Measuring social class differences in cancer patient survival: Is it necessary to control for social class differences in general population mortality? A Finnish population-based study

Paul W Dickman, Anssi Auvinen, Esko T Voutilainen, Timo Hakulinen

Abstract

***Study objectives*—Estimation of cancer patient survival by social class has been performed using observed, corrected (cause specific), and relative (with expected survival based on the national population) survival rates. Each of these measures are potentially biased and the optimal method is to calculate relative survival rates using social class specific death rates to estimate expected survival. This study determined the degree to which the choice of survival measure affects the estimation of social class differences in cancer patient survival.**

***Setting and participants*—All Finnish residents diagnosed with at least one of 10 common malignant neoplasms during the period 1977–1985 were identified from the Finnish Cancer Registry and followed up for deaths to the end of 1992.**

***Design*—Survival rates were calculated by site, sex, and age at 5, 10, and 15 years subsequent to diagnosis for each of three measures of survival; relative survival, corrected (cause specific) survival, and relative survival adjusted for social class differences in general mortality. Regression models were fitted to each set of rates for the first five years of follow up.**

***Main results*—The degree of variation in relative survival resulting from social class decreased, although did not disappear, after controlling for social class differences in general mortality. The results obtained using corrected survival were close to those obtained using relative survival with a social class correction. The differences between the three measures were largest when the proportion of deaths from other causes was large, for example, in cancers with high survival, among older patients, and for longer follow up times.**

***Conclusions*—Although each of the three measures gave comparable results, it is recommended that relative survival rates are used with expected survival adjusted for social class when studying social class variation in cancer patient survival. If this is not an available option, it is recommended that corrected survival rates are used. Relative survival rates without the social class correction overestimate social class differences and should be used with caution.**

(*J Epidemiol Community Health* 1998;**52**:727–734)

Department of Cancer Epidemiology, Karolinska Institute, Stockholm, Sweden
P W Dickman

Finnish Centre for Radiation and Nuclear Safety, Helsinki, Finland
A Auvinen

Finnish Cancer Registry, Helsinki, Finland
E T Voutilainen

Department of Public Health, University of Helsinki, Helsinki, Finland
T Hakulinen

Correspondence to:
Dr P Dickman, Department of Cancer Epidemiology, Radiumhemmet, Karolinska University Hospital, 171 76 Stockholm, Sweden.

Accepted for publication
6 February 1998

Social class has been shown to be associated with survival time following a diagnosis of cancer,[1–27] cardiovascular disease,[28 29] and HIV/AIDS[30–32] (among other diseases). Social class differences have been studied most extensively among cancer patients, the first reports dating back to the 1950s[33–35] and later research conducted in the USA,[8 21 26 27] Europe,[1 2 6 7 12] Japan,[9] and Australia.[23] Although results have been somewhat conflicting, the largest population-based studies have reported social class differences in survival for several, but not all, types of cancer. Some of the discrepancies between results are probably real, but it is possible that some results may be biased because of methodological limitations.

The two main methodological limitations that may bias social class comparisons of survival are the choice of the measure of social class and the method used for estimating survival. In this paper we will concentrate on the second factor. The three broad methods for estimating cancer patient survival with long term follow up are observed survival rates (OSR), corrected (cause specific) survival rates (CSR), and relative survival rates (RSR).[36–39] Each of these measures have been applied to studies of cancer patient survival by social class.

Relative survival has become the preferred measure of survival for cancer patients with long term follow up. One advantage of this measure is that information on cause of death is not required, thereby circumventing problems with inaccuracy[40 41] or non-availability of death certificates. The RSR is defined as the observed survival in the patient group divided by the expected survival of a comparable group from the general population.[42] It is usual to estimate the expected survival from nationwide population life tables stratified by age, sex, calendar time, and race (where applicable).[43]

Large studies from the USA,[44] UK,[45–49] Finland[50] and other countries have demonstrated unequivocally that all cause mortality is higher among lower social classes. Thus, RSRs (with expected survival based on the national population) tend to overestimate social class differences in cancer patient survival. RSRs for lower social classes are underestimated because they experience a higher general (all cause) mortality than the general population, while the RSRs for the upper social classes are overestimated.

Table 1 Number and percentage of cases in the study population, and five year relative survival rates (SRSR) for each site and social class

Site	Social class	Men			Women		
		No	*%*	*SRSR*	*No*	*%*	*SRSR*
Stomach	I	617	9	0.18	456	8	0.21
	II	1283	19	0.18	1537	28	0.19
	III	3471	52	0.18	2119	39	0.17
	IV	1307	20	0.12	1389	25	0.15
Rectum	I	344	12	0.50	315	11	0.57
	II	699	25	0.44	887	31	0.48
	III	1285	47	0.42	1017	36	0.42
	IV	428	16	0.33	644	22	0.36
Pancreas	I	368	13	0.01	314	11	0.01
	II	635	23	0.02	879	31	0.02
	III	1313	47	0.02	990	35	0.01
	IV	485	17	0.02	681	24	0.01
Lung, trachea	I	948	6	0.09	204	10	0.12
	II	2547	17	0.11	610	31	0.11
	III	8242	53	0.08	708	36	0.10
	IV	3684	24	0.07	421	22	0.07
Breast	I				2499	14	0.77
	II				7153	40	0.74
	III				5543	31	0.71
	IV				2662	15	0.69
Cervix uteri	I				172	7	0.72
	II				746	30	0.63
	III				971	39	0.67
	IV				631	25	0.57
Corpus uteri	I				573	13	0.81
	II				1611	36	0.83
	III				1600	35	0.75
	IV				729	16	0.74
Kidney	I	385	16	0.41	240	12	0.41
	II	641	26	0.36	638	33	0.44
	III	1144	47	0.33	711	36	0.38
	IV	276	11	0.27	369	19	0.34
Urinary bladder	I	441	13	0.75	136	12	0.73
	II	716	21	0.64	367	32	0.63
	III	1690	49	0.64	416	36	0.60
	IV	613	18	0.53	243	21	0.51
Skin (non-melanoma)	I	199	15	0.86	143	10	0.91
	II	316	23	0.88	436	30	0.89
	III	652	48	0.90	548	37	0.86
	IV	205	15	0.83	348	24	0.89
All sites	I	3302	9		5052	12	
	II	6837	20		14864	35	
	III	17797	51		14623	34	
	IV	6998	20		8117	19	
	all	34934	100		42656	100	

