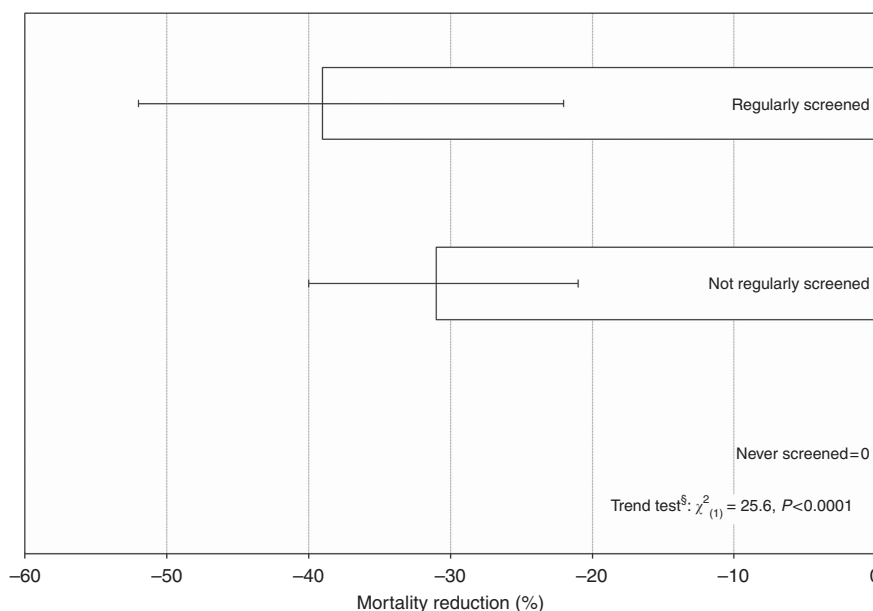


Figure 2. Differences (%)† in breast cancer mortality by mammography screening group, New Zealand women, 1999–2011.

†Adjusted for age and ethnicity by regression; and adjusted for screening selection bias (Duffy *et al*, 2002a) assuming relative risk in non-screeners to women not offered screening = 1.17 and recorded screening participation rate of 71% for 2012–2013. §Trend test of regression estimates.

in diagnosed cancers. Further, the documented mortality reductions from service screening (adjusted for screening selection bias) in this study are similar to the findings of meta-analyses of RCTs (Kerlikowske *et al*, 1995; Nystrom *et al*, 1996; Tabár *et al*, 2001, 2003; International Agency for Research on Cancer, 2016; Nelson *et al*, 2016) and previous studies of service screening (Broeders *et al*, 2012).

It is evident that the main source of the breast cancer mortality benefit in ever-screened women is the earlier detection of cancer, as indicated by more favourable prognostic indicators than in never-screened women. In particular, the mortality benefit from screening has stemmed largely from higher proportions with localised summary stage at diagnosis, and correspondingly lower proportions with regional or metastatic spread in the ever-screened compared with the never-screened women. Importantly, cancer with distant spread as a proportion of cancer with distant or regional spread at diagnosis is also significantly lower in ever-screened compared with never-screened women, which indicates an absence of artefactual stage shifting from detection by screening of possibly inconsequential subclinical cancers.

Tumour size, tumour grade, nodal status and degree of spread are inter-related. Well-differentiated cancer (grade) in a significantly higher proportion of ever-screened than never-screened women with cancer has been shown to be correlated with smaller tumour size (Duffy *et al*, 1991). Higher proportions of smaller tumour size and lower tumour grade, or less dedifferentiated cancer, in screened women compared with non-screened women, has been shown to be a consequence of early diagnosis rather than length bias (Duffy *et al*, 1991). The mortality findings are consistent with the prognostic indicators, and the evidence for a dose–response relationship is strong, with breast cancer mortality being lowest in regular screened women, higher in irregularly screened women and highest in never-screened women—consistent with the prognostic indicators.

The most direct and understandable outcome measure for this analysis is breast cancer mortality occurring each year, based on screening exposure up to the beginning of that year in women not diagnosed with breast cancer, and up to the beginning of the year of diagnosis in women with breast cancer. Some women classified as never screened up to a particular year will subsequently become

screened. The subsequent screening exposure, along with past non-screening exposure, will be relevant to subsequent breast cancer mortality, weighted by the time in each exposure category. This approach is not subject to lead-time bias as the screening exposure and non-exposure are measured as person-years from time of screening eligibility, and time of diagnosis of breast cancer is not used in the analysis. Breast cancer mortality analysed here is confined only to those breast cancer cases diagnosed after the advent of population screening mammography (1999).

While survival analysis of breast cancer cases based on time since diagnosis can be adjusted for competing causes of death and other factors, controlling for lead-time bias due to screening is exceedingly difficult. The present study has avoided artefactual contributions from lead time by not examining survival post-diagnosis but rather breast cancer mortality in relation to person-time exposed to screening up to the beginning of each successive year. Breast cancer mortality over that year is then examined. This process was repeated for each year successively, in a repeated-measures analysis of yearly breast cancer deaths.

