
Chris Robertson, a Concetta Perone, a Maja Primic-Zakelj, b Vera Pompe Kirn c and Peter Boyle a

Background Analyses of time trends in breast cancer incidence and mortality have generally revealed cohort-based changes in the rates. These have been linked to cohort-based changes in lifestyle factors. The effect of the changes in the reproductive risk factors on the changes in the rates, and the relative importance of the reproductive characteristics in Slovenia, a country which has not had much breast cancer screening, are investigated.

Methods Data on breast cancer incidence for 1971–1993 were obtained from the Cancer Registry of Slovenia (Registry). The Registry covers the whole population of the Republic of Slovenia (1.99 million on 30 June 1993). The statistical analysis uses parametric age-period-cohort models.

Results Breast cancer incidence has increased by 70% in Slovenia from 1971 to 1993. These changes are dominated by cohort effects and the cohorts born in 1907–1922 have the greatest increase in incidence. Period effects on changes in incidence were modest. The percentage of nulliparous women in the cohort and the average family size in the cohort explained 38% of the variation in the cohort effects.

Conclusions The percentage of nulliparous women in the cohort is the most important reproductive variable associated with the trends in the rates, with breast cancer risk predicted to be higher in cohorts with a larger percentage of nulliparous women. As the cohorts born 1932–1946 have a more favourable reproductive pattern as regards breast cancer risk, compared to the 1907–1922 cohorts, age-specific incidence rates in Slovenia would be predicted to decline in the future in the absence of changes in the other risk factors.

Keywords Breast cancer, age period cohort models, reproductive factors, time trends

Accepted 8 May 2000

Breast cancer incidence rates are increasing in many countries. 1, 2 They are subject to a range of influences, including the effects of the introduction of large-scale mammographic screening programmes and changes in the underlying risk factor distributions, and a clearer understanding of underlying trends may be obtained from statistical modelling. An investigation of time trends in breast cancer mortality in 20 countries using age-period and age-cohort models concluded that there was evidence of a decline in mortality in most countries which can be attributed in part to period and in part to cohort effects. 3 Evidence has been presented of a change in the cohort trend around 1925 in both US and Japanese mortality data. 4 A significant moderation of both breast cancer incidence and mortality rates in Scotland associated with cohorts of women born after World War II compared to women born between the two world wars has also been demonstrated. 5

Case-control studies and cohort studies tend to show that for an individual woman breast cancer risk is predicted to be greater if the woman is nulliparous or has a small family size, if she was older at the birth of her first child, if her age at menarche was younger and if her age at menopause was older. Dos Santos Silva and Swerdlow 6 concluded that, in England and Wales, for women born before 1914 there was a correlation over time in the decrease in fertility at young ages with the increase in breast cancer risk. However, for women born subsequently the negative relationship was replaced by a positive one. A further analysis of breast cancer mortality trends in the US and Canada noted that the decreasing birth cohort trends among women born after 1950 were in contrast to the changes in fertility patterns. 7 There was a reduction in the proportion of women aged 20–24 who were nulliparous until the cohort born

a Division of Epidemiology and Biostatistics, European Institute of Oncology, Milan, Italy.
b Epidemiology Unit, Institute of Oncology, Ljubljana, Slovenia.
c Cancer Registry of Slovenia, Ljubljana, Slovenia.
Reprint requests: Dr C Robertson, Division of Epidemiology and Biostatistics, European Institute of Oncology, Via Ripamonti 435, 20141 Milan, Italy. E-mail: chris.robertson@ieo.it
around 1940 and this increase in fertility is associated with a reduction in breast cancer mortality. For women born after 1940 there was a reduction in fertility which would be expected to be associated with an increase in breast cancer risk for these cohorts but there appears to be a continued reduction in breast cancer mortality.

Reliable breast cancer incidence data are available from some longstanding European cancer registries. In Slovenia there has been an active population-based Cancer Registry since 1950, covering a population of about 2 million at the 1991 census. Here we present the results of an analysis of time trends in breast cancer incidence in Slovenia using age-period-cohort models. The results show the changes in breast cancer incidence, and the effect of the changes in the reproductive risk factors on the changes in the rates, and the relative importance of the reproductive characteristics in a country which has not had much breast cancer screening.