In an attempt to avoid this bias, many researchers have used corrected (cause specific) survival rates, which take only deaths from the cancer in question as outcome events, while all other deaths are treated as censored events. However, the use of CSRs requires that reliably coded death certificates are available. Few studies have assessed the reliability of death certificates by social class, but it has been suggested that the accuracy of death certificates differs according to social class.[51] This probably leads to underestimation of social class differences in survival. The use of CSRs is also problematic because it is sometimes difficult to classify whether or not a death is because of cancer. For example, treatment complications may contribute to a patient's death without being the primary cause of death. Similarly, if a cancer patient has committed suicide, it is not clear whether or not to classify the death as being because of cancer. The use of RSRs avoids these problems by measuring the total excess mortality following a diagnosis of cancer, irrespective of whether the excess mortality is directly or indirectly related to the cancer.

Thus, at present, two main approaches have been applied to measuring social class specific cancer patient survival, one underestimating and the other overestimating social class differences in patient survival. Schrijvers and Mackenbach[6] reviewed 16 large, population-based studies that examined the association between social class and cancer patient survival. Five of these studies used the CSR as the measure of survival, one used the RSR (using national mortality tables), one used both CSRs and RSRs, one used standardised case fatality ratios, two used observed survival rates, and it was unclear in six studies whether a correction for non-cancer deaths was used.

The optimal method is to calculate social class specific RSRs using social class specific death rates to estimate the expected survival. This method takes into account social class variation in mortality from other diseases, but does not require information on causes of death (nor does it assume equal accuracy of that information). In this study we estimate survival using this method and compare the results with those from relative survival without the social class adjustment, and corrected survival.

Methods

Case specific data were obtained from the Finnish Cancer Registry for cases diagnosed in 10 cancer sites during the period 1971–1985 and followed up for deaths to the end of 1992 (table 1). These sites were chosen to cover a range of cancers with varying prognoses and primary sites with and without social class differences in survival.[1] The Finnish Cancer Registry is a nationwide, population-based registry established in 1952. Notification of new cancer cases is compulsory in Finland and, with the use of unique personal identification numbers, follow up procedures are extremely efficient.[52 53]

Information on the social class of each case was obtained from the 1970 population census by record linkage using the unique personal identification number given since 1967 to every resident of Finland.[54] The social class indicator was based on a person's own occupation, except for housewives, who were classified according to their husband's occupation. Retired persons were classified according to their previous occupation. The classification consisted of four groups: I (highest) included employers and higher administrative personnel, II consisted of lower administrative personnel and self employed persons, III of skilled workers, and IV (lowest) of unskilled workers.[55 56] Farmers (26 517 cases) were not included in the study, neither were those with unknown social class (6763 cases). The group of persons with unknown social class consisted mainly of census non-responders and retired persons whose previous occupation was unknown. Farmers were excluded because this group is heterogeneous with respect to social class. In descriptive studies of survival by social class it is common to classify farmers to a social class based on, for example, farm size. As the main objective of this study is to investigate the performance of various measures of survival we preferred not to do this.

Table 2 Male rectal cancer. Age specific number of cases and relative survival rates calculated after 5, 10, and 15 years of follow up for each of the three measures. NRSRs are presented along with the difference between the NRSR and the other two measures

			5 year survival			10 year survival			15 year survival		
Social class	Age	No	NRSR	SRSR-NRSR	CSR-NRSR	NRSR	SRSR-NRSR	CSR-NRSR	NRSR	SRSR-NRSR	CSR-NRSR
I	0–49	32	54.8	−0.6	0.1	47.2	−1.2	−1.9	41.7	−1.8	3.5
I	50–64	109	56.3	−1.6	0.3	46.0	−2.9	3.0	46.8	−5.1	−1.9
I	65+	203	48.6	−3.5	−9.6	41.7	−6.1	−10.7	63.4	−14.6	−32.5
II	0–49	65	47.5	−0.2	1.4	36.5	−0.4	1.2	38.6	−0.7	−1.0
II	50–64	223	46.3	−0.6	−0.6	46.0	−1.4	−3.0	41.3	−2.0	−0.2
II	65+	411	42.4	−1.3	−5.8	38.7	−2.5	−8.5	31.9	−3.6	−5.7
III	0–49	100	48.5	0.1	1.3	40.5	0.1	3.9	41.2	0.3	3.2
III	50–64	414	46.0	0.3	1.3	36.8	0.6	1.4	31.1	0.9	5.7
III	65+	771	36.7	0.9	−4.1	34.6	2.0	−6.2	30.1	3.0	−4.7
IV	0–49	24	34.5	0.9	0.7	31.6	1.8	−0.8	20.2	2.0	10.5
IV	50–64	91	31.2	1.8	4.5	26.7	3.5	1.4	22.2	5.1	4.2
IV	65+	313	30.6	2.8	−1.3	24.6	5.3	−2.2	13.6	5.4	4.6

Cause of death data were obtained from Statistics Finland. All cause death rates by age (one year intervals), sex, calendar year (five year intervals), and social class were obtained from Statistics Finland for the Finnish population enumerated in the the 1970 population census. These mortality tables were required for the estimation of expected survival to estimate relative survival. For this study, the social class specific population mortality tables were constructed from individual level death certificate data. The personal identification number included on each death certificate was used to link to the 1970 population census data to establish social class. If such linkage is not available, it is also possible to construct these mortality tables from national all cause mortality data and social class specific standardised mortality ratios (appendix 1).