As in most screening service studies, which by nature are observational, the factors contributing to differences in breast cancer mortality between those who participate in screening compared with those who do not cannot all be known or measured. Compared with women who do not screen despite its availability, it is possible that women who do screen when screening is available also have other (unmeasured and unknown) characteristics that may contribute to lower breast cancer mortality (in addition to screening itself). Although publicly funded screening and treatment services are available nationwide in New Zealand, there is some evidence of differential access or quality of care by screen-detection status and ethnic group. Surgery delays have been reported as more common in non-BSA diagnosed women and among Māori, and surgery delay in the public sector was also found to be more likely than in the private sector (Seneviratne *et al*, 2015; Lawrenson *et al*, 2016). Māori women have also been observed as less likely to have radiotherapy and/or adhere to long-term adjuvant endocrine therapy, and more likely to have a mastectomy, than non-Māori women (Lawrenson *et al*, 2016).

Table 4. Prognostic indicators for breast cancers diagnosed in ever- vs never screened^a women, New Zealand women aged 45–69 years at year of diagnosis, 2000–2011

Prognostic indicator	Ever screened		Never screened		Significance of difference
	n	%	n	%	
Grade of tumour					
Well differentiated	3725	29.6	1624	18	$\chi^2_{(1)} = 384, P < 0.0001$ $\chi^2_{(1)} = 321, P < 0.0001$
Excluding not recorded or NA (n = 1498)		31.2		19.9	
Moderately differentiated	5196	41.3	3679	40.3	
Poorly or undifferentiated or anaplastic	3014	24	2877	31.8	
Not recorded/NA	638	5.1	860	9.5	
Overall heterogeneity: $\chi^2_{(3)} = 558, P < 0.0001$					
Extent of disease					
Localised	7928	63.1	4169	46.1	$\chi^2_{(1)} = 613, P < 0.0001$ $\chi^2_{(1)} = 521, P < 0.0001$
Excluding unknown (n = 1697)		67.3		51.2	
Adjacent organ or regional lymph node	3662	29.1	3521	38.9	
Distant	190	1.5	447	4.9	
Distant±(distant+ regional)		4.9		11.3	RR = 0.44 (0.37–0.52) $\chi^2_{(1)} = 105, P < 0.0001$
Unknown	793	6.3	904	10	
Overall heterogeneity: $\chi^2_{(3)} = 724, P < 0.0001$					
Multiple tumours					
No	12 344	98.2	8697	96.2	RR = 0.48 (0.41–0.56) $\chi^2_{(1)} = 80.2, P < 0.0001$
Yes	229	1.8	344	3.8	
Maximum tumour size	mm		mm		
Mean	18.1		24.4		t = -25.8, P < 0.0001 z = -27.8, P < 0.0001 ^b
Median	15		20		

Abbreviation: NA = not available.
^aAs screened by Breastsreen Aotearoa.
^bTwo-sample median test.

Screening selection bias may also be affected in New Zealand by initial recruitment to screening through media campaigns and other health promotion activities, conducted by the Ministry of Health, rather than individual, personalised recruitment. Women more likely to respond to general health promotion activities may have lower risk of breast cancer mortality through lower breast cancer incidence and/or case fatality. While lower case fatality depends on lower cancer stage and grade, the main mechanism through which screening mammography lowers breast cancer mortality, as shown in the present study, is that breast cancers in screened women have better prognostic indicators than those in unscreened women. That is, lower case fatality is shown to be consistent with earlier detection through screening.

The principal advantage of adjustment for screening selection bias is that it accounts for possible differences between women who screen compared with women who do not screen, despite the availability of screening. An additional advantage is that this adjustment produces an estimate of the effects of screening mammography on breast cancer mortality in a population offered screening compared with a population not offered screening, to relate findings from screening service evaluation studies to RCTs in which entire populations offered screening are compared with populations not offered screening. This correction has been used extensively in published studies of screening evaluations, both in cohort studies (Tabár *et al*, 2001; Duffy *et al*, 2002b; Gabe and Duffy, 2005; Swedish Organised Service Screening Evaluation Group, 2006; Lawrence *et al*, 2009; Hofvind *et al*, 2013) and in case-control studies (Allgood *et al*, 2008; Puliti *et al*, 2008; Roder *et al*, 2008; Nickson *et al*, 2012; Otto *et al*, 2012; Paap *et al*, 2014; van der Waal *et al*, 2015).

A potential weakness in the present analysis is the use of mortality differentials from Swedish service screening studies to adjust for screening selection bias. The Swedish screening service studies found the RR of breast cancer mortality in women not screening, in spite of it being offered, compared with women not offered screening, to be 1.17 (Swedish Organised Service Screening

Evaluation Group, 2006). This RR estimate may be different to the New Zealand population, although the estimates of breast cancer mortality unadjusted for screening selection bias indicate that the anticipated direction of the effect of screening selection bias is correct, and the consequences of its application produces results that have face validity compared with data from the original trials. We have used the Swedish RR estimate in our study because of its ready availability and that it derives from screening programs in Sweden and New Zealand with apparent similarities. The aim of this study was to provide an indicator of screening performance in New Zealand based on real-life monitoring data; it is not intended to constitute the underpinning scientific justification for screening that was established by the RCTs.

