

# A controlled study of income development for breast cancer survivors in Norway

Jūratė Šaltytė Benth · Fredrik A. Dahl · Hilde Lurås · Alv A. Dahl

Received: 26 June 2013 / Accepted: 23 October 2013 / Published online: 19 December 2013  
© Springer Science+Business Media New York 2013

## Abstract

**Aims** This study seeks to assess the impact of breast cancer on survivors' annual income at 1 to 13 years of follow-up.

**Methods** A Norwegian registry-based dataset with a 13-year follow-up period was used. The dataset contained case–control pairs, where each pair consisted of one breast cancer case and a cancer-free control, matched for age, marital status, and municipality of residence. A mixed linear model was used to analyse the average income development for cases and controls adjusting for available demographic variables not used in the matching procedure, such as education level, immigration history, and number of children. Individual income was only considered relevant for cancer survivors, thus case–control pairs were censored upon death. Income development by stage of tumour at time of diagnosis was also assessed.

**Results** The income of breast cancer survivors had reduced immediately following diagnosis. At 1 year after diagnosis, income development between cases and controls became significantly different ( $p=0.006$ ). Differences increased slightly but remained significant throughout the follow-up period. At 13 years after diagnosis, the estimated cumulative income loss for survivors was 39,403 € ( $p<0.001$ ). The income development of stage I breast cancer patients was similar to their controls. For higher stage breast cancer patients, the income

differences were more pronounced but not always statistically significant.

**Conclusions** Breast cancer has a significant negative effect on the individual incomes of survivors, even at 13 years after diagnosis. Effects increased for individuals with more advanced cancer stages.

**Implications for Cancer Survivors** In Norway, breast cancer survivors experience lower income than their cancer-free controls. This may indicate a need for financial compensations in order to maintain the standard of living for this group of patients.

**Keywords** Breast cancer · Income development · Case–control pairs

## Introduction

According to the Cancer Registry of Norway, a cumulative risk of developing breast cancer by age of 75 is 7.9, and ranks highest among all cancer diagnoses in women [1]. Among all diagnoses of cancer in women, breast cancer is diagnosed in 34 % and 28 % of women in age groups 25–49 and 50–69 years, respectively. Considering the age span, many of these women are likely to be employed when malignancy determined. Moreover, since most women suffer an early stage of breast cancer, which considerably increases the survival odds, the issue of income development for female breast cancer survivors is highly relevant.

There is a substantial international literature on income development for breast cancer survivors, but since income is highly contingent upon local labour markets, we have restricted our attention to the Scandinavian context. An increased risk of experiencing income reduction was shown in one Danish [2] and three Norwegian studies [3–5]. At the same time, a Swedish study [6] found no impact of cancer on the incomes

J. Šaltytė Benth · H. Lurås  
Institute of Clinical Medicine, Campus Ahus, University of Oslo,  
Oslo, Norway

J. Šaltytė Benth (✉) · F. A. Dahl · H. Lurås  
HØKH, Research Centre, Akershus University Hospital,  
1478 Lørenskog, Norway  
e-mail: jurates@medisin.uio.no

A. A. Dahl  
National Resource Center for Late Effects, Department of Oncology,  
Oslo University Hospital, Radiumhospitalet and University of Oslo,  
Oslo, Norway

of breast cancer patients compared to cancer-free controls. This study thus aimed to assess the differences in average annual income development of breast cancer patients and matched cancer-free controls in Norway. We hypothesised that breast cancer patients would show persistently significant reduction in annual income development compared to controls.

Though a number of other studies find significant income differences, this paper adds new insight to related research in several ways. First, this study was based on high-quality merged data from two nationwide registries. Annual income development of breast cancer patients was tracked from 1 year prior to diagnosis until death or censoring nine to 13 years later, which constitutes a considerably longer follow-up period than most studies in the field. Secondly, our study resembles studies [4] and [5] but differs in key aspects. To assess income differences between cancer survivors and cancer-free controls, the difference-in-differences approach was employed in [4]. Statistically significant cancer-associated declines in earnings, not including compensatory measures, were reported for up to 9 years after diagnosis. Unlike the present study's analysis, the analysis in [4] was restricted to women who were employed in the follow-up year, thus excluding the financial disadvantage of being completely unemployed. Meanwhile, the analysis in [5] used a subset of the same case-control dataset that we analyse in the present study, only the analysis in [5] was restricted to women who did not receive either disability- or retirement pension while the follow-up time was limited to 9 years. The study [5] analysed the probability of an income reduction of more than 10 % and found a statistically significant effect of breast cancer for only 2 to 4 years following diagnosis. The present study is more sophisticated statistically, for it utilises the longitudinal structure of the data. Furthermore, since income is modelled as a continuous variable, our study avoids the information loss caused by dichotomisation. The statistical methods developed for this study can thus be considered to be an independent contribution to the field.