Cases diagnosed at necropsy (3% of the total cases) were included in the study with zero survival time. The rationale for this decision was that social class could have contributed to the cancer not being diagnosed earlier, during the lifetime of the subject. This was supported by the finding that the proportion of cases diagnosed at necropsy was higher in the lower social classes (ranging from 2.6% in social classes I and II to 4.0% in social class IV). If this is considered a bias, simply excluding these cases will not eliminate the bias, it will only reverse its direction. Case registrations based on "death certificate only" (less than 1% of the total cases) were also included in the analysis with zero survival time using similar rationale. The percentage of registrations based on death certificate only was also higher in the lower social classes (ranging from 0.3% in social classes I and II to 0.7% in social class IV).

The statistical software package developed by Hakulinen and Abeywickrama[57] was used to estimate three different sets of survival rates for annual intervals up to 15 years of follow up; CSRs, RSRs using nationwide expected survival rates (which we will abbreviate to NRSR), and RSRs using social class specific expected survival rates (SRSR). When estimating the CSR, only deaths from the cancer in question are classified as outcome events, while all other deaths are treated as censored events. Life table regression models[58] were fitted to the first five years of annual follow up using the GLIM statistical package.[59] This provided estimates of the relative risk of excess mortality resulting from cancer for each social class compared with the reference group (social class IV) corrected for age, sex, follow up, and an age by follow up interaction. The age by follow up interaction term was required to obtain an acceptable fit to the data and indicates that the age specific hazards are non-proportional for the first five years of follow up.

Results

Social class differences in survival were observed using all three approaches (NRSR, SRSR, and CSR), with the highest social classes having superior survival for most cancer types. Trends in survival across social class were consistent for each of the three measures. That is, if the NRSR survival was best for social class I, followed by II, III, and IV, then this trend was usually seen in the SRSR and the CSR. There were, however, differences

Table 3 Female breast cancer. Age specific number of cases and relative survival rates calculated after 5, 10, and 15 years of follow up for each of the three measures. NRSRs are presented along with the difference between the NRSR and the other two measures

			5 year survival			10 year survival			15 year survival		
Social class	*Age*	*No*	*NRSR*	*SRSR-NRSR*	*CSR-NRSR*	*NRSR*	*SRSR-NRSR*	*CSR-NRSR*	*NRSR*	*SRSR-NRSR*	*CSR-NRSR*
I	0–49	730	78.4	−0.1	0.2	68.6	−0.3	1.3	62.0	−0.6	1.2
I	50–64	932	77.3	−0.6	−0.5	64.2	−1.2	−0.1	57.9	−2.2	0.6
I	65+	837	78.4	−3.0	−5.5	66.0	−5.3	−6.2	57.9	−7.6	−6.4
II	0–49	2347	77.4	−0.1	0.1	65.9	−0.1	0.9	59.6	−0.2	2.0
II	50–64	2564	73.1	−0.2	1.3	60.1	−0.5	1.6	50.7	−0.9	4.3
II	65+	2242	72.0	−1.8	−0.7	56.5	−2.9	1.3	51.1	−4.0	−0.9
III	0–49	1321	75.2	0.0	0.5	62.6	0.1	1.3	57.8	0.2	3.0
III	50–64	1809	68.9	0.2	1.8	54.6	0.3	4.1	45.8	0.5	5.3
III	65+	2413	68.3	0.2	0.8	50.5	0.7	4.1	39.7	1.4	7.0
IV	0–49	403	75.5	0.2	1.1	65.0	0.4	1.4	59.5	0.7	1.1
IV	50–64	890	69.0	0.5	0.6	55.1	0.9	1.8	43.1	1.4	6.5
IV	65+	1369	63.0	1.9	1.8	47.7	3.4	3.7	39.0	5.2	6.6

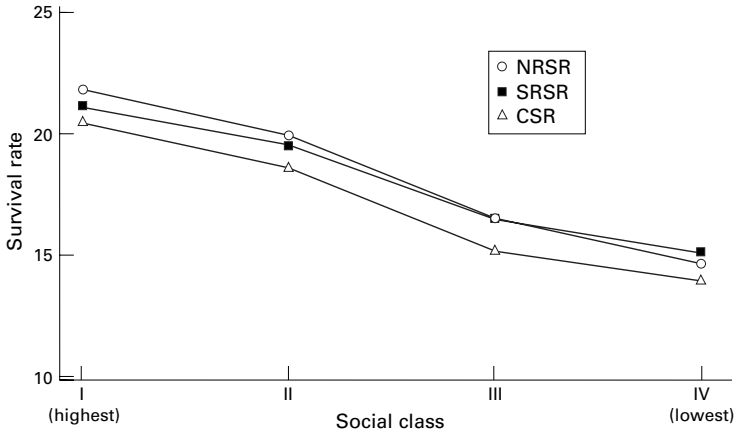

Figure 1 Female stomach cancer, social class specific five year survival rates calculated using three different measures of survival.

between the three measures and these differences were larger when the proportion of deaths from other causes was large, for example, in cancers with high survival, among older patients, and with longer follow up times (tables 2 and 3).

Observed survival is identical for the NRSR and the SRSR, so differences between these two measures are caused solely by differences in the expected survival. The SRSR is systematically lower than the NRSR in the upper social classes (I and II) and higher than the NRSR in the lower social classes (III and IV). As the differences in expected survival are larger after 10 years of follow up than five years (and after 15 years than 10), we observe corresponding differences in the NRSR and SRSR. The age effect occurs because social class differences in expected survival increase with age. Survival was generally highest in social class I and lowest in social class IV. The NRSRs exaggerate this difference because they overestimate the relative survival in social class I and underestimate the relative survival in social class IV. Hence, there is a larger amount of social class variation in the NRSRs than in the SRSRs.

