J Chron Dis 1977, Vol. 30, pp. 431–443. Pergamon Press. Printed in Great Britain

ON LONG-TERM RELATIVE SURVIVAL RATES

T. HAKULINEN

Finnish Cancer Registry, Liisankatu 21 B, SF-00170 Helsinki 17, Finland, and Department of Statistics, University of Helsinki, SF-00100 Helsinki 10, Finland

(*Received in revised form* 20 *November* 1976)

Abstract—The relative survival rate over a given period has been defined as the ratio of the proportion of survivors over the period in the patient group, to the proportion of survivors expected in a similar group of persons without the disease. If there are found consecutive relative survival rates, all calculated from the beginning of follow-up up to points separated by one year, which increase with lengthening of the interval after a specific year, it has been interpreted that the patients are affected by a smaller mortality then expected [2] after that year. An example in this paper demonstrates that this may not necessarily be true, particularly when a long-term follow-up of a heterogeneous patient population with respect to the expected survival is involved. Furthermore, the relative survival rate has been interpreted [1, 5] as the proportion of survivors, provided the only disease capable of killing patients is that of the patients. It is also shown that this interpretation is impossible for long-term relative survival rates in a heterogeneous population.

If mortality during a specific part of follow-up is of interest in a heterogeneous population, it is suggested that annual relative survival rates be examined instead of drawing a relative survival curve. If the long-term survival due to a specific disease of the patients in a heterogeneous population is desired, methods of the theory of competing risks, and the examination of relative survival curves by age and by other background variables, which make the patient groups homogeneous with respect to the expected survival, are suggested as alternative or auxiliary methods.

THE RELATIVE SURVIVAL RATE AND CURVE, DEFINITION AND INTERPRETATION

As A RULE, the evaluation of patient survival has been based upon survival rates [1–3]. Actually, the survival rate of a group of patients over a specified time interval is not a rate, but a *proportion* [4] of patients alive at the end of the interval with respect to the patients alive at its beginning.*

Commonly in use is a curve which indicates the observed survival rates over periods that begin from the diagnosis or the time of treatment (i.e. the beginning of the follow-up) [2, 3, 6]. A point that indicates the observed survival rate from the beginning to the end of the first year of the follow-up is placed at 1 yr, that which indicates the observed survival rate from the beginning of follow-up to the end of the 2nd year at 2 yr, and so on. The successive points are connected to form a line diagram which is known as the observed survival curve.

* As in this case the term 'rate' belongs to standard terminology [1–3, 5], it is here employed instead of the correct term.

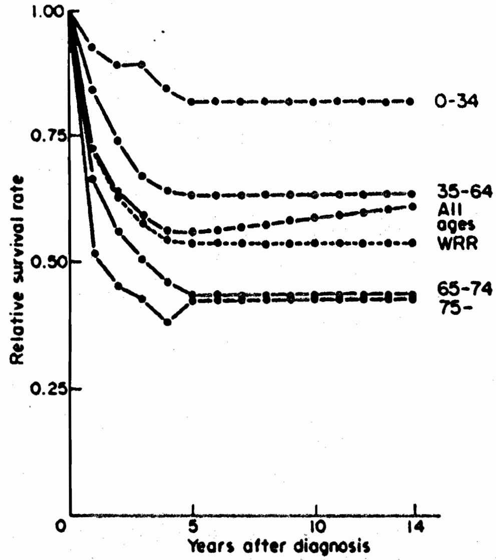

FIG. 1. A hypothetical example. The relative survival curves by age at diagnosis for female patients with localized cancer of the colon or small intestine in Finland 1953–1970, on the assumption that the 5-yr relative survival rate in each age group is the cure rate for that age group. (WRR: weighted average of the age-specific relative survival curves, with the age distribution of patients at diagnosis being used as constant weights.)

In medical follow-up studies, the observed survival rates are often insufficient to describe patient survival when the primary interest is attached to a specific disease of the patients and some of the patients die from other diseases. In order to eliminate the effect of mortality from other diseases on survival rates, the concept of relative survival rate has been introduced [1, 5]. A relative survival rate has been *defined* [1, 5] as the *ratio** of the observed survival rate in the patient group, to the survival rate expected in a group similar to the group of patients at the beginning of the interval with respect to all of the possible factors affecting the survival, except the disease under study. The relevant factors recommended to be taken into account are sex, race, age, calendar period and domicile [5].