## Material and methods

### Data sources

This study was based on data from two nationwide Norwegian registries whose merging was possible due to the unique personal identity numbers given to all Norwegian citizens. The Cancer Registry of Norway (CRN, established 1953) has collected data on all new cases of cancer as per statutory registration requirements for all hospitals, laboratories, and general practitioners. Diagnostic data from the CRN are nearly complete and of high quality [7]. FD-Trygd, Statistics Norway's event database (established 1992) contains information regarding demographics, social security benefits from

the National Insurance Scheme, education level, employment status, and income for the entire Norwegian population.

### Sampling of patients and cancer-free controls

Patient sampling for the study started in 1992, the first year when dynamic data on national insurance contributions became available. Eligibility criteria for identification of breast cancer patients in the CRN were women aged 45 to 54 who had been diagnosed with breast cancer as their first life-time malignancy between 1992 and 1996. This 5-year recruitment period was determined in order to obtain a reasonable follow-up time, while the age span assured a high proportion of employed women, since the formal retirement age in Norway is 67 years. The CRN identified 2,052 women who fulfilled these criteria.

Cancer-free controls at inclusion and during follow-up matched for age, marital status, and municipality of residence were drawn randomly from the Norwegian population registry by Statistics Norway, one control per case. The controls were cancer-free, based on checks in the CRN. Throughout this article, the term 'year of diagnosis' also refers to controls, as well as primarily referring to the year of breast cancer diagnosis for their matched case.

Among the 4,104 women in the sample, 14 did not have a registered income and 22 had missing income data for some years during the study period. These 36 women were excluded from the analyses along with their matched controls. Six controls died before the time of cancer diagnosis and were also excluded along with their matched cases. Thus, our sample consisted of 4,020 women: 2,010 breast cancer patients and 2,010 matched controls. By comparison, the cohorts of 1992 to 1996 contained 342 (17.0 %), 362 (18.0), 373 (18.6), 436 (21.7), and 497 (24.7) pairs, respectively.

### Time to events

For each case and control, observation time began the pre-diagnosis year. All women were followed until death or censoring at the end of the observation period, 31 December 2005. Hence, the maximum follow-up time varied between nine and 13 years depending on the year of diagnosis (i.e. 1992–1996).

### Income

The income variable 'pension qualifying income', in this study, is referred to simply as income. It is defined as personal income excluding transfers of social security benefits, such as sick-leave pay, unemployment benefits, work assessment allowance, and disability pension. The income variable was adjusted for the consumer price index (CPI) for 2012.

## Socio-demographic variables and stages

Age was used as a continuous variable. Level of education was coded as ‘elementary school’, ‘secondary school’ or ‘higher education’. Immigration history was divided into either ‘no immigration history’ or ‘immigration history’. Family type was coded into three categories: ‘married or in-live companion with children’, ‘married or in-live companion without children’ and ‘other’, while number of children aged less than 18 at the time of inclusion was categorised into ‘zero’, ‘one’, ‘two’, and ‘three or more’. The socio-demographic variables at year of diagnosis were employed throughout the analyses.

Staging of breast cancer has been described previously in [5]. The Classification of Malignant Tumours (TNM) system classifies cancer into four stages (i.e. I to IV), each characterised by tumour size (T), involvement of regional lymph nodes (N), and distant metastasis (M).

## Constructing the income variable

To study how cancer affects income, the following computational steps were performed for each woman in the dataset. To align income data, year of diagnosis was set at year 0 for all included women. The income year before diagnosis (year –1) was used as reference income. For years 1 to 13, income development was defined as that year’s income minus the reference income. Hence, a woman’s income development for a certain year represents the change (i.e. increase or decrease) in income from 1 year pre-diagnosis to the specific post-diagnosis year. Women who died within the given follow-up year were excluded from the analysis of that and sub-subsequent years. To preserve the comparability achieved by the matching procedure, data for the pair were excluded during the year when one of the women died, as well as all subsequent years.

## Statistical analyses

Descriptive statistics were used to present the socio-demographic variables; a *t* test for independent samples and a  $\chi^2$  test were used to compare cases and controls at inclusion.

Differences in income development between cases and controls were assessed by *t* test for independent samples, separately for each year during follow-up. Such an unadjusted comparison of income development suggests a clear interpretation and shows how the income of two groups develops over time.