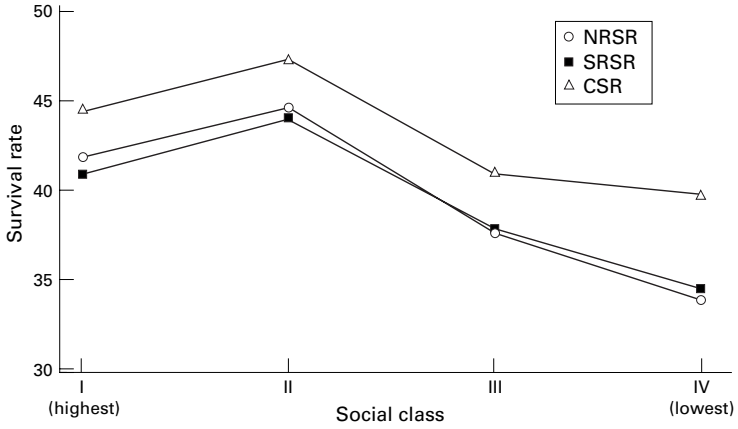

Figure 2 Female kidney cancer, social class specific five year survival rates calculated using three different measures of survival.

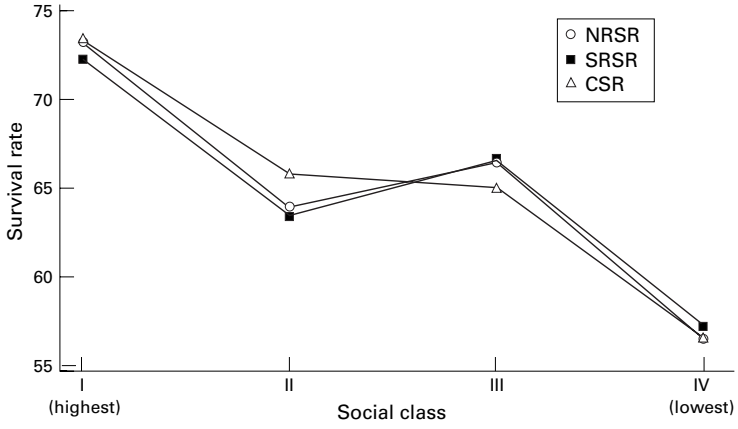

Figure 3 Cervical cancer, social class specific five year survival rates calculated using three different measures of survival.

The differences between the NRSRs and the SRSRs are similar for all sites and can be categorised systematically, as they depend solely on differences in the expected survival rates. This is not the case for the CSRs. The CSRs were systematically higher than the NRSRs and SRSRs in two sites (kidney and skin), systematically lower in two sites (stomach and rectum), and did not differ systematically in the other six sites (pancreas, lung, breast, cervix, corpus, and bladder). This same pattern was observed for the 5 year, 10 year, and 15 year survival rates, although the differences between the three measures were larger in the longer follow up periods. These three scenarios are presented diagrammatically in figures 1–3. It is notable in the figures that the NRSR is higher than the SRSR in social classes I and II but lower in social classes III and IV. Even though the CSRs are systematically different from the SRSRs and NRSRs in figures 1 and 2, the general trend in survival across social class is the same for all three measures. The site with the biggest differences between the three measures was non-melanotic skin cancer (fig 4). On examining the NRSR estimates, it seems that social class IV has a distinct survival disadvantage compared with social class I but this difference nearly disappears when the SRSR is used as the measure of survival.

The life table regression model with predictor variables social class, age, sex, follow up, and an age by follow up interaction provided an adequate fit to the data for all sites except breast cancer, where the p value of the goodness of fit test statistic (log-likelihood ratio statistic, also called scaled deviance (D)[60]) was 0.03 for the NRSR, and 0.07 for the SRSR and CSR. Adding a social class by follow up interaction term provided an adequate fit to the data ($D=28$, $df=30$, $p=0.57$). The interaction effect was that social class IV had the lowest survival in the first two years of follow up but the best survival in follow up years 4 and 5. There were statistically significant ($p < 0.01$) differences in survival by social class for eight of the 10 sites studied. The differences were not statistically significant (social class was not a significant factor in the model at the $\alpha=0.1$ significance level) for cancers of the pancreas and skin. The statistical significance of each factor in the model was similar for each of the three measures of survival. As expected, the level of social class variation in survival was smaller in

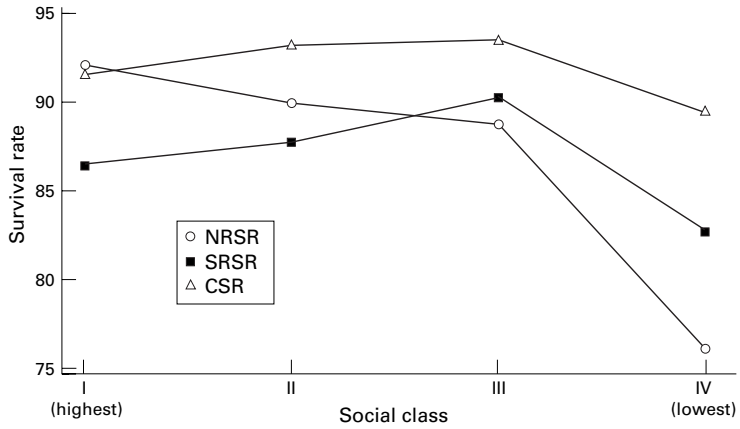

Figure 4 Male skin cancer (non-melanoma), social class specific five year survival rates calculated using three different measures of survival.

KEY POINTS

- The association between general mortality and social class needs to be considered when comparing cancer patient survival between social classes.
- Social class specific estimates of the relative survival rates (RSRs) should be adjusted for social class differences in general mortality.
- When death certificates are accurate, corrected (cause specific) survival rates provide similar estimates to the RSRs.
- The RSRs with the social class adjustment are recommended, especially where there are doubts concerning the accuracy of death certificates.

the SRSRs than the NRSRs, although social class was still a significant factor in the model using the SRSRs for the same eight sites.