Materials and Methods

Data on breast cancer incidence for the years 1971 through 1993 were obtained from the Cancer Registry of Slovenia (Registry). The Registry covers the whole population of the Republic of Slovenia (1.99 million on 30 June 1993) and is located at the Institute of Oncology in Ljubljana, where more that 50% of cancer patients in Slovenia are admitted. The main data sources are notifications gathered from all hospitals and diagnostic centres in Slovenia. This information is completed by death certificates and autopsy protocols stating cancer diagnosis, and cases were confirmed microscopically.

The Registry provided individual anonymous records of all breast cancer patients diagnosed and registered in the period 1974–1993. Population data were available from Population Registry of the Republic of Slovenia. When the methods of Tarone and Chu were used, the data were arranged in 29 2-year age groups from 26 to 83 and in 11 2-year time periods from 1972 to 1993. In all there were 39 birth cohorts beginning 1888–1891, 1890–1893, etc. There is some overlap in the limits of these cohorts but with 75% of the births centred in the middle two years, the cohorts are referred to by the first of these two years, namely, 1889, 1891, ... 1965.

The analysis which incorporates the demographic information about cohorts uses 5-year age groups from 30 to 84 and 5-year time periods from 1974 to 1993 giving overlapping cohorts of 10 years which are considered to be centred on the middle 5 years. This is necessary as the demographic information about cohorts is only available for these 5-year cohorts.

Demographic information from two sources was used. The first source was the data on the percentage of women who were nulliparous and the average number of children born to each woman covering cohorts from 1875 until 1950. These data are available from a study analysing the evolution of fertility in Slovenia from the end of the 18th century and were obtained from population surveys and vital statistical data, named in the tables as Official Statistics.

The second source of demographic information comes from three large case-control studies of breast cancer which were carried out in Slovenia between 1960 and 1990. Data from the 2286 controls (mean age 52.9) in the first study carried out in 1965–1967, 1989 (mean age 44.5) in the second carried out in 1980–1983 and 624 (mean age 44.9) in the third carried out in 1988–1990 were used to provide estimates of the reproductive variables. This source has the advantage of providing information over a slightly wider range of cohorts as the control women in the studies had an age range of 25–80+ (first study), 25–54 (second and third studies) thus giving information on cohorts from 1880 until 1965. In all studies the study region was all of Slovenia. The first two studies used hospital-based studies while the last used population-based controls. In all studies the controls were matched to the cases on age and region and, for the hospital-based studies, on data of admission to hospital. In Slovenia, matching on region is a surrogate for matching on social class.

For the cohorts 1955 and 1960, the information on the percentage of nulliparous women and on the average number of children per woman is incomplete (no information from Official Statistics), as these women are still in the reproductive phase of their lives. The three cells in the 11 age groups by four time periods table of rates corresponding to these cohorts are excluded from the analysis.

There are two strands to the statistical analysis. Initially we use an age-period-cohort model among the cases aged 26–83 years to estimate the change in the trends associated with period and cohort. It is assumed that age, calendar period and birth cohort have an additive effect on the log incidence rate,

\[
\log \lambda_{ijk} = \mu + \alpha_i + \pi_j + \gamma_k
\]

where the incidence rates are represented by \(\lambda_{ijk}\), the age effects by \(\alpha_i\), the period effects by \(\pi_j\) and the cohort effects by \(\gamma_k\).

It is well known that the linear dependence among the three time factors, age, period and cohort, in the model leads to the non-identifiability of the parameters. Unlike the slopes of the linear trend in birth cohorts (or calendar periods), the differences in the slopes are identifiable. Therefore particular contrasts in birth cohort (or calendar period) parameters can be defined which, as they are invariant to the particular parameterization chosen, allow the estimation of identifiable parameters.