Actually, the relative survival rate thus is not a rate, but a ratio between two proportions (here termed rates). On the other hand, the relative survival rate has been *interpreted* [1, 5] as the proportion of patients alive at the end of the interval with respect to the patients alive at the beginning of the interval, provided that the only disease capable of killing patients is that being studied. Consequently, if proportions are termed rates, the relative survival rate also can be considered to be a rate. The basis for this interpretation is a model which assumes that the patients are subject to two causes of death: that under study, and other causes which act independently of each other [1, 5].

The above interpretation provides a basis for drawing, by the application of rules for drawing the observed survival curve, and relative survival rates, an

* As in this case the term 'rate' belongs to standard terminology [1–3, 5], it is here employed instead of the correct term.

observed survival curve for patients in the fictitious situation in which only the disease under study can kill patients [2]. The resulting line diagram is here called the *relative survival curve.*

Let us next suppose that in this fictitious situation all of the patients alive at t yr after the beginning of the follow-up study are cured of the disease under study. As this disease was the only one capable of killing individuals, none of the patients will die after t yr of follow-up, and consequently after that point the relative survival curve will be a horizontal line. The t-yr relative survival rate is termed the cure rate [7], viz. the proportion of cured patients among those patients alive at the beginning of the follow-up (Fig. 1, disregard the curve for 'all ages').

With the above interpretation of the relative survival rate as a proportion, in principle, the relative survival curve is, either declining or horizontal with increasing follow-up time. In the following, an example is given in which this does not hold true. Reasons for this will be considered.

EXAMPLE OF A RELATIVE SURVIVAL CURVE RISING DESPITE OF PATIENTS HAVING HIGHER MORTALITY THAN EXPECTED

Patients with localized cancer of the colon or small intestine diagnosed in Finland during the years 1953–1970 [8] are used as an example (Table 1). The material

TABLE 1. LIFE TABLE FOR FEMALE PATIENTS WITH LOCALIZED CANCER OF THE COLON OR SMALL INTESTINE IN FINLAND 1953–1970†

i	l_i	d_i	w_i	p_i	$_{i+1}p_0$	p_i^*	$_{i+1}p_0^*$	r_i	$_{i+1}r_0$
0	1247	364	69	0.6998	0.6998	0.9607	0.9607	0.7285	0.7285
1	814	122	79	0.8425	0.5896	0.9682	0.9216	0.8702	0.6397
2	613	70	50	0.8810	0.5194	0.9678	0.8826	0.9103	0.5884
3	493	42	53	0.9100	0.4726	0.9663	0.8441	0.9417	0.5599
4	398	18	36	0.9526	0.4502	0.9663	0.8060	0.9858	0.5586
5	344	20	30	0.9392	0.4229	0.9672	0.7685	0.9711	0.5503
6	294	14	37	0.9492	0.4014	0.9661	0.7315	0.9825	0.5487
7	243	9	29	0.9606	0.3856	0.9676	0.6953	0.9927	0.5545
8	205	9	22	0.9536	0.3677	0.9676	0.6598	0.9855	0.5572
9	174	6	21	0.9633	0.3542	0.9666	0.6253	0.9966	0.5664
10	147	6	29	0.9547	0.3381	0.9639	0.5917	0.9905	0.5715
11	112	5	18	0.9515	0.3217	0.9674	0.5590	0.9835	0.5755
12	89	3	18	0.9625	0.3097	0.9690	0.5274	0.9933	0.5871
13	68	3	12	0.9516	0.2947	0.9654	0.4970	0.9858	0.5929
14	53								

†l_i = number of patients alive at i yr after diagnosis.
d_i = number of patients died between i and $i+1$ yr after diagnosis.
w_i = number of patients withdrawn alive between i and $i+1$ yr after diagnosis.
p_i = 1-yr observed survival rate for patients alive at i yr after diagnosis.
$_{i+1}p_0$ = $i+1$-yr observed survival rate for patients alive at diagnosis.
p_i^* = 1-yr expected survival rate for patients alive at i yr after diagnosis.
$_{i+1}p_0^*$ = $i+1$-yr expected survival rate for patients alive at diagnosis.
r_i = 1-yr relative survival rate for patients alive at i yr after diagnosis.
$_{i+1}r_0$ = $i+1$-yr relative survival rate for patients alive at diagnosis.