The social class differences in survival apparent in the survival rates (figures 1–4) can also be seen in the relative risks (RRs) of excess death resulting from a diagnosis of cancer (table 4). Confidence intervals (CIs) are given for the RRs based on the SRSRs to give an indication of the level of uncertainty in the estimated RRs. The widths of the CIs are nearly identical for the NRSRs and the SRSRs because the width of the CI depends on the number of deaths and the estimated level of survival. The CIs for the CSRs are slightly narrower because the survival times for patients dying of causes other than the cancer in question are treated as fixed in the calculation of the CSR when they are actually random. For most sites there is a clear trend in the RRs, with social class I having the best survival and social class IV the worst. A RR of 0.7 for social class I/IV indicates that the excess risk of death because of cancer is 30% less in social class I than it is in social class IV. It can also be seen that the level of social class variation is lower among the SRSRs than the NRSRs. For example, the RR for social class I/IV (best/worst) is closer to 1.0 in the SRSR analysis than the NRSR analysis for all sites. The amount of social class variation in the CSRs is similar to that of the SRSRs.

Table 4 Estimated relative risks of excess mortality because of cancer for models fitted to each of the three measures of survival along with 95% confidence intervals for the relative risk based on the SRSR. Estimates are from the model containing follow up time, social class, sex (where applicable), age, and an age by follow up interaction

Site	*Parameter*	*CSR*	*NRSR*	*SRSR*	*95% CI*
Stomach	I/IV	0.78	0.76	0.79	(0.72,0.86)
	II/IV	0.86	0.86	0.87	(0.81,0.93)
	III/IV	0.90	0.89	0.90	(0.85,0.96)
Rectum	I/IV	0.64	0.57	0.62	(0.53,0.72)
	II/IV	0.76	0.68	0.72	(0.64,0.81)
	III/IV	0.85	0.80	0.82	(0.74,0.90)
Pancreas	I/IV	0.90	0.89	0.90	(0.79,1.01)
	II/IV	0.99	0.99	1.00	(0.90,1.10)
	III/IV	0.98	0.98	0.98	(0.89,1.08)
Lung, trachea	I/IV	0.97	0.95	0.97	(0.90,1.05)
	II/IV	0.94	0.92	0.93	(0.88,0.99)
	III/IV	1.01	0.99	1.00	(0.96,1.05)
Breast	I/IV	0.73	0.65	0.71	(0.61,0.82)
	II/IV	0.80	0.79	0.84	(0.75,0.94)
	III/IV	0.90	0.92	0.95	(0.85,1.07)
Cervix uteri	I/IV	0.59	0.55	0.59	(0.41,0.85)
	II/IV	0.83	0.87	0.89	(0.74,1.08)
	III/IV	0.80	0.75	0.76	(0.64,0.91)
Corpus uteri	I/IV	0.60	0.55	0.62	(0.45,0.86)
	II/IV	0.67	0.66	0.71	(0.56,0.90)
	III/IV	0.90	0.96	0.99	(0.80,1.23)
Kidney	I/IV	0.76	0.75	0.79	(0.67,0.92)
	II/IV	0.83	0.81	0.83	(0.73,0.95)
	III/IV	0.92	0.93	0.95	(0.84,1.07)
Urinary bladder	I/IV	0.45	0.43	0.52	(0.40,0.67)
	II/IV	0.67	0.67	0.74	(0.62,0.88)
	III/IV	0.73	0.73	0.77	(0.66,0.89)
Skin (non-melanoma)	I/IV	0.94	0.46	0.89	(0.42,1.89)
	II/IV	0.83	0.54	0.89	(0.53,1.47)
	III/IV	0.78	0.66	0.87	(0.57,1.32)

Discussion

Although there have been some conflicting studies for individual sites, it is clear that the higher social classes have a survival advantage for many cancers. Methodological limitations have been listed as a possible factor contributing to these differences.[25 61–63] The two main methodological limitations are in the choice of the measure of social class and the method of estimating survival. We have used one of the strongest possible measures of social class, a four category scale based on individual occupation data.[1] Occupation provides a sound basis for assessing social status as it is relatively stable over time, yet offers more variation than measures based on education.[55 56 64] The data maintained by the Finnish Cancer Registry are of high quality, so are ideal for examining differences between various measures of survival.

There are many factors that are thought to contribute to the social class differences in cancer patient survival. The main factors are social class differences in tumour biology, and delay in diagnosis, treatment, and host resistance.[29] Discussion of the mechanisms leading to social class differences in cancer patient survival is, however, outside the scope of this paper. Detailed coverage of this topic can be found in many other sources.[1 10 29 62 65 66]

Each of the survival measures presented in this paper adjusts the observed survival rate to reflect true social class differences in cancer patient survival. To effectively interpret the results, one must understand what adjustments are being made and any factors that may bias the results. The NRSRs are adjusted for the expected survival of a comparable group in the general population stratified by age, sex, and calendar period of diagnosis. Social class

differences in NRSRs could result from all types of excess mortality such as that from the cancer in question, treatment complications, or unrelated causes of death. CSRs ignore non-cancer deaths even though some could be related to treatment and, as such, should be considered in the outcome. In addition, corrected rates are dependent on the accuracy of cause of death information, which may vary by social class.[51]

Our results indicate that NRSRs overestimate social class differences in survival. Therefore, when studying cancer patient survival by social class, we recommend adjusting for social class differences in general mortality, producing what we have called SRSRs. In addition to the adjustments inherent in the NRSR, the SRSR is adjusted for some lifestyle factors, such as the higher mortality in the lower social classes resulting from the higher prevalence of tobacco and alcohol use.[67–69] The results obtained using CSRs were close to those obtained using SRSRs, suggesting a high level of accuracy in the coding of cause of death in Finland. Thus, CSRs can be regarded as a good measure of social class specific survival in Finland. It is difficult, however, to generalise this conclusion. We cannot unconditionally recommend CSRs for use in all countries as they are heavily dependent on the accuracy of cause of death data, which varies between countries. If cause of death information is considered to be reliable, and it is not possible to estimate the SRSR, we would suggest using the CSR as the next best option to the SRSR.

We could possibly expect the CSR to be systematically higher than the NRSR and the SRSR for smoking related cancers (lung and bladder). This is because the expected survival estimates based on general population mortality do not adjust for the additional smoking related mortality experienced by these patients (although the SRSRs will be adjusted in part because of the higher prevalence of smoking in the lower social classes). This was not however the case; the CSRs did not differ systematically from the NRSRs and SRSRs for cancers of the bladder or lung. We cannot offer any explanation for this finding, other than the fact that any correction is small during the short (five year) follow up time.