A linear regression model was estimated to assess the impact of cancer diagnosis on income, adjusted for the available socio-demographic variables that the matching procedure had not utilised (level of education, immigration history, and number of children <18 years). Each data record was based on both a case–control pair and a given number of follow-up

years during which both members of each pair were alive. The dependent variable was difference in income development between a case and her control. The main predictors were dummy variables encoding the number of years since diagnosis, which offered 14 indicator variables associated with years 0 to 13 after diagnosis by using the pre-diagnosis year (–1) as a reference category. To adjust for a possibly non-linear effect of age on income, the pair’s age was encoded using dummy variables representing all possible age values in the study period by using the youngest women (aged 44 in the pre-diagnosis year) as reference. The socio-demographic variables at baseline for each case and control were included as confounding variables.

Income data demonstrated strong clustering, in the sense that differences in income development over time tend to be similar for a given pair. Thus, the standard linear regression assumption of independent observations was violated. Therefore, a linear mixed model accounting for cluster effect in data was developed with a random effect for pairs (see [Appendix](#)). Regression coefficients associated with each year during follow-up were presented with the corresponding 95 % confidence intervals (CI). These coefficients should be interpreted as the effect that different numbers of follow-up years has on the difference between average income development for cases and controls. Results were presented graphically by plotting the regression coefficients for each dummy variable with the corresponding 95 % CI.

Identical analyses stratified by stage of breast cancer were also executed. Also estimating the cumulative loss in income at different follow-up years is relevant. This cannot be reliably estimated by adding the income loss estimates for the different follow-up years, as these estimates are related to different subsamples. For example, the income development in the first follow-up year might be more favourable among women that were alive eight years after diagnosis compared to all women alive one year after diagnosis. A different version of the regression model was therefore estimated, where the dependent variable was the difference in cumulative income development between alive cases and controls. The main predictors of this model were interpreted as the average total loss in income since cancer diagnosis, for the different number of years survived.

All analyses were performed with SPSS version 20 and SAS version 9.2. The level of significance was set at 5 %. All tests were two-sided.

## Ethics

Data produced by national administrative agencies registering the entire population are transferred to Statistics Norway according to the Statistics Act in relation to official statistics. Individuals cannot refuse to have their personal data included. Privacy concerns are addressed by the Personal Data Act that

sets forth procedures for handling and processing of such data. No personal informed consent from the individuals constituting the sample of this study is thus needed according to Norwegian legislation. Data were delivered in a de-identified form from Statistics Norway. The study was approved by the Regional Committee for Health Research of South-eastern Norway.

## Results

Of a total 4,020 women, 638 (592 patients and 46 controls) had died at the time of censoring (i.e. 31 December 2005). Table 1 presents descriptive statistics. There were statistically significantly more controls than cases with education at the lowest level ( $p=0.001$ ), with immigration history ( $p<0.001$ ), and with three or more children ( $p=0.025$ ). There were no significant differences in age or marital status. As expected, due to cancer-related mortality, the proportions alive for 1, 5,

10, and 13 years after diagnosis were higher among controls. Most cases (91.9 %) were diagnosed at stage I or II (see Table 1).

At inclusion, the mean CPI adjusted income was significantly higher in cases than controls (14,563 € with standard deviation (SD) 9,993 € versus 13,344 € with SD 10,123 €;  $p<0.001$ ). The differences in average CPI adjusted income for cases and controls (see Table 2 columns 3 and 4) were small and remained quite stable throughout the follow-up period. Table 2 presents the average income development for cases and controls as well as differences between the two (columns 5–7) for each year during follow-up. Figure 1a illustrates the income development for cases and controls. The income for cases was rapidly reduced immediately after diagnosis, and 1 year after diagnosis the income between cases and controls became significantly different ( $p<0.001$ ) according to the independent samples  $t$  test (see Table 2, column 8). This pattern remained persistent for the rest of the follow-up period. In Fig. 1b, the results of the regression analysis are plotted

**Table 1** Socio-demographic characteristics of women with breast cancer and matched controls

Characteristic	Cases		Controls		$p$ value
	$N$	%	$N$	%	
Age (years), mean (sdv)	2010	49.8 (2.8)	2010	49.8 (2.8)	0.991
<i>Level of education</i>	1971		1929		
Elementary school	558	28.3	640	33.2	0.001
Secondary school	979	49.7	939	48.7	
Higher	434	22.0	350	18.1	
<i>Immigration history</i>	2010		2009		
No immigration history	1884	93.7	1775	88.4	<0.001
Immigration history	126	6.3	234	11.6	
<i>Marital status</i>	1990		2009		
Married/live-in partner with children	1189	59.8	1246	62.0	0.335
Married/live-in partner without children	279	14.0	263	13.1	
Other	522	26.2	500	24.9	
<i>Children &lt;18 years</i>	1990		2009		
0	1168	58.7	1181	58.8	0.025
1	515	25.9	512	25.5	
2	257	12.9	233	11.6	
3 or more	50	2.5	83	4.1	
<i>Tumour stage at diagnosis</i>					
Stage I (T1N0M0)	972	48.4			
Stage II (T2N0M0; T0-2N1M0)	875	43.5			
Stage III (T0-2N2-3M0; T3-4N0-3M0)	48	2.4			
Stage IV (T1-4N0-3M1)	105	5.2			
Stage unknown	10	0.5			
<i>Survival</i>					
1 year after inclusion	1917	95.4	2007	99.9	
5 years after inclusion	1631	81.1	1994	99.2	
10 years after inclusion <sup>a</sup>	1188	78.5	1489	98.4	
13 years after inclusion <sup>b</sup>	215	62.9	334	97.7	