Published evaluations of mammography service screening include case-control and cohort studies, although several cohort studies use a quasiexperimental study design with categorical exposure to screening (or non-screening) in aggregate for all individuals, such as before and after studies in the same population, or contemporaneous comparisons of populations in different geographic areas. Although the outcome is measured as a cohort mortality rate, the exposure is ecological. Such studies are susceptible to bias and confounding because of different characteristics of comparative populations that usually are not randomly selected.

Cohort studies of service screening using screening exposure measured in individuals have been conducted in Finland (Hakama *et al*, 1997; Anttila *et al*, 2002), Denmark (Olsen *et al*, 2005) and Sweden (Duffy *et al*, 2002b). The Finnish studies, focusing on women aged 50–59 years, used linked screening, cancer and death registry data to examine breast cancer mortality outcomes. Since screening was implemented in Finland in different municipalities at different times, with women invited according to even or odd year of birth, such quasirandomised cohorts can be compared.

A Finnish study, published in 1997, of breast cancer mortality from cancers diagnosed only after the implementation of screening found 24% (95% CI (RR): 0.53–1.09; marginally nonsignificant)

Table 5. Prognostic indicators for breast cancers diagnosed in regular vs irregularly screened^a women, New Zealand women aged 45–69 years at year of diagnosis, 2003–2011^b

Prognostic indicator	Regularly screened		Irregularly screened		Significance of difference
	n	%	n	%	
Grade of tumour					
Well differentiated	1697	30.6	2160	29.7	$\chi^2_{(1)} = 1.38, P = 0.2403$ $\chi^2_{(1)} = 0.38, P = 0.5381$
Excluding not recorded or NA (n = 696)		32.1		31.6	
Moderately differentiated	2315	41.8	2960	40.6	
Poorly or undifferentiated or anaplastic	1277	23	1724	23.7	
Not recorded/NA	255	4.6	441	6.1	
Overall heterogeneity: $\chi^2_{(3)} = 14.7, P = 0.0021$					
Extent of disease					
Localised	3738	67.4	4333	59.5	$\chi^2_{(1)} = 85.2, P < 0.0001$ $\chi^2_{(1)} = 74.1, P < 0.0001$
Excluding unknown (n = 847)		71.6		64.1	
Adjacent organ or regional lymph node	1433	25.9	2310	31.7	
Distant	53	1	115	1.6	
Distant÷(distant + regional)		3.6		4.7	RR = 0.75 (0.55–1.03)
Unknown	320	5.8	527	7.2	$\chi^2_{(1)} = 3.10, P = 0.0784$
Overall heterogeneity: $\chi^2_{(3)} = 93, P < 0.0001$					
Multiple tumours					
No	5475	98.8	7127	97.8	RR = 0.57 (0.43–0.76) $\chi^2_{(1)} = 15.5, P < 0.0001$
Yes	69	1.2	158	2.2	
Maximum tumour size	(mm)	(mm)			
Mean	16.3	19.3			t = -12.0, P < 0.0001
Median	14	15			z = -10.5, P < 0.0001 ^c

Abbreviation: NA = not available.

^aAs screened by Breastscreen Aotearoa.

^b2003 is the earliest year that regular and irregular screening can be established.

^cTwo-sample median test.

lower breast cancer mortality in women invited to screening during 1987–1989 compared with those not invited (Hakama *et al*, 1997). For invited women aged under 56 years, breast cancer mortality was found to be 44% (95% CI: 0.33–0.95) lower (significant). This study was not affected by screening selection bias.

A later individual-based Finnish study (Anttila *et al*, 2002) of service screening in the Helsinki, published in 2002, compared birth cohorts not offered screening (born 1930–1934), with birth cohorts offered screening who were exposed to most screening rounds with the longest follow-up time (born 1935–1939). This study found 19% (95% CI: 0.62–1.05; marginally nonsignificant) lower breast cancer mortality in the screened cohort after adjusting for screening selection bias (Anttila *et al*, 2002). When the Finnish results were combined in a meta-analysis, breast cancer mortality in women invited to screen was estimated to be 23% (95% CI (RR): 0.40–1.00) lower than uninvited women (statistically significant) (Irvin and Kaplan, 2014).

A Danish cohort study of Copenhagen women aged 50–69 years with individual measurement of screening exposure for the first decade of screening (1991–2001), with up to 10 years follow-up, found 25% (95% CI: 0.63–0.89) lower breast cancer mortality in those invited for screening, compared with women not invited for screening (Olsen *et al*, 2005). The result was derived from a contemporaneous national cohort of uninvited Danish women, historical cohorts of Copenhagen women in the decade before screening and historical cohorts of Danish women. Breast cancer mortality was found to be 37% (approximate 95% CI: 0.52–0.77) lower in women who actually screened compared with those not screening, and after adjusting for screening selection bias by Gabe and Duffy (2005), it was 30% lower.