However, we found that the CSRs for cancer of the kidney and non-melanotic skin cancer were systematically higher than the NRSRs and SRSRs. This situation could occur if some cancer deaths among these patients were systematically recorded as non-cancer on death certificates, thereby artificially increasing the CSR. We can see no obvious reason why this would occur for cancers of the kidney and skin but not in other sites. We would also observe systematically higher CSRs if patients presenting with cancer of the kidney or skin experienced systematically higher mortality than the general population for causes other than the cancer in question. Some risk factors for cancer of the kidney, most importantly smoking and obesity, also increase the risk of other diseases, which could account for the finding. For non-melanotic skin cancer, an increased risk of lymphoma has been reported.[70] This is probably because ultraviolet radiation is a risk factor for both of these cancers. It is possible that this could partially explain the systematically higher corrected survival rates for patients first diagnosed with non-melanotic skin cancer. Excess mortality resulting from skin cancer is extremely low, so the additional mortality from subsequent lymphomas will, although also low, have a greater effect on the RSRs for skin cancer patients than for other sites.

We also cannot offer a plausible explanation as to why the CSRs for cancer of the stomach and rectum were systematically lower than the NRSR and SRSR. This situation could occur if these patients experienced systematically lower mortality than the general population for causes other than the cancer in question. It could also occur if non-cancer deaths were incorrectly classified as cancer deaths. We can see no reason why either of these two scenarios would occur for cancers of the stomach and rectum but not for other cancers.

The NRSR will also be slightly biased when estimating survival for all social classes combined, because of the dependence of incidence on social class for some cancers.[71] For example, breast cancer is more common in the upper social classes, with a standardised incidence ratio (SIR) of 1.39 for females in social class I, while stomach cancer is more common in the lower social classes (SIR=1.20 for men and 1.12 for women in social class IV). All cause death rates for the general population are therefore not strictly representative of the cancer cases for calculating expected survival, and the resulting NRSRs will be slightly biased. We found this bias to be extremely small. The differences in the estimated SRSRs and NRSRs were so small that either of these measures can be used for estimating cancer patient survival at the aggregate level (all social classes combined).

As expected, there were larger differences between the survival measures for social class specific rates. If a single measure is to be used for estimating social class specific survival, we recommend the SRSR, RSRs with expected survival adjusted for social class differences in general mortality. If this is not an available option, we recommend using the CSR (assuming cause of death information is considered to be accurate). RSRs without the social class correction overestimated social class differences and should be used with caution where social class specific estimates of survival are required. Observed survival rates are the measure most prone to bias and we do not recommend their use in assessing social class differences in survival from cancer.

Funding: the work of Paul Dickman was financially supported by the Swedish Cancer Foundation. The work of Anssi Auvinen was financially supported by the Finnish Cancer Society and the Duodecim Society. The study was partly financed by the Ida Montin Fund.

Conflicts of interest: none.

Appendix 1

ESTIMATING SOCIAL CLASS SPECIFIC ANNUAL DEATH PROBABILITIES

The statistical software package for estimating relative survival developed by Hakulinen and Abeywickrama[57] requires an input file containing annual probabilities of death (or survival) and life expectancy of the population from which the patients are drawn stratified by age, sex, and calendar year. Other software packages require similar input files in order to estimate relative survival. This information is usually tabulated for the national population but we require it individually for each social class. In this study, we have estimated social class specific annual probabilities of death directly from case specific (individual level) mortality data. If these data are not available, it is possible to estimate the social class specific annual probabilities of death by applying social class specific "standardised" mortality ratios (SMR) by sex and age to the nationwide probabilities of death.

We denote the SMR for social class r as

$$c_r = \frac{\mu_r}{\mu_{wc}} \tag{1}$$

where μ_r is the death rate in social class r and μ_{wc} is the death rate rate for the whole country. The subscripts for age, sex, and time period are omitted from c_r, μ_r and μ_{wc}.

The probability of surviving a given age sex time period specific annual interval (from x to x+1 years) for a member of the general population[72] is given by $p_{wc}=e^{-\mu_{wc}}$. The corresponding probability for a member of the general population in social class r is given by

$$\begin{aligned} p_r &= \exp\left[-\int_x^{x+1} \mu_r dt\right] \\ &= e^{-\mu_r} \\ &= e^{-c_r \mu_{wc}} \text{ (from equation 1)} \\ &= [e^{-\mu_{wc}}]^{c_r} \\ &= [\,p_{wc}\,]^{c_r} \end{aligned} \tag{2}$$

The subscripts for age, sex, and period are also omitted from p_{wc} and p_r. Note that p_{wc} and p_r are calculated for annual intervals while the SMRs (c_r) usually cover five year intervals. Note that explicit values for μ_r and μ_{wc} are not required.

We have q_{wc} and require q_r, the annual probability of death in social class r, which is given by

$$\begin{aligned} q_r &= 1 - p_r \\ &= 1 - [p_{wc}]^{c_r} \\ &= 1 - [1 - q_{wc}]^{c_r} \end{aligned} \tag{3}$$