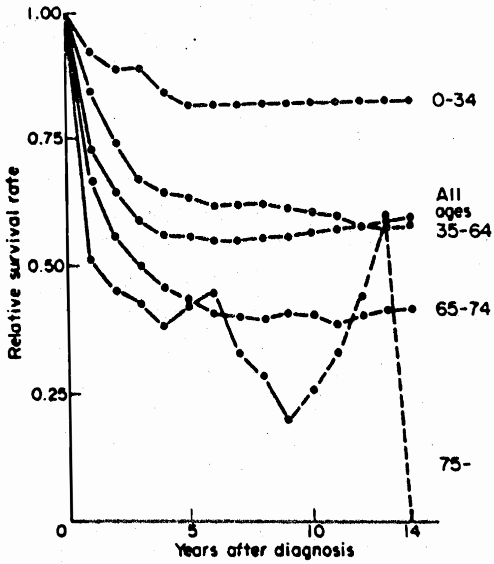

FIG. 2. The relative survival curves by age at diagnosis for female patients with cancer of the colon or small intestine in Finland 1953–1970. A broken line indicates rates based upon <5 patients.

was compiled by the Finnish Cancer Registry, which is a cancer registry that covers the whole of Finland, and which has been in operation since 1952 [9]. Cases based upon death certificate or autopsy only have been excluded from the material.

Life tables were calculated for the whole group (Table 1) and for 4 age groups (Appendix table). The observed rates were obtained by application of the actuarial method [10]. The expected survival rates were calculated by means of the 'exact' method ([5], p. 110), viz. by averaging the expected survival rates for individual patients alive at the beginning of the period concerned, with sex, age and calendar period taken as factors which specify the life table of the general (approximately non-diseased [1, 5]) population [11]. The relative survival rates were obtained as ratios of the observed and expected rates. Annual (1-yr) survival rates, observed and expected, were calculated, using as material the patients alive at the beginning of the follow-up year concerned. Annual relative survival rates were obtained as ratios of the annual observed and expected survival rates.

As the annual relative survival rates (quantities r_i, Table 1) were below 1 during the 14-yr follow-up, the mortality observed exceeded that expected in every follow-up year. However, the relative survival curve was rising after 6 yr of follow-up (quantities ${}_{i+1}r_0$ in Table 1, the curve for all ages in Fig. 2). Consequently, in this example a rising relative survival curve was found, and even in a situation in which the patients were still affected by excess mortality.

REASONS FOR A RISE IN THE RELATIVE SURVIVAL CURVE. THE AIM OF THE STUDY

The traditional explanation for a rising relative survival curve is that the group of patients is affected by a mortality less than that expected ([2], p. 2). This may occur if the expected mortality is very low, and the number of patients is so small that the most probable number of patients dying is zero ([2], the patients aged 0–34 yr in Appendix table and Fig. 2). This may also occur when improved medical care of the patients, or the frequent use of medical facilities by patients, prevents them from dying from diseases other than that of their primary illness [8].

Another possible explanation is that the expected survival rate is not the survival rate of a similar, non-diseased population. In a study on the causes of death of breast-cancer patients [8], a possible selection of the patients with respect to social class was introduced as one of the possible reasons for mortality from 'other causes of death' which was less than that expected among patients aged 75 or more at diagnosis. Those of the higher social classes may experience less mortality than those of lower ones. As in general life tables by social class are not available, for a patient group which mainly consists of patients belonging to high social classes, the expected survival rates are less than the appropriate ones which results in relative survival rates that are too high. In this case, the disease under study and other diseases do not act independently of each other.

Nevertheless, these explanations do not necessarily apply to the above example, as the mortality observed among the patients was higher than that expected. The aim of this paper is to show that a rising relative survival curve may also result from a methodological effect, particularly when it is a matter of the long-term survival of a heterogeneous population. In what follows, possible means of improving the result and alternative methods for the avoidance of this effect are discussed and related to interpretation of the relative survival rate.

THE EFFECT OF A HETEROGENEOUS POPULATION WITH RESPECT TO AGE ON THE RELATIVE SURVIVAL CURVE

The relative survival rate was defined as the ratio between the observed and expected survival rates. Let us think of a relative survival rate over a long period of time, e.g. 30 yr. Subsequent to 30 yr of follow-up, those people expected to be alive are virtually the young patients at the beginning of follow-up. The old patients have died, whether cured or not, and are expected to have died. Consequently, a long-term relative survival rate of a population is that of its young age groups. This is formally shown in the Appendix, which further demonstrates that the relative survival rate for the whole group of patients is expressible as a weighted average of the age-specific relative survival rates, but with weights that are not constant. With increase in the period of follow-up, the relative survival rate for the whole group gradually converges towards that of young patients.