An appropriate contrast is to compare the linear trend in the eight cohorts from 1900 to 1914 to the eight cohorts born from 1924 onwards:

\[
C_{\text{1900-1924}} = 7\lambda_{1938} + 5\lambda_{1936} + 3\lambda_{1934} + 1\lambda_{1932} - 1\lambda_{1930} - 5\lambda_{1926} - 7\lambda_{1924} - 7\lambda_{1914} + 5\lambda_{1912} + 3\lambda_{1910} + 1\lambda_{1908} - 1\lambda_{1906} - 3\lambda_{1904} - 5\lambda_{1902} - 7\lambda_{1900}.
\]

This is the same as contrast \(C_2\) of Tarone and Chu. If the estimated value is zero then the linear trend in the cohort effects among the 1900–1914 cohort era is the same as that in the 1924–1938 era. If \(C_{\text{1900-1924}}\) is negative then the trend in the younger cohort era is not as steep as in the older cohort era and this would provide evidence that there had been a moderation of breast cancer risk among the younger cohort era. This is not interpreted as a reduction in risk, though this may have happened, but merely that the rate of increase in risk is not as steep.

In the notation adopted here for the contrast \(C_{\text{m, y, z}}\), \(y\) is the beginning of the first cohort era, \(z\) is the beginning of the second
cohort era, and \( m \) is the number of cohorts in each era. These cohorts are all of equal width (\( w \) years) and so the first era is from \( y \) to \( y + w(m - 1) \) with the second from \( z \) to \( z + w(m - 1) \). This uses the convention that the cohort is referred to by the first of the two central years in it. The difference between \( y \) and \( z \) must be at least \( w(m + 1) \) years to avoid any overlap in the cohort eras. Similar contrasts can be defined for periods and these are denoted \( P^m_{y,z} \) where \( y \) is the beginning of the first period, \( z \) is the beginning of the second period, and \( m \) is the number of periods in each block, for example,

\[
P^4_{1972,1982} = \frac{2\pi_{1990} + \pi_{1988} - \pi_{1984} - 2\pi_{1982}}{(2\pi_{1980} + \pi_{1978} - \pi_{1974} - 2\pi_{1972})}.
\]

If the periods are all of width \( w \) years then the difference between \( y \) and \( z \) must be at least \( wm \) years to avoid any overlap in the period blocks. This is different from the cohort eras as the periods have no overlap whereas there is overlap in the cohorts.

In our analysis the estimates of age, period and cohort effects were obtained under the constraint that the first age and period effect and the first and the last cohort effects are zero, though it is immaterial which constraint is used as the contrasts \( C^m_{y,z} \) and \( P^m_{y,z} \) are invariant to the constraint.4

Secondly, we include data on some of the main reproductive risk factors of breast cancer within an age-period-cohort model.13 The cohort parameters, \( \gamma_k \), are replaced by the demographic variables in an attempt to link the changes in the trend associated with birth cohort to changes in reproductive characteristics.

### Results

Breast cancer incidence increased significantly from an age-standardized rate of 55.6 in 1971 to 94.8 in 1993. The age-specific rates (Figure 1) show that there has been an increase in incidence in recent years in women over 40 and particularly among women over 60. This is consistent with a cohort effect as seen from the right hand panel of the graph.

We calculate the same contrasts, in terms of years, as Tarone and Chu,4 corresponding to the 1901 birth cohort and the 1925 birth cohort, \( C^8_{1901,1925} \) takes the value \(-0.095\) with a standard error of 0.114 (\( P \)-value 0.20). The contrast \( C^8_{1901,1923} = -3.602 \) with a standard error of 1.043 (\( P < 0.01 \)) and \( C^8_{1923,1947} = 3.159 \) with a standard error of 2.533 (\( P = 0.11 \)). The series of contrasts calculated over all the range of cohorts (Table 1) provides evidence of a highly significant moderation of incidence risk beginning with women born around 1923; the biggest change in the birth cohort trend occurred around 1925: \( C^6_{1899,1925} = -4.084 \) with a standard error of 0.68 (\( P < 0.01 \)).

\( P^4_{1972,1982} \) takes the value 0.282 with a standard error of 0.144, \( P^4_{1974,1984} \) takes the value 0.301 with a standard error of 0.140, indicating a slight increase in the calendar period slope since the mid-1980s. There is a suggestion that some of the increase in the risk of breast cancer may be associated with period effects.