In a follow-up of mammography screening in Florence (Italy), breast cancer mortality was estimated to be 19% lower (borderline significant) in invited women diagnosed with breast cancer between 1990 and 1993 (follow-up to 1999) compared with

that in women not yet invited to screening (Paci *et al*, 2002b). However, breast cancer mortality in ‘uninvited’ women was estimated by application of case fatality rates to breast cancers expected to be diagnosed in the population not yet invited to screening.

In a Swedish study (Duffy *et al*, 2002b) published in 2002, breast cancer mortality in invited women was found to be 30% lower than in uninvited women from cancers diagnosed entirely in the screening epoch (comparable to the present study). Breast cancer mortality in women who actually screened was found to be 39% lower compared with that in women who did not screen (statistically significant), after adjusting for screening selection bias (Duffy *et al*, 2002b).

From a systematic review by Gabe and Duffy (2005), results of both aggregate and individual-based cohort studies were meta-analysed and breast cancer mortality RR was estimated as 0.74 (26% lower) for invitation to screening, after adjustment for potential confounding and bias (11 studies; Peer *et al*, 1995; Dijk *et al*, 1997; Hakama *et al*, 1997; UK Trial of Early Detection of Breast Cancer Group, 1999; Jonsson *et al*, 2000, 2001; Anttila *et al*, 2002; Paci *et al*, 2002a; Duffy *et al*, 2002b; Tabár *et al*, 2003; Olsen *et al*, 2005). For actual screening, breast cancer mortality RR of 0.57 was estimated (mortality reduction 43%; five studies; Gabe and Duffy, 2005). After adjusting further for screening selection bias, a combined 32% breast cancer mortality reduction was found associated with screening attendance (from 10 studies; Gabe and Duffy, 2005). However, it was not stated explicitly which value of RR for unscreened relative to uninvited women was used for adjusting for screening selection bias.

Since the Gabe review (2005; Gabe and Duffy, 2005), similar results were produced from an individual-based cohort study in Norway of 50–79-year-old women over 1986–2009 (Weedon-Fekjær *et al*, 2014) showing 28% lower breast cancer mortality in women invited to screening than in women not invited, after

adjusting for confounders. Another Norwegian study comparing cohorts in early- and late-starting mammography screening areas found somewhat smaller and nonsignificant differences using differing historical and/or regional controls (RR range 0.89–0.93). However, national survey data indicated that 40% had already participated in regular screening before the introduction of screening mammography in Norway (Olsen *et al.*, 2013), and 64% of first attendees reported having a mammogram before participating in the organised screening program (Lyng *et al.*, 2011; Olsen *et al.*, 2013).

A further meta-analysis of quasiexperimental aggregate studies (Irvin and Kaplan, 2014), including the Norwegian study above (Weedon-Fekjær *et al.*, 2014), with incidence-based breast cancer mortality as the outcome for women age 50–69 years, showed breast cancer mortality to be: 43% lower in historical comparisons (two studies; Swedish Organised Service Screening Evaluation Group, 2006; Ascunce *et al.*, 2007); 22% lower in geographical comparisons (two studies; UK Trial of Early Detection of Breast Cancer Group, 1999; Jonsson *et al.*, 2007); and 13% lower in geographical–historical comparisons (five studies; Jonsson *et al.*, 2001, 2003; Olsen *et al.*, 2005, 2013; Parvinen *et al.*, 2006). These meta-analysed estimates were all statistically significantly different from zero breast cancer mortality difference.

An aggregate cohort evaluation of BreastScreen Australia for 1990–2004, using small area incidence-linked breast cancer mortality, correlated with lagged mammography screening participation rates in each area and each year (Poisson regression analysis), or breast cancer mortality subsequent to screening in these areas and years (Cox proportional hazard regression), found 25–34% lower breast cancer mortality associated with screening mammography (respectively), adjusted for confounders when projected to the target screening participation of 70% (Taylor *et al.*, 2009; Morrell *et al.*, 2012).

Data linkage for BSA has allowed an unprecedented examination of the efficacy of screening mammography in a real-world population setting. While mismatches based on the NHI are rare, the extent of duplicate (i.e., different) NHIs applying to the same person across different data sources is not well established; however, records are de-duplicated when discovered (Lewis, 2016 personal communication). Accordingly, the effect on screening estimates derived in the present study may bias them away from the null. For instance, a woman with one NHI recorded on the BSA screening register and another NHI recorded on the death and cancer registers would be misclassified as ever-screened and still alive, thus favouring the ever-screened. However, the likelihood of these exceptions is very low (Lewis, 2016 personal communication).

Apart from the Finnish, Danish and Swedish studies, the authors are not aware of other cohort studies of established screening mammography that have used individual-based information on screening exposure linked with cancer diagnosis and mortality from breast cancer. Further, to our knowledge no previous individual cohort studies have taken account of changes in screening exposure during follow-up of study populations, and none have compared regular screening with less frequent screening and never screening to examine a dose–response effect.