<sup>a</sup> Among those diagnosed between 1992 and 1995 (1,188 pairs)

<sup>b</sup> Among those diagnosed in 1992 (342 pairs).

Stages: T0, no evidence of primary tumour; Tumour size (T): T1,  $\leq 20$  mm; T2,  $>20$  mm but  $\leq 50$  mm; T3,  $>50$  mm; T4, tumour of any size with direct extension to the chest wall and/or skin; regional lymph nodes (N): N0, no N metastases; N1, movable ipsilateral axillary lymph node(s) metastases; N2, metastases in ipsilateral axillary lymph node(s) attached to each other or other structures; N3, metastases in ipsilateral internal mamma node(s); distant metastasis (M): M0, no presence of distant metastasis; M1, presence of distant metastasis

**Table 2** Means and standard deviations (SD) of CPI adjusted income by the year after diagnosis, and the impact that breast cancer has on a survivors' income (and cumulative income) development, adjusted for level of education, immigration history, and number of children <18 years

Year after diagnosis	N <sup>a</sup>	Mean (SD) CPI adjusted income (€)		Differences in income development <sup>b</sup> (€)		Mean difference (SE)	p value <sup>c</sup>	Coefficient <sup>c</sup> (95 % CI)	p value <sup>c</sup>	Coefficient <sup>c</sup> (95 % CI)	p value <sup>c</sup>		
		Differences in income development <sup>b</sup> (€)		Mean difference (SE)	p value <sup>c</sup>							Coefficient <sup>c</sup> (95 % CI)	p value <sup>c</sup>
		Case	Control										
-1	2010	14,563 (9,993)	13,344 (10,123)	0	0	-	-	-	-	0	-		
0	1984	14,757 (10,384)	13,534 (10,609)	236 (6,491)	284 (6,799)	-48 (211)	0.820	28 (-906; 963)	0.953	-629 (-5,384; 4,127)	0.796		
1	1914	13,987 (10,571)	13,507 (10,347)	-1,307 (9,782)	232 (7,807)	-1,539 (286)	<0.001	1,500 (438; 2,561)	0.006	429 (-4,991; 5,850)	0.877		
2	1838	13,274 (11,086)	13,807 (10,786)	-2,669 (12,840)	548 (9,623)	-3,216 (375)	<0.001	3,247 (2,021; 4,474)	<0.001	3,249 (-3,038; 9,535)	0.311		
3	1746	13,708 (11,959)	14,090 (11,073)	-1,853 (14,726)	1,111 (11,366)	-2,964 (446)	<0.001	3,071 (1,651; 4,491)	<0.001	6,066 (-1,233; 13,366)	0.103		
4	1675	13,646 (12,339)	14,278 (11,323)	-2,017 (16,124)	1,405 (12,092)	-3,422 (493)	<0.001	3,673 (2,042; 5,303)	<0.001	9,503 (1,104; 17,902)	0.027		
5	1619	13,452 (11,998)	14,339 (11,865)	-2,446 (16,540)	1,479 (13,388)	-3,925 (530)	<0.001	4,293 (2,441; 6,146)	<0.001	13,673 (4,115; 23,230)	0.005		
6	1564	13,552 (12,818)	14,367 (12,412)	-2,304 (18,346)	1,529 (14,838)	-3,833 (598)	<0.001	4,280 (2,197; 6,364)	<0.001	17,795 (7,035; 28,554)	0.001		
7	1510	13,466 (12,821)	14,265 (12,706)	-2,347 (18,861)	1,364 (16,060)	-3,711 (639)	<0.001	4,340 (2,019; 6,662)	<0.001	21,893 (9,891; 33,893)	<0.001		
8	1476	13,173 (12,852)	14,078 (12,793)	-3,064 (19,495)	830 (17,198)	-3,894 (678)	<0.001	4,612 (2,048; 7,175)	<0.001	26,204 (12,944; 39,463)	<0.001		
9	1443	12,845 (13,409)	13,737 (13,219)	-3,590 (20,563)	75 (18,318)	-3,665 (726)	<0.001	4,544 (1,734; 7,354)	0.002	29,591 (15,049; 44,132)	<0.001		
10	1042	12,066 (13,123)	13,138 (13,451)	-4,532 (20,955)	-970 (19,790)	-3,562 (895)	<0.001	4,985 (1,897; 8,074)	0.002	32,106 (16,122; 48,090)	<0.001		
11	717	11,555 (13,121)	12,062 (13,009)	-4,932 (20,722)	-2,621 (20,548)	-2,311 (1,092)	0.034	4,561 (1,172; 7,951)	0.008	29,747 (12,210; 47,284)	<0.001		
12	449	10,898 (13,144)	11,731 (13,213)	-6,612 (21,736)	-3,583 (20,574)	-3,029 (1,416)	0.033	5,691 (1,957; 9,425)	0.003	34,350 (15,041; 53,658)	<0.001		
13	211	10,714 (13,839)	10,508 (13,311)	-7,270 (22,882)	-4,667 (20,772)	-2,604 (2,133)	0.223	4,999 (746; 9,253)	0.021	39,403 (17,447; 61,360)	<0.001		