As an example let us examine a simplified version of the example discussed above. Let us assume that patients with localized cancer of the colon or small intestine are cured, provided that they have survived up to 5 yr subsequent to diagnosis, and that after this the mortality of each age group is exactly that expected (Fig. 1). Accordingly, the relative survival curves for each age group are horizontal after 5 yr of follow-up. As in this example the relative survival rate for young patients (ages 0–34) is higher than the average of relative survival rates (WRR), the relative survival curve for all patients rises from this average towards the relative survival curve for ages 0–34. This convergence brings about a rising relative survival curve for the whole group after 5 yr of follow-up.

If the relative survival curve for all ages shown in Fig. 1 were the only one available, the false interpretation might be made that the patients are affected by a mortality which is lower than that expected. In order to make the relative survival curve for all ages horizontal after 5 yr of follow-up, in each age group the annual observed survival rate should be about 99% of that expected. To preclude misinterpretation, accordingly, relative survival curves should be avoided. and annual relative survival rates presented when, in a heterogeneous population, the relative survival rates have a gradient with respect to age at the beginning of follow-up.

Heterogeneity does not necessarily have to be that with respect to age. A similar effect will be produced by any factor, e.g. sex, race or calendar period which is taken into account in the determination of expected survival rates. The relative survival rate for the whole group will converge towards that of persons with the highest expected survival.

It might be asked whether the situation described above is common enough to warrant attention. In many other primary sites, e.g. lip, stomach, rectum, lung and corpus uteri, particularly as concerns localized tumours, the relative survival rates diminish with increasing age at diagnosis [2, 3, 6]. Accordingly, alternatives to the relative survival rate, calculated as the ratio between the observed and expected survival rates, are worth considering.

ALTERNATIVES TO THE RELATIVE SURVIVAL RATE

Weighting age-specific observed survival rates

The 'exact' method of calculation [5] employed in this study involves computation of the expected survival rate as the average of the expected survival rates for individuals alive at the beginning of the period for which the relative survival rate is calculated. This is conceptually equal to calculation of the average of the expected survival rates for various age groups. On the other hand, the observed survival rate is obtained from a single life table designed for the whole group. This may be inadequate for determination of the survival rates for all patients, since a withdrawal alive (w_i, Table 1) is assumed to have a survival similar to that of patients on the average, after the year of withdrawal [12]. If the age distribution of patients admitted to the study changes with calendar time of admission or if for some other reason there occurs a change in survival by the time of the beginning of follow-up, the survival experience of a withdrawal is different from the average survival of patients living beyond the year of withdrawal [13].

The effect of a changed age distribution of patients was studied in the above example by the calculation of a life table for each age group (Appendix table), and the observed survival rates for the whole group by weighting the age-specific observed survival rates with weights proportional to the number of patients belonging to the age groups at diagnosis [cf. eqn (1) in the Appendix]. The rise in the resulting relative survival curve persisted, although its magnitude diminished (cf. WOR vs RSR in Fig. 3).

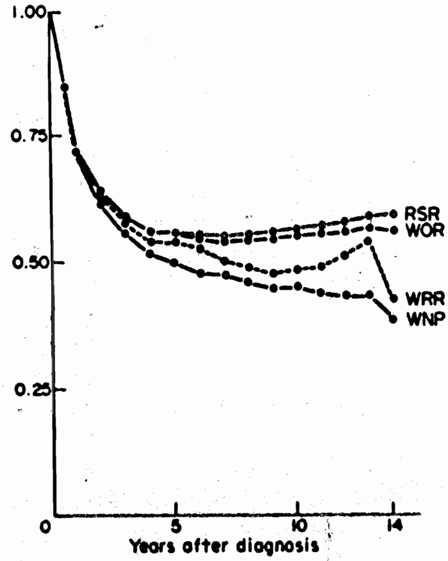

FIG. 3. Alternatives for the relative survival curve (RSR) for female patients with localized cancer of the colon or small intestine in Finland 1953–1970: WRR: weighted average of the age-specific relative survival curves; WOR: ratio between the weighted average of the age-specific observed survival rates and the expected survival rate; WNP: weighted average of the age-specific net probabilities of survival. The age distribution of patients at diagnosis has been used as weights in weighting of the age-specific quantities.