The cohort-based demographic information is presented in Table 2. The figures for the average number of children are relatively consistent over the two sources bearing in mind that the Official Statistics give the average number of children per woman whereas the Control Series gives the average number of children per parous woman. There is considerable discrepancy in the estimated percentage of nulliparous women among the older cohorts though the trends are largely similar from the 1907 cohort onwards. In Table 2 we see that the percentage of nulliparous women was greatest at the 1907–1922 cohorts which coincides with the cohorts showing greatest evidence of increases in the rates in Figure 1. There is no reliable information on the percentage of women nulliparous in the 1957–1961 cohort, nor on the average number of children per woman for

![Figure 1](plot_of_age_specific_incidence_rates.png)
the cohorts 1952–1956 and 1957–1961 as most of these women are still in the reproductive phase of their lives. The three cells in the table corresponding to these cohorts must be excluded from this analysis.

From Table 2, we see that fertility is decreasing. The decline in completed family size was consistent until the 1927–1931 cohort since when it has been level at 2.0 or below. Although the proportion of women having at least one child has increased since the cohorts born before 1900 (where less than 80% were parous) the proportion of women having two or more children has decreased. Diminishing family size was accompanied by a shorter interval in which the family was completed with women having their children at younger ages.9

The deviances and degrees of freedom for the models (Table 3) show that there is evidence that there are significant cohort-based curvatures in the time trends in Slovenia ($\chi^2 = 38.3, 10$ d.f.). Using the Official Statistics demographic data we see that both demographic variables have a significant effect on their own with the percentage of women nulliparous being the more important. These two variables explain 38% of the deviance associated with the cohorts. The predicted rate of breast cancer from model 5, Table 3, is greater in cohorts with a higher percentage of nulliparous women (an increase in log rate of 0.036 [0.014] for a one per cent increase), and lower in cohorts with a higher average number of children per woman (a decrease in log rate of –0.273 [0.162] for a unit increase in average number of children). This model is not as good a fit as the full age-period-cohort model ($\chi^2 = 20.31, 8$ d.f.). The main lack of fit comes from the 1927–1931 cohort which has lower rates than predicted.

Using the Control Series data gives a similar picture. The percentage of nulliparous women is the most important variable linked with the changes in the rates. The three demographic

### Table 1: Tarone and Chu cohort contrasts comparing non-overlapping cohorts in blocks of six 2-year cohorts

<table>
<thead>
<tr>
<th>First cohort in first block</th>
<th>First cohort in second block</th>
<th>Contrast</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1889</td>
<td>1915</td>
<td>–0.989</td>
<td>3.054</td>
</tr>
<tr>
<td>1891</td>
<td>1917</td>
<td>–2.812</td>
<td>2.156</td>
</tr>
<tr>
<td>1893</td>
<td>1919</td>
<td>–3.053</td>
<td>1.627</td>
</tr>
<tr>
<td>1895</td>
<td>1921</td>
<td>–3.765</td>
<td>1.343</td>
</tr>
<tr>
<td>1897</td>
<td>1923</td>
<td>–5.618</td>
<td>1.222</td>
</tr>
<tr>
<td>1899</td>
<td>1925</td>
<td>–4.084</td>
<td>0.680</td>
</tr>
<tr>
<td>1901</td>
<td>1927</td>
<td>–2.013</td>
<td>1.066</td>
</tr>
<tr>
<td>1903</td>
<td>1929</td>
<td>–2.041</td>
<td>1.076</td>
</tr>
<tr>
<td>1905</td>
<td>1931</td>
<td>–2.689</td>
<td>1.054</td>
</tr>
<tr>
<td>1907</td>
<td>1931</td>
<td>–3.127</td>
<td>1.046</td>
</tr>
<tr>
<td>1909</td>
<td>1935</td>
<td>–3.567</td>
<td>1.067</td>
</tr>
<tr>
<td>1911</td>
<td>1937</td>
<td>–3.551</td>
<td>1.132</td>
</tr>
<tr>
<td>1913</td>
<td>1939</td>
<td>–4.085</td>
<td>1.260</td>
</tr>
<tr>
<td>1915</td>
<td>1941</td>
<td>–2.635</td>
<td>1.406</td>
</tr>
<tr>
<td>1917</td>
<td>1943</td>
<td>–3.728</td>
<td>1.741</td>
</tr>
<tr>
<td>1919</td>
<td>1945</td>
<td>–3.812</td>
<td>2.127</td>
</tr>
<tr>
<td>1921</td>
<td>1947</td>
<td>–3.034</td>
<td>2.330</td>
</tr>
<tr>
<td>1923</td>
<td>1949</td>
<td>–6.686</td>
<td>4.145</td>
</tr>
<tr>
<td>1925</td>
<td>1951</td>
<td>0.270</td>
<td>4.396</td>
</tr>
</tbody>
</table>

