

2 Methods summary

The objective of this report is to estimate relative survival proportions for women with breast cancer in Australia. The women included in this analysis were those diagnosed with breast cancer from 1982 to 1994 inclusive with survival follow-up to 31 December 1994. The analysis examines relative survival estimates, taking into account the year of diagnosis, age at diagnosis, State or Territory of usual residence, and urban or rural area of usual residence.

To achieve this objective four key elements are required:

- data relating to women diagnosed with breast cancer and their vital status;
- population-based lifetables;
- relative survival analysis software to combine the first two elements; and
- staff with knowledge of the data systems, their linkage and survival analysis.

The methods used to combine these elements are described briefly here and in more detail in Appendix A.

Preparation of breast cancer data and lifetables

Breast cancer incidence data were derived from the National Cancer Statistics Clearing House (NCSCCH) database (Table 1). As at February 1998, the NCSCCH contained identifiable information relating to new breast cancer cases and deaths between 1982 and 1994. This group of women is referred to throughout the report as the breast cancer 'cohort'. Queensland cancer incidence and mortality data were not complete at the time of analysis and were therefore excluded.

Table 1: Summary of breast cancer data initially extracted from the NCSCCH, 1982-1994

State/Territory	New cases (cohort)	Per cent of new cases (cohort)	Total deaths (cohort)	Per cent of total deaths (cohort)
New South Wales	31,792	42.2	9,992	42.1
Victoria	23,696	31.4	7,623	32.1
Western Australia	8,089	10.7	2,533	10.7
South Australia	8,038	10.7	2,476	10.4
Tasmania	2,293	3.0	765	3.2
Australian Capital Territory	1,212	1.6	263	1.1
Northern Territory	278	0.4	83	0.3
Australia^(a)	75,398	100.0	23,735	100.0

(a) Excludes Queensland data.

A death clearance of all breast cancer incident cases was undertaken using probabilistic linkage between the NCSCCH and the National Death Index (NDI), which contains data on all deaths in Australia from 1980. State and Territory cancer registries validated this death clearance by following up proposed matches through a range of sources, e.g. hospitals,

pathology laboratories, other cancer registries and treating doctors. A deduplication probabilistic linkage on the final database ensured that cases were not counted in more than one State or Territory. Quality assurance procedures were used to ensure consistency of data within the project and with previously published work. In handling the breast cancer data, strict confidentiality protocols (required under the Australian Institute of Health and Welfare Act 1987) were adhered to, with no identifiable information passed beyond the NCSCCH or the State and Territory cancer registries. No identifying information was included on the final data analysis file.

Lifetables were derived for Australia, each State and Territory, and urban and rural populations using deaths data and estimates of resident population. Queensland data were excluded from each of these lifetables as the data were not included in the survival analysis.

Relative survival analysis

Relative survival is defined as the ratio of the observed survival rate for a given cohort of patients to the expected survival rate (Ederer et al. 1961). The expected survival rate is that which the patient group should have experienced based on the lifetable of the general population from which they were diagnosed (Estève et al. 1990). A relative survival of less than 100% implies that the cohort survived for less time than would be expected for the general population. A relative survival of 100% implies that survival in the cohort is no different from that in the general population.

The relative survival analysis was undertaken using the RELSURV (v2.0) software, which was written by Guy Hédelin of Louis Pasteur University, Strasbourg, France (Hédelin 1995). The program calculates expected survival using the lifetable method and estimates relative survival using a Cox proportional hazards regression model.

In undertaking the relative survival analysis some key assumptions were made which are important in the interpretation of the results:

1. Records with the following characteristics were excluded from the survival analysis (see Appendix A):
 - any woman whose age at diagnosis was not known or was missing;
 - any woman aged 100 years or over at diagnosis;
 - all cases diagnosed in Queensland;
 - any woman whose diagnosis or death date could not be resolved;
 - death certificate only cases (i.e. cases diagnosed at death);
 - any woman who died within a month of diagnosis; and
 - any woman diagnosed in December 1994.
2. After excluding 8,088 records with the characteristics described in (1), the data file contained 73,827 records. However the software was not able to handle files larger than 65,535 records. To compensate for this in the analysis the following strategies were used.
 - (a) To estimate relative survival at the 'Australia, all ages, 1982-1994' level, a sample of 65,500 records (88.7% of all records) was systematically selected and used.
 - (b) To estimate relative survival at all other levels, the data file was split according to the level of analysis being undertaken. For example, for analysis by State and Territory, the data file was split into two files, one containing all records for New South Wales and Victoria, and the other containing all records for the remaining

States and Territories. The former file was used to estimate relative survival for New South Wales and Victoria and the latter file was used for all other States and Territories.

3. All cases were followed up to 31 December 1994, which was the censoring date. Unless a woman diagnosed with breast cancer was known to have died before 31 December 1994, it was assumed that she was still alive (i.e. censored). The impact of this censoring date, combined with the modeling methods used, is that survival proportions can be estimated for women for a number of years following their diagnosis: women diagnosed between 1982 and 1987, up to 10 years; those diagnosed 1988 to 1992, up to 7 years; and those diagnosed 1993 to 1994, up to 2 years.