Weighting age-specific relative survival rates

Theoretically attractive results are obtained by weighting the age-specific relative survival curves by weights that do not vary with respect to the length of follow-up time (WRR in Fig. 1). This method provides constant relative survival rates when the corresponding rates in each age group are constant. Conceptually, when the relative survival rate is interpreted as the survival rate of patients, provided that only the disease of the patients is capable of killing, the correct solution is to weight the age-specific relative survival rates with weights that are proportional to the numbers of persons in each age group at the beginning of follow-up.

In practice, however, random variation begins to play an important role in the relative survival curve of an old age group (ages 75– in Fig. 2). As the weighting is constant, this also applies to the combined curve, provided that the proportion of old patients is not small (WRR in Fig. 3). It is not even quite certain that the relative survival curve for an old age group can be estimated throughout the entire period of follow-up. The last subinterval with individuals alive at the beginning may contain an individual withdrawing alive. In this case, the weighted average could in principle be calculated by solving the missing age-specific relative survival rate, e.g. ${}_ir_0(m)$ from eqn (2) in the Appendix and applying it in calculations. In practice, however, the solution does not exist when the expected survival rate for the *m*th age group is zero. Indeed, it is difficult, if not impossible, to answer a question which relates, e.g. to the 30-yr survival rate of 80-yr old patients, provided that only the disease of the patients is capable of killing.

Combining annual relative survival rates

Heise [14] has considered and rejected the calculation of cumulative relative survival rates by multiplication of the annual relative survival rates. By employment of the notation of Table 1, this means that

$$ {}_i r_0^{X} = \prod_{j=0}^{i-1} r_j = \prod_{j=0}^{i-1} (p_j/p_j^*). \tag{1} $$

Heise rejected this method, since the weights given to annual age-specific expected survival rates involved in the right hand side of eqn (1) varied from year to year, being proportional to the number of patients in each age group at the beginning of the follow-up year. A referee has presented the possibility of using weighted annual age-specific relative survival rates for the quantities r_j in (1), and employing the same weights as Heise used for weighting of the annual age-specific expected rates.

Both the method presented by Heise [14], and that presented by the referee, are conceptually different from both the relative survival rate and the alternatives already presented. In both methods, the structure of the patient population at the beginning of follow-up is not sufficient for determination of the expected survival rate for the group, i.e. the observed survival also affects that expected.

When eqn (1) is applied to the data in Table 1, a strictly decreasing relative survival curve is obtained, as all of the annual relative survival rates are <1. When the follow-up time becomes long, the difference between this curve and the relative survival curve in Table 1 becomes large. The 10-yr figure is 0.498, as compared with the 0.566 of Table 1, the 14-yr figures being 0.475 and 0.593 respectively.

Using the theory of competing risks

As the desired interpretation of the relative survival rate is the observed survival rate in the fictitious situation in which only the disease of the patients is capable of killing them, the net probability of survival in the theory of competing risks [15] provides a possible alternative. It differs from the alternatives previously considered, in so far that an expected survival rate is unnecessary in calculations. This is an advantage in situations in which no appropriate life tables for a comparable general population exist. On the other hand, knowledge of the causes of death of the patients is needed, or at least knowledge of whether or not a given death is attributable to the disease under study. When the age-specific net probabilities of survival are weighted in proportion to the numbers of persons in each age group at the beginning of follow-up in order to obtain a net probability for the whole group, there arise practical difficulties similar to those discussed in the section on weighting age-specific relative survival rates. For old patients in particular it may be difficult to estimate, or even to think of long-term survival probabilities in a situation in which a specific disease of the patients is the only disease killing patients. The attractive feature of the net survival probability is that it is non-increasing when the time of follow-up is lengthened.

The application of Chiang's model for competing risks [15] to the data in the example resulted in strictly non-rising curves for the net probability of survival

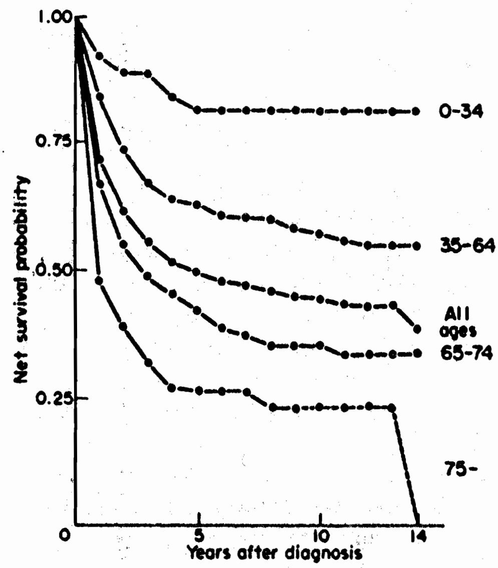

FIG. 4. Net probabilities of survival by age at diagnosis for female patients with localized cancer of the colon or small intestine in Finland 1953–1970, on the assumption that cancer of the colon and that of the small intestine were the only possible causes of death. The figures for all ages combined have been calculated as weighted averages of the age-specific net-probabilities, with the age distribution of patients at diagnosis being used as constant weights. A broken line indicates estimates based upon <5 patients.