It is probable that a small proportion of women classified as never-screened with BSA may have screened privately, or had participated in screening pilots conducted in the early 1990s and have not subsequently screened with BSA. The extent of private opportunistic screening in New Zealand outside BSA is difficult to determine because private mammography is not subsidised or provided without charge, and is paid for by the individual or their insurance company. Consequently, it would be expected that *de facto* or previous pilot screening in never-BSA screened women, who would be misclassified as never screened, would bias any mortality benefits found in screened women toward the null.

While case–control studies do not provide the strongest evidence for causation, they use only exposures and outcomes measured in individuals. For this reason, they are not subject to the same levels of misclassification of exposure that can beset aggregate studies or quasiexperimental cohort studies. An advantage of case–control studies over cohort studies is they are not affected by attrition, which can produce biased estimates of the effect if loss to follow-up is systematically associated with exposure to the study factor.

When comparing screened with unscreened women, results from case–control studies are consistent with those found in the present study. In the review by Gabe and Duffy (2005), of seven case–control studies, the estimated odds ratios (OR) of breast cancer death in screened compared with unscreened women ranged from 0.42 to 0.75. A case–control study in South Australia found OR of 0.59 (95% CI: 0.47–0.74) for screening participants compared with non-participants, which was ~0.70 when corrected for screening participation bias, similar to the present study (Roder *et al.*, 2008).

There is some disagreement over the relative roles in the population breast cancer mortality decline of organised mammography screening, compared with breast cancer treatment improvements, especially addition and/or extended use of chemotherapy and tamoxifen therapy to surgery for the treatment of primary breast cancer (Early Breast Cancer Trialists' Collaborative Group (EBCTCG), 2005; Burton *et al.*, 2012). This coincided with the advent of mammography screening programs in many Western countries around 1990. Nonetheless, service studies of mammography screening in Sweden and the Netherlands found significant mortality reductions coincident with the introduction of mammography screening that preceded widespread changes in the primary treatment of breast cancer (Tabár *et al.*, 2001, 2003; Otto *et al.*, 2003). In the case of New Zealand, the implementation of screening mammography (1999) occurred after most of these treatment improvements had become widespread (late 1980s–early 1990s), and lower breast cancer mortality has nonetheless been found to be associated with screening.

To conclude, this study has shown that screening mammography in New Zealand is associated with significantly lower breast cancer mortality for ever-screened women compared with women who have never screened with BSA; that there is evidence of a dose–response relationship between greater screening exposure and lower breast cancer mortality; and that the breast cancer mortality reductions are consistent with prognostic indicators in relation to screening exposure, and with evidence from RCTs and other service studies of screening effectiveness.

ACKNOWLEDGEMENTS

We would like to acknowledge the valuable advice given to this evaluation by Stephen Duffy, Professor of Screening, Wolfson Institute of Preventive Medicine, University of London, London, UK, as the independent peer reviewer of the evaluation study, and the contributions of Chris Lewis, New Zealand Ministry of Health, and Kerry Sexton and the late Amanda Borich from the National Screening Unit, Ministry of Health, Wellington, New Zealand.

CONFLICT OF INTEREST

Four of the authors (SM, RT, DR and BR) were commissioned by the NZ Ministry of Health to conduct a comprehensive, independent evaluation of BSA from which this paper is derived (Morrell *et al.*, 2015). Two of the authors formerly provided independent screening monitoring reports for BSA. One of the

authors (BR) currently performs this role. Other coauthors (MG and KC) are employees of NZ Ministry of Health.

