

SUPPLEMENTARY DATA AVAILABLE ON THE CANCER EPIDEMIOLOGY JOURNAL
WEBSITE: TRENDS IN ETHNIC AND SOCIOECONOMIC INEQUALITIES
IN CANCER SURVIVAL, NEW ZEALAND, 1991-2004

Supplementary Table 1. Number and percentage of adults with a primary cancer diagnosis included in excess mortality rate analyses, by exclusion criteria and period of diagnosis

	Patients diagnosed 1991-1995/1996		Patients diagnosed 1996-2000/2001		Patients diagnosed 2001-2004	
	Number	%	Number	%	Number	%
Number of linked records	50,064		59,967		51,462	
Data restrictions						
Aged below 15 years of age at time of diagnosis	264	1	291	<1	237	<1
Aged above 99 years of age at time of diagnosis	15	<1	18	<1	21	<1
Zero survival or death certificate only, or unknown basis of diagnosis	4746	9	3864	6	2889	6
Tumour <i>in situ</i>	129	<1	0	<1	6	<1
Sex-site incompatibility	78	<1	78	<1	84	<1
Missing ethnicity	159	<1	495	1	522	1
Missing income	5484	11	8364	14	7659	15
Patients accepted for analyses in excess mortality modelling	39,219	78	47,013	78	40,245	78

Statistical Appendix: Methods used to calculate ‘pseudo’ average excess mortality rates over five years and ‘pseudo’ five-year relative survival ratios

To capture changes in inequalities on an absolute rate difference scale, we calculated regression-based changes over a decade in the pseudo-excess mortality rate difference (pseudo-EMRD) comparing low- and high-income groups, and likewise Māori and non-Māori, by cancer site (but not pooled over cancer sites). These are ‘pseudo’ because we are using an average EMR. Because the excess mortality rate decreases rapidly post-diagnosis, we used average excess mortality rates over five years (hereafter simply ‘EMR’) derived from five year relative survival risks (hereafter simply ‘RSR’). For example, the high income group’s average EMR at the beginning of the decade is:

$$EMR_{High,Year=0} = \frac{-\ln[RSR_{High,Year=0}]}{5}$$

where $RSR_{High, Year=0}$ is the RSR for the high-income group in 1991-96, directly estimated using standard relative survival methods [43]. The regression-based estimate of the high income group’s EMR in a decades time is then:

$$EMR_{High,Year=10} = EMR_{High,Year=0} \times EMRR_{year}$$

where $EMRR_{year}$ is the ratio reduction (usually) per decade in the high income (reference group for income) group’s EMR from model 4 above. The regression-based estimate of the low income group’s EMR at the beginning of the decade is:

$$EMR_{Low,Year=0} = EMR_{High,Year=0} \times EMRR_{income}$$

where $EMRR_{income}$ is the rate ratio comparing low to high income from model 4 above. And the regression- based estimate of the low income group’s EMR in a decades time is:

$$EMR_{Low,Year=10} = EMR_{Low,Year=0} \times EMRR_{year} \times EMRR_{income \times year}$$

**SUPPLEMENTARY DATA AVAILABLE ON THE CANCER EPIDEMIOLOGY JOURNAL
WEBSITE: TRENDS IN ETHNIC AND SOCIOECONOMIC INEQUALITIES
IN CANCER SURVIVAL, NEW ZEALAND, 1991-2004**

where $EMRR_{\text{income} \times \text{year}}$ is the rate ratio for the interaction of income and calendar year from model 4 above. From these four formula, EMRs can be calculated psuedo-EMRDs at the beginning and end of a decade, and thence the absolute change over a decade in the psuedo-EMRD. The same process was used for ethnicity. We did not attempt to calculate confidence intervals about these measures.

Finally, using the empirically calculated RSRs for high income (or non-Māori) at the beginning of the decade (i.e. the same starting estimate as above) and the back-calculated RSRs from the three above regression-based estimates of EMRs (using the formula $RSR = \exp[-EMR \times 5]$), we likewise estimated changes in the absolute differences by income (and ethnicity) in RSRs over a decade.