The first row of this table gives the estimated contrast $C_{1889,1915}$ and its standard error. This contrast compares the linear slope for the cohorts 1889 and 1899 with the linear slope for the cohorts from 1915 until 1925. The estimated value is negative which means that the rate of increase in the rates for the 1889 to 1899 cohorts is greater than the rate of increase in the rates for the 1915 to 1925 cohorts.

### Table 2: Demographic information on cohorts

<table>
<thead>
<tr>
<th>Birth cohorts</th>
<th>Official Statistics</th>
<th>Control Series data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average no. of</td>
<td>No.</td>
</tr>
<tr>
<td></td>
<td>children per woman</td>
<td></td>
</tr>
<tr>
<td>1882–1886</td>
<td>4.0</td>
<td>45</td>
</tr>
<tr>
<td>1887–1891</td>
<td>3.6</td>
<td>129</td>
</tr>
<tr>
<td>1892–1896</td>
<td>3.4</td>
<td>255</td>
</tr>
<tr>
<td>1897–1901</td>
<td>3.3</td>
<td>283</td>
</tr>
<tr>
<td>1902–1906</td>
<td>3.0</td>
<td>317</td>
</tr>
<tr>
<td>1907–1911</td>
<td>2.8</td>
<td>288</td>
</tr>
<tr>
<td>1912–1916</td>
<td>2.6</td>
<td>203</td>
</tr>
<tr>
<td>1917–1921</td>
<td>2.4</td>
<td>293</td>
</tr>
<tr>
<td>1922–1926</td>
<td>2.3</td>
<td>696</td>
</tr>
<tr>
<td>1927–1931</td>
<td>2.1</td>
<td>659</td>
</tr>
<tr>
<td>1932–1936</td>
<td>2.0</td>
<td>574</td>
</tr>
<tr>
<td>1937–1941</td>
<td>1.9</td>
<td>410</td>
</tr>
<tr>
<td>1942–1946</td>
<td>1.8</td>
<td>251</td>
</tr>
<tr>
<td>1947–1951</td>
<td>6.8</td>
<td>99</td>
</tr>
<tr>
<td>1952–1956</td>
<td>6.8</td>
<td>26</td>
</tr>
</tbody>
</table>

\(^a\) Estimated.  
\(^b\) Overestimates as these women are still in the reproductive phases of their lives at the time of the earliest survey.  
\(^c\) Slightly biased as these women are still in the reproductive phases of their lives at the time of the earliest survey.  

The data calculated from the census and vital registrations were reported in Sircelj9 for cohorts 1883–1887, 1888–1892, etc. These are slightly different, by one year, from the cohort groups used for the control series data. This difference is not likely to have any major influence.
variables in model 9, Table 3, explain 66% of the deviance associated with cohort. The parameter estimates from this model are: percentage nulliparous 0.021 (0.005); average number of children per parous woman –0.124 (0.116) and median age at first birth –0.022 (0.039). The rates increase as the percentage of nulliparous women increases, as the average number of children per parous women decreases and there is no influence of the median age at first birth. In the model with only age, period and the median age at first birth the estimated effect of the median age at first birth is 0.069 (0.032) which implies that the rates increase as the median age at first birth increases.

Discussion

The analysis has demonstrated that the increases in breast cancer rates among women in Slovenia were dominated by cohort-based changes. There is some evidence of a period-based increase in the rates in the last 10 years compared to the first 10 years. The 1907–1921 cohort is associated with the major increases in the rates. This cohort of women had extremely high levels of nulliparity, approaching 20% of the cohort.

Demographic variables summarizing the reproductive characteristics of the cohort are associated with the changes in the rates. The rates increase in cohorts with a high median age at first birth, decrease in cohorts with a greater average number of children per woman and increase in cohorts with a high percentage of nulliparous women. The latter variable is the strongest predictor.