REFERENCES

- Alexander FE, Anderson TJ, Brown HK, Forrest AP, Hepburn W, Kirkpatrick AE, McDonald C, Muir BB, Prescott RJ, Shepherd SM (1994) The Edinburgh randomised trial of breast cancer screening: results after 10 years of follow-up. *Br J Cancer* **70**(3): 542–548.
- Allgood PC, Warwick J, Warren RML, Day NE, Duffy SW (2008) A case-control study of the impact of the East Anglian breast screening programme on breast cancer mortality. *Br J Cancer* **98**(1): 206–209.
- Anttila A, Koskela J, Hakama M (2002) Programme sensitivity and effectiveness of mammography service screening in Helsinki, Finland. *J Med Screen* **9**(4): 153–158.
- Asuncion EN, Moreno-Iribas C, Barcos Urriaga A, Ardanaz E, Ederra Sanz M, Castilla J, Eg̃os N (2007) Changes in breast cancer mortality in Navarre (Spain) after introduction of a screening programme. *J Med Screen* **14**(1): 14–20.
- Blanks RG, Moss SM, McGahan CE, Quinn MJ, Babb PJ (2000) Effect of NHS breast screening programme on mortality from breast cancer in England and Wales, 1990–98: comparison of observed with predicted mortality. *BMJ* **321**(7262): 665–669.
- Broeders M, Moss S, Nyström L, Njor S, Jonsson H, Paap E, Massat N, Duffy S, Lyng E, Paci E, E.U.R.O.S.C.R.E.E.N Working Group (2012) The impact of mammographic screening on breast cancer mortality in Europe: a review of observational studies. *J Med Screen* **19**(Suppl 1): 14–25.
- Broeders MJ, Peer PG, Straatman H, Beex LV, Hendriks JH, Holland R, Verbeek AL (2001) Diverging breast cancer mortality rates in relation to screening? A comparison of Nijmegen to Arnhem and the Netherlands, 1969–1997. *Int J Cancer* **92**(2): 303–308.
- Burton RC, Bell RJ, Thiagarajah G, Stevenson C (2012) Adjuvant therapy, not mammographic screening, accounts for most of the observed breast cancer specific mortality reductions in Australian women since the national screening program began in 1991. *Breast Cancer Res Treat* **131**(3): 949–955.
- Dijkstra JAV, Verbeek AL, Beex LV, Hendriks JH, Holland R, Mravunac M, Straatman H, Werre JM (1997) Breast-cancer mortality in a non-randomized trial on mammographic screening in women over age 65. *Int J Cancer* **70**(2): 164–168.
- Duffy S, Cuzick J, Tabár L, Vitak B, Hsiu-Hsi Chen T, Yen M-F, Smith R (2002a) Correcting for noncompliance bias in case-control studies to evaluate cancer screening programmes. *Appl Stat* **51**: 235–243.
- Duffy SW, Tabár L, Chen H-H, Holmqvist M, Yen M-F, Abdsalah S, Epstein B, Frodis E, Ljungberg E, Hedborg-Melander C, Sundbom A, Tholin M, Wiese M, Akerlund A, Wu H-M, Tung T-S, Chiu Y-H, Chiu C-P, Huang C-C, Smith RA, Ros̃on M, Stenbeck M, Holmberg L (2002b) The impact of organized mammography service screening on breast carcinoma mortality in seven Swedish counties. *Cancer* **95**(3): 458–469.
- Duffy SW, Tabar L, Fagerberg G, Gad A, Gr̃ntoft O, South MC, Day NE (1991) Breast screening, prognostic factors and survival? results from the Swedish two county study. *Br J Cancer* **64**(6): 1133–1138.
- Duffy SW, Tabár L, Vitak B, Yen MF, Warwick J, Smith RA, Chen HH (2003) The Swedish Two-County Trial of mammographic screening: cluster randomisation and end point evaluation. *Ann Oncol* **14**(8): 1196–1198.
- Early Breast Cancer Trialists' Collaborative Group (EBCTCG) (2005) Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet (London, England)* **365**: 1687–1717.
- Freedman DA, Petitti DB, Robins JM (2004) On the efficacy of screening for breast cancer. *Int J Epidemiol* **33**(1): 43–55.
- Frisell J, Lidbrink E, Hellström L, Rutqvist LE (1997) Followup after 11 years?update of mortality results in the Stockholm mammographic screening trial. *Breast Cancer Res Treat* **45**(3): 263–270.
- Gabe R, Duffy SW (2005) Evaluation of service screening mammography in practice: the impact on breast cancer mortality. *Ann Oncol* **16**(Suppl 2): ii153–ii162.
- Göttsche P, Jørgensen K (2013) Screening for breast cancer with mammography. *Cochrane Database Syst Rev* **6**: CD001877.
- Göttsche PC, Nielsen M (2006) Screening for breast cancer with mammography. *Cochrane Database Syst Rev* **4**: CD001877.
- Göttsche PC, Nielsen M (2009) Screening for breast cancer with mammography. *Cochrane Database Syst Rev* **4**, CD001877.
- Göttsche PC, Nielsen M (2011) Screening for breast cancer with mammography. *Cochrane Database Syst Rev* **1**: CD001877.