1 Auvinen A, Karjalainen S, Pukkala E. Social class and cancer patient survival in Finland. *Am J Epidemiol* 1995;**142**:1089–102.
2 Schrijvers CT, Mackenbach JP, Lutz JM, *et al*. Deprivation, stage at diagnosis and cancer survival. *Int J Cancer* 1995;**63**:324–9.
3 Schrijvers CT, Mackenbach JP, Lutz JM, *et al*. Deprivation and survival from breast cancer. *Br J Cancer* 1995;**72**:738–43.
4 Schrijvers CT, Coebergh JW, van der Heijden LH, *et al*. Socioeconomic status and breast cancer survival in the southeastern Netherlands, 1980–1989. *Eur J Cancer* 1995; **31A**:1660–4.
5 Schrijvers CT, Coebergh JW, van der Heijden LH, *et al*. Socioeconomic variation in cancer survival in the southeastern Netherlands, 1980–1989. *Cancer* 1995;**75**:2946–53.
6 Schrijvers CT, Mackenbach JP. Cancer patient survival by socioeconomic status in seven countries: a review for six common cancer sites. *J Epidemiol Community Health* 1994; **48**:441–6.
7 Lamont DW, Symonds RP, Brodie MM, *et al*. Age, socio-economic status and survival from cancer of cervix in the West of Scotland 1980–87. *Br J Cancer* 1993;**67**: 351–7.
8 Gordon NH, Crowe JP, Brumberg DJ, *et al*. Socioeconomic factors and race in breast cancer recurrence and survival. *Am J Epidemiol* 1992;**135**:609–18.
9 Kato I, Tominaga S, Ikari A. The role of socioeconomic factors in the survival of patients with gastrointestinal cancers. *Jpn J Clin Oncol* 1992;**22**:270–7.
10 Auvinen A. Social class and colon cancer survival in Finland. *Cancer* 1992;**70**:402–9.
11 Brenner H, Mielck A, Klein R, *et al*. The role of socioeconomic factors in the survival of patients with colorectal cancer in Saarland/Germany. *J Clin Epidemiol* 1991;**44**:807–15.
12 Kogevinas M, Marmot MG, Fox AJ, *et al*. Socioeconomic differences in cancer survival. J *Epidemiol Community Health* 1991;**45**:200–3.
13 Karjalainen S, Pukkala E. Social class as a prognostic factor in breast cancer survival. *Cancer* 1990;**66**:819–26.
14 Murphy M, Goldblatt P, Thornton-Jones H, *et al*. Survival among women with cancer of the uterine cervix: Influence of marital status and social class. *J Epidemiol Community Health* 1990;**44**:293–6.
15 Kogevinas M. *Longitudinal study: socio-demographic differences in cancer survival.* LS no 5. London: Office of Population Censuses and Surveys, 1990.
16 Vågerö D, Persson G. Cancer survival and social class in Sweden. *J Epidemiol Community Health* 1987;**41**:204–9.
17 Dayal HH, Polissar L, Dahlberg S. Race, socioeconomic status, and other prognostic factors for survival from colorectal cancer. *J Chron Dis* 1987;**40**:857–64.
18 Bassett MT, Krieger N. Social class and black-white differences in breast cancer survival. *Am J Public Health* 1986;**76**:1400–3.
19 Keirn W, Metter G. Survival of cancer patients by economic status in a free care setting. *Cancer* 1985;**55**:1552–5.
20 Dayal HH, Polissar L, Dahlberg S. Race, socioeconomic status, and other prognostic factors for survival from prostate cancer. *J Natl Cancer Inst* 1985;**49**:2208–16.
21 Chirikos TN, Homer RD. Economic status and survivorship in digestive system cancers. *Cancer* 1985;**56**:210–17.
22 Chirikos TN, Reiches NA, Moeschberger ML. Economic differentials in cancer survival: A multivariate analysis. *J Chron Dis* 1984;**37**:183–93.
23 Bonett A, Roder D, Esterman A. Determinants of case survival for cancers of the lung, colon, breast and cervix in South Australia. *Med J Aust* 1984;**141**:705–9.
24 Dayal HH, Power RN, Chiu C. Race and socio-economic status in survival from breast cancer. *J Chron Dis* 1982;**35**: 675–83.
25 Berg JW, Ross R, Latourette HB. Economic status and survival of cancer patients. *Cancer* 1977;**39**:467–77.
26 Linden G. The influence of social class in the survival of cancer patients. *Am J Public Health and the Nations Health* 1969;**59**:267–74.
27 Lipworth L, Abelin T, Connelly RR. Socio-economic factors in the prognosis of cancer patients. *J Chron Dis* 1969;**23**:105–16.
28 Wilkinson RG. Income distribution and life expectancy. *BMJ* 1992;**304**:165–8.
29 Leon D, Wilkinson RG. Inequalities in prognosis: socioeconomic differences in cancer and heart disease survival. In: Fox J, ed. *Health inequalities in European countries.* Aldershot: Gower, 1989; 280–300.
30 Hogg RS, Strathdee SA, Craib KJ, *et al*. Lower socioeconomic status and shorter survival following HIV infection. *Lancet* 1994;**344**:1120–4.
31 Cayla JA, Artazcoz L, Iglesias B, *et al*. Epidemiology of AIDS in Barcelona (1981–1991) (II). A study of mortality and survival. [In Spanish]. *Med Clin (Barc)* 1994;**102**:129–35.
32 Curtis JR, Patrick DL. Race and survival time with AIDS: A synthesis of the literature. *Am J Public Health* 1993;**83**: 1425–8.
33 Cohart EM. Socioeconomic distribution of cancer of the female sex organs in New Haven. *Cancer* 1955;**8**:34–41.
34 Cohart EM. Socioeconomic distribution of cancer of the lung in New Haven. *Cancer* 1955;**8**:1126–9.
35 Cohart EM. Socioeconomic distribution of cancer of the gastrointestinal tract in New Haven. *Cancer* 1955;**8**:379–88.
36 Berrino F, Estève J, Coleman MP. Basic issues in estimating and comparing the survival of cancer patients. In: Berrino F, Sant M, Verdecchia A, *et al*, eds. *Survival of cancer patients in Europe: The EUROCARE Study.* IARC Scientific Publications no 132. Lyon: International Agency for Research on Cancer, 1995:1–14.
37 Henson DE, Ries LA. The relative survival rate. *Cancer* 1995;**76**:1687–8.
38 Estève J. Benhamou E, Raymond L. *Statistical methods in cancer research. Vol IV: Descriptive epidemiology.* IARC Scientific Publications no 128. Lyon: International Agency for Research on Cancer, 1994.

39 Parkin DM, Hakulinen T. Analysis of survival. In: Jensen OM, Parkin DM, MacLennan R, *et al*, eds. *Cancer registration: principles and methods.* IARC Scientific Publications no 95. Lyon: International Agency for Research on Cancer, 1991:159–76.