<sup>a</sup>Number of pairs in the analysis

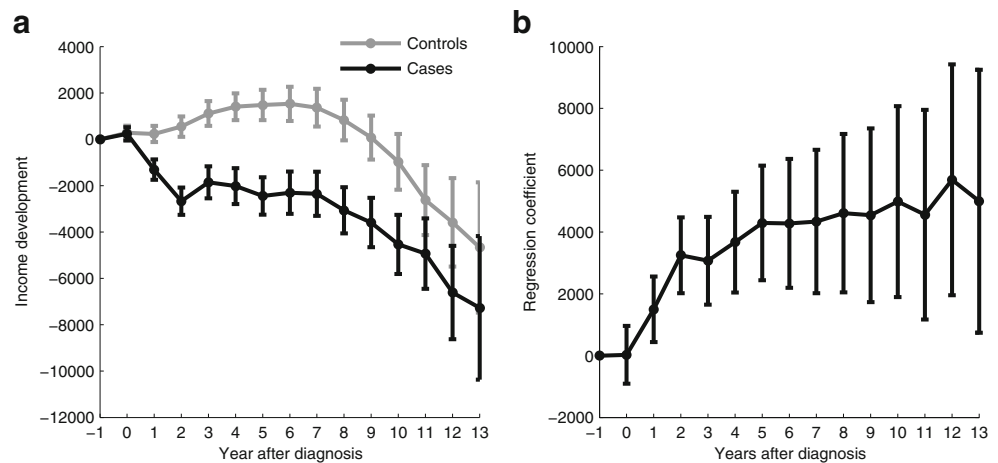
<sup>b</sup>Differences in income development between cases and controls separately for each year in the follow-up

<sup>c</sup>p value from t test for independent samples

<sup>d</sup>Differences in cumulative income development between cases and controls separately for each year in the follow-up

<sup>e</sup>Results of mixed linear model analysis

**Fig. 1** Development of income for cases (solid black line) and controls (solid grey line) after diagnosis with the corresponding 95 % CI (a) and differences in development of income adjusted for level of education, immigration history, and number of children presented as regression coefficients with the corresponding 95 % CI (b);  $-1$  corresponds to the year before diagnosis



along follow-up years together with the corresponding 95 % CI. The regression coefficients associated to the indicator variable are to be interpreted as the effect that the different number of follow-up years has on the difference between average income development for both cases and controls. Clearly, the income development adjusted for the socio-demographic characteristics showed the same tendency (see Fig. 1a). Table 2 presents numerical results from adjusted regression analysis (columns 9 and 10).

As shown in Table 2 (columns 11 and 12), at the fourth year after diagnosis, the difference in cumulative income development became significantly different between cases and controls and preserved this tendency throughout the follow-up period.

Analysis of income development by stage demonstrated marginal differences between stage I cases and their matched controls. With increasing stage, differences became more pronounced though not always statistically significant. There were few stage III and IV cases in the sample, and though large, the differences were not significant (see Table 3 and Fig. 2 for results).

## Discussion

The hypothesis that cases with breast cancer would show a persistent and significant reduction of annual income compared to controls was supported throughout the follow-up period of 13 years. At 1 year after diagnosis, the income level for cases had already become significantly lower than for controls, and this difference persisted for the remainder of the observation period. Even for those relatively few case–control pairs ( $n=211$ ) alive 13 years after diagnosis, a statistically significant effect of cancer diagnosis on their income development occurred. The effect of cancer on the cumulative income showed a stronger effect for longer follow-up periods, with an estimated total income loss of 39,403 € after 13 years

( $p<0.001$ ). Analyses of income development stratified by stage showed a similar tendency, with stronger effects for more advanced stages. Income development among stage I cases was similar to controls, and no statistically significant differences were found. For stage II to IV cases, differences were large but not always statistically significant due to few cases with more advanced stages of disease.