(WNP in Fig. 3, Fig. 4). These are somewhat lower than the corresponding relative survival rates (Fig. 2). The data used merits attention in regard to causes of death. In the above calculations, use was made of the causes of death, checked and if necessary corrected by the Finnish Cancer Registry [8]. If no corrections were introduced, e.g. the 10-yr net probability of survival would increase from 0.447 to 0.497, and the 14-yr net probability from 0.384 to 0.424 respectively. An increase takes place, as most of the corrections made by the Registry are changes from cancer other than that of the primary site, to cancer of the primary site [8]. One additional difficulty in long-term survival studies arises from deaths from new tumours of the same primary site. Calculation of the net survival probabilities implies that a decision is required whether these are recurrences of the original tumour, or new tumours.

CONCLUSION

The relative survival rate over a given period has been defined as the ratio of the proportion of survivors observed in the patient group, to the proportion of survivors expected in a group similar to the group of patients at the beginning of the period with respect to all possible factors that affect survival, except for the disease being studied. When a relative survival curve is drawn, a possible rise in the curve when follow-up time is lengthened solely indicates a rise in the ratio between the two aforementioned proportions. It also indicates a decrease in the relative excess risk of death during the follow-up time with the prolongation of the follow-up. The comparison between consecutive points of the relative survival curve is, however, confounded by the age-specific expected proportions of survivors. In addition to age, a similar confounding effect will be produced by any factor which is taken into account in the determination of expected survival rates.

For heterogeneous patient populations, a possible rise in the relative survival curve, with increase in the follow-up time does not necessarily indicate that the patients are affected by a lower mortality than that expected. Consequently, relative survival curves should be avoided, and annual relative survival rates presented, if mortality during a specific part of follow-up is under study.

The relative survival curve must not be interpreted as the survival curve for patients in the case that only the disease of patients can be a cause of death, since it can, in principle, be rising. If truly non-rising curves for this interpretation are required, the theory of competing risks should be applied. If proper information on the causes of death of the patients is not available, relative survival curves should be examined by age, and by other background variables, to make the groups homogeneous with respect to the expected survival.

REFERENCES

1. Berkson J, Gage RP: Calculation of survival rates for cancer. **Proc Staff Meet Mayo Clin** 25: 270–286, 1950
2. Axtell LM, Cutler SJ, Myers MH: End results in cancer. Report No. 4. 217 pp. Bethesda: U.S. Department of Health, Education and Welfare. National Cancer Institute, 1972
3. The Cancer Registry of Norway: **Survival of Cancer Patients. Cases Diagnosed in Norway 1953–1967.** Oslo: The Norwegian Cancer Society, 1975
4. Elandt-Johnson RC: Definition of rates: some remarks on their use and misuse. **Am J Epid** 102: 267–271, 1975
5. Ederer F, Axtell LM, Cutler SJ: The relative survival rate: A statistical methodology. **Nat Can Inst Mono** 6: 101–121, 1961
6. Waterhouse JAH: **Cancer Handbook of Epidemiology and Prognosis.** Edinburgh: Churchill Livingstone, 1974
7. Cutler SJ, Axtell LM: Partitioning of a patient population with respect to different mortality risks. **J Am Stat Ass** 58: 701–712, 1963
8. Hakulinen T, Teppo L: Causes of death among female patients with cancer of the breast and intestines. **Ann Clin Res** 9: 15–24, 1977
9. Teppo L, Hakama M, Hakulinen T *et al.*: Cancer in Finland: Incidence, mortality, prevalence. **Acta Path Microbiol Scand Sect A: Suppl** 252: 1–79, 1975
10. Cutler SJ, Ederer F: Maximum utilization of the life table method in analyzing survival. **J Chron Dis** 8: 699–712, 1958
11. Official Statistics of Finland VI A: 114, 121, 126, 134. Life Tables 1951–1970. Helsinki: Central Statistical Office of Finland 1957, 1963, 1968, 1974
12. Merrell M, Shulman LE: Determination of prognosis in chronic disease, illustrated by systemic lupus erythematosus. **J Chron Dis** 1: 12–32, 1955
13. Heise H: The effect of a temporary change in survival on the life table method for computing survival rates. Methodological Note No. 8. End Results Evaluation Section. National Cancer Institute, February 4, 1959
14. Heise H: Computing expected survival rates: all ages combined. Methodological Note No. 12. End Results Evaluation Section. National Cancer Institute, January, 1960
15. Chiang CL: **Introduction to Stochastic Processes in Biostatistics.** New York: Wiley, 1968