40 Percy CL, Stanek E, Gloeckler L. Accuracy of cancer death certificates and its effect on cancer mortality statistics. *Am J Public Health* 1981;**71**:242–50.

41 Goldacre MJ. Cause-specific mortality: understanding uncertain tips of the disease iceberg. *J Epidemiol Community Health* 1993;**47**:491–6.

42 Ederer F, Axtell LM, Cutler SJ. The relative survival rate: A statistical methodology. *National Cancer Institute Monograph* 1961;**6**:101–21.

43 Berkson J, Gage RP. Calculation of survival rates for cancer. *Proceedings of Staff Meetings of the Mayo Clinic* 1950;**25**: 270–86.

44 Pappas G, Queen S, Hadden W, *et al.* The increasing disparity in mortality between socioeconomic groups in the United States, 1960 and 1986. *N Engl J Med* 1993;**329**: 103–9.

45 Smith GD, Leon D, Shipley MJ, *et al*. Socioeconomic differentials in cancer among men. *Int J Epidemiol* 1991;**20**:339–45.

46 Smith GD, Shipley MJ, Rose G. Magnitude and causes of socioeconomic differentials in mortality: Further evidence from the Whitehall study. *J Epidemiol Community Health* 1990;**44**:265–70.

47 Fox AJ, Goldblatt PO, Jones DR. Social class mortality differentials: artefact, selection or life circumstances? *J Epidemiol Community Health* 1985;**39**:1–8.

48 Fox AJ, Jones DR, Goldblatt PO. Approaches to studying the effect of socio-economic circumstances on geographic differences in mortality in England and Wales. *Br Med Bull* 1984;**40**:309–14.

49 Fox AJ. *Longitudinal study: socio-demographic mortality differentials 1971–1975.* LS no 1. London: Office of Population Censuses and Surveys, 1982.

50 Valkonen T, Martelin T, Notkola V, *et al*. *Socio-economic mortality differences in Finland 1981–90.* Helsinki: Statistics Finland, 1993.

51 Samphier ML, Robertson C, Bloor MJ. A possible artefactual component in specific cause mortality gradients: social class variations in the clinical accuracy of death certificates. *J Epidemiol Community Health* 1988;**42**: 138–43.

52 Teppo L, Pukkala E, Lehtonen M. Data quality and quality control of a population-based cancer registry. Experience in Finland. *Acta Oncol* 1994;**33**:365–9.

53 Hakulinen T. Health care system, cancer registration and follow-up of cancer patients in Finland. In: Berrino F, Sant M, Verdecchia A, *et al,* eds. *Survival of cancer patients in Europe: The EUROCARE Study.* IARC Scientific Publications no 132. Lyon: International Agency for Research on Cancer, 1995:53–4.

54 Central Statistical Office of Finland. *Population census 1970. Vol 9. Occupation and social position*. Helsinki: Central Statistical Office of Finland, 1974.

55 Ranhala U. *The social stratification of Finnish society.* [In Finnish]. Porvoo: Society for Social Policy, 1966.

56 Central Statistical Office of Finland. *Classification of socio-eonomic satus.* Handbook no 17. [In Finnish].Helsinki: Central Statistical Office of Finland, 1989.

57 Hakulinen T, Abeywickrama K. A computer program package for relative survival anal*ysis. Computer Programs in Biomedicine* 1985;**19**:197–207.

58 Hakulinen T, Tenkanen L. Regression analysis of relative survival rates. *Applied Statistics* 1987;**36**:309–17.

59 Francis B, Green M, Payne C, eds. *The GLIM system: generalized linear interactive modelling.* Oxford: Oxford University Press, 1993. Release 4.

60 Nelder JA, Wedderburn RWM. Generalized linear models. *Journal of the Royal Statistical Society A* 1972;**135**:370–84.

61 Greenwald HP, Borgatta EF, McCorkle R, *et al*. Explaining reduced cancer survival among the disadvantaged. *Milbank Q* 1996;**74**:215–38.

62 Auvinen A, Karjalainen S. Possible explanations for social class differences in cancer patient survival. In: Kogevinas M, Pearce N, Susser M, *et al.* eds. *Social inequalities and cancer.* IARC Scientific Publications no 138. Lyon: International Agency for Research on Cancer, 1997: 377–97.

63 Greenwald HP, Polissar NL, Borgatta EF, *et al*. Detecting survival effects of socioeconomic status: problems in the use of aggregate measures. *J Clin Epidemiol* 1994;**47**:903–9.

64 Liberatos P. Link BG, Kelsey JL. The measurement of social class in epidemiology. *Epidemiol Rev* 1988;**10**:87–121.

65 Karjalainen S. *Equity and cancer patient survival.* [PhD Thesis]. Tampere: University of Tampere, Finland, 1991. Acta Universitatis Tamperensis, ser A vol 316.

66 Schrijvers C. *Socioeconomic inequalities in cancer survival in the Netherlands and Great Britain: small-area based studies using cancer registry data.* [PhD Thesis]. Rotterdam: Erasmus University, 1996.

67 Lowry R. Kann L, Collins JL, *et al*. The effect of socioeconomic status on chronic disease behaviour among US adolescents. *JAMA* 1996;**276**:792–7.

68 Rahkonen O, Berg MA, Puska P. Relationship between educational status, gender and smoking in Finland, 1978–1992. *Health Promotion International* 1995;**10**:115–20.

69 Luoto R, Pekkanen J, Uutela A, *et al*. Cardiovascular risks and socioeconomic status: Differences between men and women in Finland. *J Epidemiol Community Health* 19;**48**: 348–54.

70 Teppo L, Pukkala E, Saxén E. Multiple cancer—an epidemiologic exercise in Finland. *J Natl Cancer Inst* 1985; **75**:207–17.

71 Pukkala E. *Cancer risk by social class and occupation: A survey of 109,000 Finns of working age. Vol 7. Contributions to epidemiology and biostatistics.* Basel: Karger, 1995.

72 Chiang CL. *Introduction to stochastic processes in biostatistics.* New York: John Wiley, 1968.