The cohort-based risk has been declining such that the younger cohorts are at a reduced risk of breast cancer. This is at a time when the average family size has been decreasing with cohort. However, the percentage of the cohort who were nulliparous has been decreasing and the median age at first birth has also been decreasing. These changes are compensating for the effect of a reduction in family size.

One implication of this analysis is that we should expect a slight reduction in the age-specific incidence rates of breast cancer in Slovenia in the not too distant future. As the population is ageing the absolute numbers will still increase. In 1990, women of the 1907–1921 cohort were aged 69–83 and by the year 2000 most will be dead. Incidence rates in the early years of the 21st century will be based upon cohorts born in the period 1930–1950 and these cohorts have a more favourable distribution of some reproductive variables but not all, e.g. age at first full-term pregnancy. This implication is based upon the assumptions that these cohort effects remain constant and that the model is fully specified. The latter is not the case as other risk factors such as age at menarche and diet were not taken into account in our analysis.

The cross-sectional analyses of fertility in Slovenia show that, as in most European countries, fertility has been decreasing in the last 100 years, but not evenly. The latest period of sharp decrease started in 1981 and is characterized by the continuous decrease in total fertility, continuous decrease in fertility among women under 25 years and in the increase in the mean age of women at the birth of their first child. While the mean age of mother at birth of her first child was 23.2 in 1985, it was 24.8 in 1994. There is very little information in our analyses about these recent trends in fertility in Slovenia and should the trends result in an increased percentage of the recent cohorts nulliparous then in the absence of changes of other factors which influence breast cancer risk, an increase in rates can be expected.

The main lack of fit of the age-period model with the demographic information is associated with the 1927–1931 cohort. This cohort would have been aged around 20 just after the end of World War II and would have been at the ages of puberty during the war. Similar changes in the World War II cohorts have also been observed in Norway, chiefly for endometrial cancer risk factors, and in the Netherlands. It may well be that the alteration of eating patterns due to rationing in wartime and the immediate post-war period influenced certain breast cancer risk factors, in particular delaying menarche and so influenced the breast cancer rates for this cohort.

Each cohort is treated as having a fixed reproductive effect. This is fine for the cohorts which have completed their reproductive phase before the beginning of the study period, i.e. the cohorts from 1892 until 1932, and possibly 1937, where the reproductive phase is largely over. However, in the younger cohorts the percentage of nulliparous will decrease as the cohort ages. Thus the 1942–1946 cohort will have a greater percentage
nulliparous in 1974–1978 when the average age is 32 than in 1979–1983 when the average age is 37. This time-dependent effect has not been included in the modelling as sufficient data were not available at each time period. The sensitivity of the estimates to the mispecification among the younger cohorts was investigated by fitting the model to only women over 40. No appreciable difference in parameter estimates was observed.

Cohort-based changes in breast cancer risk have been noted in a large number of studies: in Japan,21 in Saskatchewan, Canada,22 in Sweden,23 in Singapore,24 and in Scotland.25 Similar reports have also been made for mortality: in Connecticut, US,25 in Taiwan,26 in Sweden27 and in the US.28 These studies have generally concluded that the birth cohort changes indicate persistent secular changes in largely unknown risk factors associated with lifestyle. From epidemiological studies of breast cancer risk the likely candidate risk factors are reproductive factors and dietary habits, including alcohol use.

In England and Wales dos Santos Silva and Swerdlow6 noted that there was a negative relationship between the rate of breast cancer mortality and the percentage of the cohort who were nulliparous by the age of 40 for cohorts born from 1875 until 1925. In recent years the percentage of the cohorts who were nulliparous continued to decrease but mortality rates also decreased. They argue that the lack of agreement in this and other reproductive trends is a result of the influence of other risk factors; in particular they discuss the possible protective role of oral contraceptive use and changes in dietary fat intake.

Ewertz and Duffy28 reported that the changes in the breast cancer incidence rates in Denmark for 1943–1989 could not be explained wholly by changes in fertility but that a considerable part may be explained by changes in dietary habits and alcohol consumption. They did not, however, use the most important reproductive predictor namely the percentage of nulliparous women. In contrast we find that the changes in the demographic variables can explain a substantial portion of the cohort effect. This analysis does not exclude the possibility that there are other lifestyle changes which also influence the changes in breast cancer risk.

References