APPENDIX

FORMAL DEMONSTRATION OF THE EFFECT UPON THE RELATIVE SURVIVAL CURVE OF A HETEROGENEOUS POPULATION WITH RESPECT TO AGE

Let

$_ip_0$ = the observed survival rate in the group of patients from the beginning of follow-up to the end of the ith year of follow-up,

$_ip_0^*$ = the corresponding expected survival rate,

and

$_ir_0 = {_ip_0}/{_ip_0^*}$ = the relative survival rate.

Let the patients be divided into m age groups, with quantities corresponding to those for the whole group $_ip_0(a)$, $_ip_0^*(a)$ and $_ir_0(a)$, $a = 1, \ldots, m$.

For the whole group of patients,

$$_ip_0 = \sum_{a=1}^{m} w_a \cdot {_ip_0(a)}, \tag{1}$$

in which

w_a = the proportion of patients belonging to age group a at the beginning of follow-up.

As

$$_ip_0^* = \sum_{a=1}^{m} w_a \cdot {_ip_0^*(a)},$$

it follows that

$$\begin{aligned} _ir_0 &= \left[\sum_{a=1}^{m} w_a \cdot {_ip_0(a)}\right] \Big/ \left[\sum_{a=1}^{m} w_a \cdot {_ip_0^*(a)}\right] \\ &= \left[\sum_{a=1}^{m} w_a \cdot {_ip_0^*(a)} \cdot {_ir_0(a)}\right] \Big/ \left[\sum_{a=1}^{m} w_a \cdot {_ip_0^*(a)}\right] \\ &= \left[\sum_{a=1}^{m} W_a(i) \cdot {_ir_0(a)}\right] \Big/ \left[\sum_{a=1}^{m} W_a(i)\right], \end{aligned} \tag{2}$$

in which

$$W_a(i) = w_a \cdot {_ip_0^*(a)}. \tag{3}$$

It is thus apparent that the relative survival rate for the whole group is a weighted average of the relative survival rates in various age groups. The weights are proportional both to the sizes of the age groups at the beginning of follow-up, and to the expected survival rates in the age groups, i.e. to the expected numbers of persons alive in the age groups at i yr after the beginning of follow-up.

Let us then, for the sake of convenience, assume that no secular changes occur in the annual expected survival rates. This is essential only for simplicity of the formulae. Let us in addition assume that the patients in each age group are homogeneous with respect to the expected mortality. In principle, this is not a restriction, since it is always possible to choose sufficiently narrow age groups to make the assumption valid. Thus

$$_ip_0^*(a) = \prod_{j=0}^{i-1} p_j^*(a)$$

in which

$p_j^*(a)$ = the annual expected survival rate for a person in age group a j yr after the beginning of follow-up.

If the index of the youngest age group is denoted by b, then for $a \neq b$ it follows that

$$_ip_0^*(a)/{_ip_0^*(b)} = \left[\prod_{j=0}^{i-1} p_j^*(a)\right] \Big/ \left[\prod_{j=0}^{i-1} p_j^*(b)\right].$$

For sufficiently large i, the above ratio is equal to

$$\left[\prod_{j=0}^{i-\Delta_{ba}-1} p_j^*(a) \cdot \prod_{j=i-\Delta_{ba}}^{i-1} p_j^*(a)\right] \Big/ \left[\prod_{j=0}^{\Delta_{ba}-1} p_j^*(b) \cdot \prod_{j=\Delta_{ba}}^{i-1} p_j^*(b)\right],$$

in which

Δ_{ba} = the age difference between the groups a and b.