It is important to note that the income variable used in this study does not account for transfers of social security benefits, such as sick-leave pay, unemployment benefits, work assessment allowance, or disability pension. In Norway, one can be on sick leave for up to 1 year after diagnosis and still receive up to 100 % of the salary. As a basis to calculate basic benefits, supplementary pension, and supplement for spouse and children, six times the basic amount of the National Insurance Scheme is used. Only persons with income below such an amount are fully compensated by sickness benefit. Consequently, the economic hardship for cancer patients usually begins the year after the 1 year of receiving sick-leave pay when their opportunities for economic compensation are considerably reduced. The analysis in this paper therefore might overestimate the effect that breast cancer has on the affected women's economic situation, particularly the first year after diagnosis. Nevertheless, findings are mostly in accordance with comparable studies [4, 5]. Even with simpler statistical analyses, income differences between cancer survivors and cancer-free controls for up to 9 years after diagnosis were detected in [4]. In study [5], while assessing the probability of an income reduction of more than 10 %, statistically significant effects of breast cancer on income were shown only 2 to 4 years after diagnosis. The discrepancy between [5] and the present study, however, is not surprising since there is considerable information loss due to dichotomisation in [5].

The content of the income variable was changed from 1 January 2002, by also including work-rehabilitation benefits. However, since this study assessed income development between two comparable groups, modifying the income variable

**Table 3** The analysis of the impact breast cancer has on a survivors' income development, adjusted for level of education, immigration history, and number of children <18 years stratified by stage

Year after diagnosis	Stage I cases vs matched controls			Stage II cases vs matched controls			Stage III+IV cases vs matched controls				
	N <sup>a</sup>	Mean difference (SE) <sup>b</sup>	Coefficient <sup>c</sup> (95 % CI) p <sup>c</sup>	N <sup>a</sup>	Mean difference (SE) <sup>b</sup>	Coefficient <sup>c</sup> (95 % CI) p <sup>c</sup>	N <sup>a</sup>	Mean difference (SE) <sup>b</sup>	Coefficient <sup>c</sup> (95 % CI) p <sup>c</sup>		
-1	972	0	-	875	0	-	153	0	-		
0	968	159 (294)	-281 (-1,680;1,118)	0.694	871	-338 (342)	0.598	138	319 (580)	-225 (-3,524;3,075)	0.894
1	958	-1,091 (415) <sup>d</sup>	808 (-1,680;1,118)	0.322	842	-1,668 (423) <sup>d</sup>	0.024	107	-4,573 (1,185) <sup>d</sup>	4,629 (626;8,633)	0.024
2	942	-2,549 (513) <sup>d</sup>	2,027 (-791;2,394)	0.034	800	-3,532 (587) <sup>d</sup>	<0.001	89	-7,561 (1,566) <sup>d</sup>	7,859 (3,232;12,486)	<0.001
3	921	-2,208 (641) <sup>d</sup>	1,767 (-419;3,952)	0.113	745	-3,396 (648) <sup>d</sup>	<0.001	73	-8,047 (2,118) <sup>d</sup>	7,431 (2,070;12,793)	0.007
4	901	-2,795 (702) <sup>d</sup>	2,430 (-101;4,961)	0.060	702	-3,807 (722) <sup>d</sup>	<0.001	65	-7,787 (2,478) <sup>d</sup>	6,604 (480;12,728)	0.035
5	884	-3,425 (731) <sup>d</sup>	3,222 (328;6,117)	0.029	671	-4,507 (792) <sup>d</sup>	<0.001	57	-5,435 (3,427)	3,316 (-3,694;10,326)	0.353
6	871	-3,230 (810) <sup>d</sup>	2,964 (-306;6,235)	0.076	638	-4,428 (910) <sup>d</sup>	<0.001	49	-7,199 (4,161)	4,669 (-3,266;12,604)	0.248
7	853	-3,564 (850) <sup>d</sup>	3,413 (-244;7,070)	0.067	607	-3,784 (1,009) <sup>d</sup>	0.005	44	-5,867 (3,996)	3,346 (-5,522;12,214)	0.458
8	839	-3,666 (909) <sup>d</sup>	3,521 (-528;7,571)	0.088	588	-4,104 (1,060) <sup>d</sup>	0.005	43	-6,287 (4,226)	3,047 (-6,694;12,788)	0.538
9	829	-3,491 (962) <sup>d</sup>	3,297 (-1,150;7,743)	0.146	566	-3,660 (1,162) <sup>d</sup>	0.012	42	-7,952 (4,093)	5,488 (-5,168;16,144)	0.311
10	591	-4,373 (1,188) <sup>d</sup>	4,256 (-628;9,140)	0.088	414	-2,283 (1,436)	0.030	32	-5,796 (4,797)	2,096 (-9,724;13,916)	0.727
11	417	-2,976 (1,464)	3,536 (-1,807;8,878)	0.194	272	-1,776 (1,728)	0.040	24	1,778 (5,923)	-1,773 (-14,825;11,279)	0.789
12	263	-3,477 (1,897)	4,697 (-1,156;10,550)	0.116	167	-2,180 (2,269)	0.067	16	-7,788 (6,966)	13,335 (-1,415;28,085)	0.076
13	123	-2,061 (2,799)	3,098 (-3,480;9,675)	0.356	76	-4,345 (3,329)	0.073	11	2,727 (13,749)	7,297 (-9,353;23,906)	0.388