- Göttsche PC, Olsen O (2000) Is screening for breast cancer with mammography justifiable? *Lancet* **355**(9198): 129–134.
- Hakama M, Pukkala E, Heikkil̃ M, Kallio M (1997) Effectiveness of the public health policy for breast cancer screening in Finland: population based cohort study. *BMJ* **314**(7084): 864–867.
- Hofvind S, Ursin G, Tretli S, Sebuodeg̃rd S, Møller B (2013) Breast cancer mortality in participants of the Norwegian Breast Cancer Screening Program. *Cancer* **119**(17): 3106–3112.
- International Agency for Research on Cancer (2002) *IARC Handbooks of Cancer Prevention: Breast Cancer Screening*. IARC, Lyon, France.
- International Agency for Research on Cancer (2016) *IARC Handbooks of Cancer Prevention, Volume 15, Breast Cancer Screening*. IARC, Lyon, France.
- Irvin VL, Kaplan RM (2014) Screening mammography and breast cancer mortality: meta-analysis of quasi-experimental studies. *PLoS One* **9**(6): e98105.
- Jonsson H, Bord̃s P, Wallin H, Nyström L, Lenner P (2007) Service screening with mammography in Northern Sweden: effects on breast cancer mortality—an update. *J Med Screen* **14**(2): 87–93.
- Jonsson H, Nyström L, Törnberg S, Lenner P (2001) Service screening with mammography of women aged 50–69 years in Sweden: effects on mortality from breast cancer. *J Med Screen* **8**(3): 152–160.
- Jonsson H, Törnberg S, Nyström L, Lenner P (2000) Service screening with mammography in Sweden? evaluation of effects of screening on breast cancer mortality in age group 40–49 years. *Acta Oncol* **39**(5): 617–623.
- Jonsson H, Törnberg S, Nyström L, Lenner P (2003) Service screening with mammography of women aged 70–74 years in Sweden. effects on breast cancer mortality. *Cancer Detect Prev* **27**(5): 360–369.
- Kerlikowske K, Grady D, Rubin SM, Sandrock C, Ernster VL (1995) Efficacy of screening mammography. A meta-analysis. *JAMA* **273**(2): 149–154.
- Lawrence G, Wallis M, Allgood P, Nagtegaal ID, Warwick J, Cafferty FH, Houssami N, Kearins O, Tappenden N, O'Sullivan E, Duffy SW (2009) Population estimates of survival in women with screen-detected and symptomatic breast cancer taking account of lead time and length bias. *Breast Cancer Res Treat* **116**(1): 179–185.
- Lawrenson R, Seneviratne S, Scott N, Peni T, Brown C, Campbell I (2016) Breast cancer inequities between M̃ori and non-M̃ori women in Aotearoa/New Zealand. *Eur J Cancer Care* **25**: 225–230.
- Lyng E, Braaten T, Njor SH, Olsen AH, Kumle M, Waaseth M, Lund E (2011) Mammography activity in Norway 1983 to 2008. *Acta Oncol* **50**(7): 1062–1067.
- Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M (2013) The benefits and harms of breast cancer screening: an independent review. *Br J Cancer* **108**(11): 2205–2240.
- Miller AB, Baines CJ, To T, Wall C (1992) Canadian national breast screening study: 2. breast cancer detection and death rates among women aged 50 to 59 years. *CMAJ* **147**(10): 1477–1488.
- Morrell S, Taylor R, Roder D, Dobson A (2012) Mammography screening and breast cancer mortality in Australia: an aggregate cohort study. *J Med Screen* **19**(1): 26–34.
- Morrell S, Taylor R, Roder D, Robson B (2015) Cohort and Case Control Analyses of Breast Cancer Mortality: BreastScreen Aotearoa 1999–2011. New Zealand Ministry of Health: Wellington, New Zealand. Available at: https://www.nsu.govt.nz/system/files/resources/bnszmortcohcceval_final7_8_dec.pdf.
- Nelson HD, Fu R, Cantor A, Pappas M, Daeges M, Humphrey L (2016) Effectiveness of breast cancer screening: Systematic review and meta-analysis to update the 2009 U.S. preventive services task force recommendation. *Ann Intern Med* **164**: 244–255.
- Nickson C, Mason KE, English DR, Kavanagh AM (2012) Mammographic screening and breast cancer mortality: a case-control study and meta-analysis. *Cancer Epidemiol Biomarkers Prev* **21**(9): 1479–1488.
- Nystrom L, Larsson LG, Wall S, Rutqvist LE, Andersson I, Bjurstram N, Fagerberg G, Frisell J, Tabar L (1996) An overview of the Swedish randomised mammography trials: total mortality pattern and the representivity of the study cohorts. *J Med Screen* **3**(2): 85–87.
- Olsen AH, Lyng E, Njor SH, Kumle M, Waaseth M, Braaten T, Lund E (2013) Breast cancer mortality in Norway after the introduction of mammography screening. *Int J Cancer* **132**(1): 208–214.