By reason of the absence of secular changes in the annual expected survival rates,

$$p_j^*(a) = p_{\Delta_{ba}+j}^*(b),\ 0 \leq j \leq i - \Delta_{ba} - 1,$$

and thus

$$ {}_ip_0^*(a)/{}_ip_0^*(b) = \left[\prod_{j=i-\Delta_{ba}}^{i-1} p_j^*(a)\right] \Big/ \left[\prod_{j=0}^{\Delta_{ba}-1} p_j^*(b)\right]. \tag{4}$$

The numerator of the right hand side of eqn (4) converges towards zero, whereas the denominator is a constant when i increases. Hence, the ratio ${}_ip_0^*(a)/{}_ip_0^*(b)$ converges towards zero when i increases.

It follows from this that the ratio $W_a(i)/W_b(i)$, $a \neq b$, converges towards zero when i increases, and that ${}_ir_0$ converges towards ${}_ir_0(b)$ with an increase in i. The relative survival rate for the whole group thus gradually reaches the relative survival rate of the youngest age group.

APPENDIX TABLE. LIFE TABLES BY AGE AT DIAGNOSIS FOR FEMALE PATIENTS WITH LOCALIZED CANCER OF THE COLON OR SMALL INTESTINE IN FINLAND, 1953–1970†

i	l_i	d_i	w_i	${}_{i+1}p_0$	${}_{i+1}p_0^*$	${}_{i+1}r_0$
0–34						
0	70	5	7	0.9248	0.9992	0.9255
1	58	2	10	0.8899	0.9984	0.8913
2	46	—	10	0.8899	0.9975	0.8921
3	36	2	1	0.8398	0.9967	0.8426
4	33	1	—	0.8143	0.9958	0.8178
5	32	—	2	0.8143	0.9949	0.8185
6	30	—	—	0.8143	0.9939	0.8193
7	30	—	3	0.8143	0.9928	0.8202
8	27	—	4	0.8143	0.9917	0.8211
9	23	—	2	0.8143	0.9905	0.8221
10	21	—	3	0.8143	0.9892	0.8232
11	18	—	1	0.8143	0.9878	0.8244
12	17	—	7	0.8143	0.9862	0.8257
13	10	—	3	0.8143	0.9846	0.8271
14	7					

35–64						
0	518	83	30	0.8350	0.9917	0.8420
1	405	51	33	0.7254	0.9827	0.7381
2	321	33	16	0.6489	0.9729	0.6670
3	272	13	29	0.6161	0.9623	0.6403
4	230	5	13	0.6024	0.9507	0.6336
5	212	8	22	0.5784	0.9380	0.6166
6	182	2	27	0.5715	0.9243	0.6184
7	153	2	17	0.5636	0.9092	0.6199
8	134	4	14	0.5459	0.8927	0.6114
9	116	3	16	0.5307	0.8749	0.6066
10	97	3	17	0.5127	0.8554	0.5994
11	77	4	13	0.4836	0.8343	0.5797
12	60	2	9	0.4662	0.8117	0.5744
13	49	1	8	0.4558	0.7873	0.5790
14	40					
65–74						
0	408	142	22	0.6423	0.9629	0.6670
1	244	46	21	0.5158	0.9231	0.5587
2	177	25	16	0.4395	0.8807	0.4990
3	136	16	20	0.3837	0.8357	0.4591
4	100	10	15	0.3422	0.7886	0.4339
5	75	9	3	0.3003	0.7391	0.4063
6	63	5	8	0.2748	0.6876	0.3997
7	50	4	9	0.2507	0.6344	0.3951
8	37	2	4	0.2364	0.5805	0.4072
9	31	3	3	0.2123	0.5262	0.4035
10	25	3	7	0.1827	0.4724	0.3867
11	15	1	3	0.1692	0.4193	0.4034
12	11	1	2	0.1522	0.3673	0.4145
13	8	1	1	0.1319	0.3176	0.4154
14	6					
75–						
0	251	134	10	0.4553	0.8822	0.5161
1	107	23	15	0.3500	0.7713	0.4538
2	69	12	8	0.2854	0.6674	0.4276
3	49	11	3	0.2193	0.5712	0.3840
4	35	2	8	0.2052	0.4830	0.4248
5	25	3	3	0.1790	0.4031	0.4440
6	19	7	2	0.1094	0.3319	0.3295
7	10	3	—	0.0766	0.2698	0.2838
8	7	3	—	0.0438	0.2157	0.2029
9	4	—	—	0.0438	0.1694	0.2582
10	4	—	2	0.0438	0.1304	0.3355
11	2	—	1	0.0438	0.0984	0.4446
12	1	—	—	0.0438	0.0732	0.5974
13	1	1	—	—	0.0535	—
14	—					

† For symbols see footnote of Table 1.