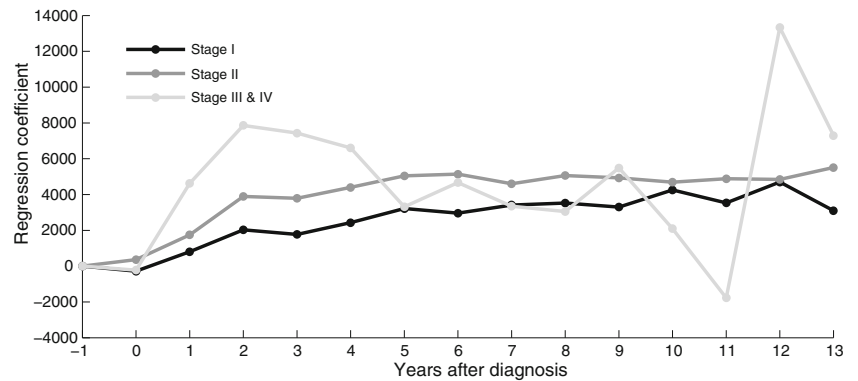
<sup>a</sup> Number of pairs in the analysis

<sup>b</sup> Mean differences (SE) in income development between cases and controls

<sup>c</sup> Results of mixed linear model analysis

<sup>d</sup> Results significant at 5 % level

**Fig. 2** Differences in development of income adjusted for level of education, immigration history, and number of children presented as regression coefficients with the corresponding 95 % CI (B);  $-1$  corresponds to the year before diagnosis



for the last 3 years of the study most likely had a negligible effect on results. Moreover, analysis of yearly age-adjusted income development in controls showed no substantive change between 2001 and 2002.

Level of education and number of children has been reported to be associated with breast cancer [3], and these variables are statistically different for cases and controls in the dataset. Regression modelling is therefore warranted to control for these effects. It should be noted, however, that these effects are relatively small, thus a crude comparison of income development in breast cancer–control cohorts matched on age, marital status, and municipality of residence may be justified in future studies when further socio-demographic variables are unavailable.

In conclusion, breast cancer has a significant impact on survivors’ incomes, even 13 years after diagnosis. The completeness of data from the CRN, which provides the opportunity to find matched cancer-free controls in registries of Statistics Norway, is this study’s chief strength. Moreover, an extended follow-up period of 9 to 13 years of income data contributes substantially to the field. Finally, the statistical method applied in this study fully explores the longitudinal nature of the data.

Our findings suggest that breast cancer survivors may require financial compensation in order to maintain the same standard of living as their cancer-free controls. Such compensation could take the form of a government-financed program or private insurance. In either case, this study offers a starting point for estimating the associated costs in the Norwegian context.

**Declaration of conflicting interests** The authors declare that there is no conflict of interest.

**Appendix A: Mathematical formulation of the model**

The data were organized in pairs of matched cases and controls. The index  $i$  was used to identify such a pair.

Define  $I_{i,j}$  as the income for the case of pair  $i, j$  years after diagnosis, where  $j \in \{-1, 0, 1, 2, \dots\}$ . The income development  $j$  years after diagnosis was calculated as  $\Delta_{i,j} = I_{i,j} - I_{i,-1}$ . Similarly, define  $\hat{I}_{i,j}$  as the income of the cancer-free control of pair  $i, j$  years after the diagnosis of the corresponding case, and let  $\hat{\Delta}_{i,j} = \hat{I}_{i,j} - \hat{I}_{i,-1}$  be income development.