- Olsen AH, Njor SH, Vejborg I, Schwartz W, Dalgaard P, Jensen M-B, Tange UB, Blichert-Toft M, Rank F, Mouridsen H, Lynge E (2005) Breast cancer mortality in Copenhagen after introduction of mammography screening: cohort study. *BMJ* **330**(7485): 220.
- Otto SJ, Fracheboud J, Looman CWN, Broeders MJM, Boer R, Hendriks JHCL, Verbeek ALM, de Koning HJ, National Evaluation Team for Breast Cancer Screening (2003) Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systematic review. *Lancet* **361**(9367): 1411–1417.
- Otto SJ, Fracheboud J, Verbeek ALM, Boer R, Reijerink-Verheij JCIY, Otten JDM, Broeders MJM, de Koning HJ. National Evaluation Team for Breast Cancer Screening (2012) Mammography screening and risk of breast cancer death: a population-based case-control study. *Cancer Epidemiol Biomarkers Prev* **21**(1): 66–73.
- Paap E, Verbeek ALM, Botterweck AAM, van Doorne-Nagtegaal HJ, Imhof-Tas M, de Koning HJ, Otto SJ, de Munck L, van der Steen A, Holland R, den Heeten GJ, Broeders MJM (2014) Breast cancer screening halves the risk of breast cancer death: a case-referent study. *Breast* **23**(4): 439–444.
- Paci E, Duffy SW, Giorgi D, Zappa M, Crocetti E, Vezzosi V, Bianchi S, del Turco MR (2002a) Quantification of the effect of mammographic screening on fatal breast cancers: The Florence Programme 1990–96. *Br J Cancer* **87**(1): 65–69.
- Paci E, Giorgi D, Bianchi S, Vezzosi V, Zappa M, Crocetti E, Rosselli del Turco M (2002b) Assessment of the early impact of the population-based breast cancer screening programme in Florence (Italy) using mortality and surrogate measures. *Eur J Cancer* **38**(4): 568–573.
- Page A, Morrell S, Taylor R (2014) *Screening and assessment report of women attending BreastScreen Aotearoa: Women screened July 2010–June 2012*. BreastScreen Aotearoa, Wellington. Independent Monitoring Group Report.
- Parvinen I, Helenius H, Pylkkönen L, Anttila A, Immonen-Röihö P, Kauhava L, Röihö O, Klemi PJ (2006) Service screening mammography reduces breast cancer mortality among elderly women in Turku. *J Med Screen* **13**(1): 34–40.
- Peer PG, Werre JM, Mravunac M, Hendriks JH, Holland R, Verbeek AL (1995) Effect on breast cancer mortality of biennial mammographic screening of women under age 50. *Int J Cancer* **60**(6): 808–811.
- Puliti D, Miccinesi G, Collina N, De Lisi V, Federico M, Ferretti S, Finarelli AC, Foca F, Mangone L, Naldoni C, Petrella M, Ponti A, Segnan N, Sigona A, Zarcone M, Zorzi M, Zappa M, Paci E, I.M.P.A.C.T Working Group (2008) Effectiveness of service screening: a case-control study to assess breast cancer mortality reduction. *Br J Cancer* **99**(3): 423–427.
- Quinn M, Allen E (1995) Changes in incidence of and mortality from breast cancer in England and Wales since introduction of screening. United Kingdom Association of Cancer Registries. *BMJ* **311**(7017): 1391–1395.
- Roder D, Houssami N, Farshid G, Gill G, Luke C, Downey P, Beckmann K, Iosifidis P, Grieve L, Williamson L (2008) Population screening and intensity of screening are associated with reduced breast cancer mortality: evidence of efficacy of mammography screening in Australia. *Breast Cancer Res Treat* **108**(3): 409–416.
- Seneviratne S, Campbell I, Scott N, Coles C, Lawrenson R (2015) Treatment delay for Māori women with breast cancer in New Zealand. *Ethn Health* **20**: 178–193.
- Shapiro S (1997) Periodic screening for breast cancer: the HIP Randomized Controlled Trial. Health Insurance Plan. *J Natl Cancer Inst Monogr* **22**(22): 27–30.
- Swedish Organised Service Screening Evaluation Group (2006) Reduction in breast cancer mortality from organized service screening with mammography: 1. further confirmation with extended data. *Cancer Epidemiol Biomarkers Prev* **15**(1): 45–51.
- Tabár L, Fagerberg G, Duffy SW, Day NE, Gad A, Grøntoft O (1992) Update of the Swedish two-county program of mammographic screening for breast cancer. *Radiol Clin N Am* **30**(1): 187–210.
- Tabár L, Vitak B, Chen HH, Yen MF, Duffy SW, Smith RA (2001) Beyond randomized controlled trials: organized mammographic screening substantially reduces breast carcinoma mortality. *Cancer* **91**(9): 1724–1731.
- Tabár L, Yen M-F, Vitak B, Chen HHT, Smith RA, Duffy SW (2003) Mammography service screening and mortality in breast cancer patients: 20-year follow-up before and after introduction of screening. *Lancet* **361**(9367): 1405–1410.
- Taylor R, Morrell S, Estoesta J, Brassil A (2004) Mammography screening and breast cancer mortality in New South Wales, Australia. *Cancer Causes Control* **15**(6): 543–550.
- Taylor R, Morrell S, Roder D (2009) Screening Monograph No. 4/2009 ? Mortality (Ecological) Study. Commonwealth Department of Health and Aging (Australia). Available at: http://s3.amazonaws.com/zanran_storage/www.cancerscreening.gov.au/ContentPages/999890649.pdf?last (accessed 26 August 2016).
- Threlfall AG, Collins S, Woodman CBJ (2003) Impact of NHS breast screening on advanced disease and mortality from breast cancer in the North West of England. *Br J Cancer* **89**(1): 77–80.
- UK Trial of Early Detection of Breast Cancer Group (1999) 16-year mortality from breast cancer in the UK Trial of Early Detection of Breast Cancer. *Lancet* **353**(9168): 1909–1914.
- van der Waal D, Broeders MJM, Verbeek ALM, Duffy SW, Moss SM (2015) Case-control studies on the effectiveness of breast cancer screening: Insights from the UK Age Trial. *Epidemiology* **26**(4): 590–596.
- Weedon-Fekjær H, Romundstad PR, Vatten LJ (2014) Modern mammography screening and breast cancer mortality: population study. *BMJ* **348**: g3701.



This work is licensed under the Creative Commons Attribution-Non-Commercial-Share Alike 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/4.0/>

© The Author(s) named above 2017