Define  $Y_{i,j} = \Delta_{i,j} - \hat{\Delta}_{i,j}$ , the difference in income development  $j$  years after diagnosis, between pair  $i$ ’s case and control. It represents the advantage that pair  $i$ ’s cancer free control has over the corresponding breast cancer patient, corresponding to increase of income since one year prior to diagnosis.

Define the following linear regression equation for  $Y_{i,j}$ :

$$Y_{i,j} = \alpha + \beta_j + \gamma_{\alpha(i)} + \varphi x_i + \hat{\varphi} \hat{x}_i + \delta_i + \varepsilon_{i,j} \tag{1}$$

Here,  $\alpha$  is a constant term and  $\beta_j, j \in \{-1, 0, 1, 2, \dots\}$  is the effect that the number of years after diagnosis has on  $Y$ . These are the primary parameters, interpreted as the effect that the cancer diagnosis has had on the case’s income development,  $j$  years after diagnosis. The pre-diagnosis year ( $-1$ ) was used as reference category.

The parameters  $\gamma_{\alpha(i)}, \alpha \in \{44, 45, \dots, 67\}$  represent the effect that the pair’s age at the time of income has on  $Y$ . Youngest women (age 44 in the pre-diagnosis year) were used as reference group. Age may have a non-linear relation to income development and is therefore included in to the model as confounder. The vectors  $x_i$  and  $\hat{x}_i$  represent level of education, immigration history, and number of children for cases and controls, respectively. Here, the vectors  $x_i$  and  $\hat{x}_i$  represent confounding socio-demographic factors that were not used in the matching procedure. Parameter vectors for cases and controls are denoted by  $\varphi$  and  $\hat{\varphi}$ , respectively.

The error term is given as the sum  $\delta_i + \varepsilon_{i,j}$ , where both  $\delta_i \sim N(0, \tau^2)$  and  $\varepsilon_{i,j} \sim N(0, \sigma^2)$  are independently identically distributed. Here,  $\delta_i$  is a random error associated with pair  $i$  and is constant for all years  $j$ . It captures the clustering effect of correlated observations from the same pair  $i$  over time. The



term  $\varepsilon_{i,j}$  is a random error associated with each individually observed pair  $i$  at year  $j$ . This error structure represents multi-level modelling, with the pair as the clustering unit.

Note that the right-hand side of Equation (1) can be expressed in a standard form of a sum of parameters and predictors as follows:

$$Y_{i,j} = \alpha + \sum_j \beta_j d_{j,j'} + \sum_{\alpha'} \gamma_{\alpha'} e_{i,j,\alpha} + \sum_k \varphi_{j,k} x_{i,k} + \sum_k \hat{\varphi}_{j,k} \hat{x}_{i,k} + \delta_i + \varepsilon_{i,j}$$

Here,  $d_{j,j'}$  is a dummy variable, equal to 1 if  $j=j'$ , and 0 otherwise. Similarly,  $e_{i,j,\alpha} = 1$ , when pair  $i$  was of age  $\alpha j$  years after diagnosis, and 0 otherwise.

In the model estimation, a combination  $ij$  of a pair  $i$  and number years  $j$  after diagnosis was only used if both the case and control were alive at that time.

## References

1. Cancer in Norway 2010. Cancer incidence. Mortality, survival and prevalence in Norway. 2010 ([http://www.kreftregisteret.no/Global/CIN\\_2010.pdf](http://www.kreftregisteret.no/Global/CIN_2010.pdf))
2. Rayce SL, Christensen U, Hougaard CØ, Diderichsen F. Economic consequences of incident disease: the effect on loss of annual income. *Scand J Publ Health*. 2008;36:258–64.
3. Syse A, Tønnessen M. Cancer’s unequal impact on incomes in Norway. *Acta Oncol*. 2012;51:480–89.
4. Syse A, Tretli S, Kravdal Ø. Cancer’s impact on employment and earnings – a population-based study from Norway. *J Cancer Surviv*. 2008;2:149–58.
5. Hauglann B, Šaltytė Benth J, Fosså SD, Dahl AA. A cohort study of permanently reduced work ability in breast cancer patients. *J Cancer Surviv*. 2012;6:345–56.
6. Eaker S, Wigertz A, Lambert PC, Bergkvist L, Ahlgren J, Lambe M. Breast cancer, sickness absence, income and marital status. A study on life situation 1 year prior diagnosis compared to 3 and 5 years after diagnosis. *PLoS ONE*. 2011;6:e18040.
7. Larsen IK, Småstuen M, Johannesen TB, Langmark F, Parkin DM, Bray F, et al. Data quality at the Cancer Registry of Norway: an overview of comparability, completeness, validity and timeliness. *Eur J Cancer*. 2009;45:1218–31.