A Short Textbook of

Medical Statistics

Sir Austin Bradford Hill

C.B.E., D.Sc., Ph.D., Hon. D.Sc. (Oxon.), Hon. M.D. (Edin.), F.F.C.M. (Hon.), F.R.C.P. (Hon.), F.R.S.

Professor Emeritus of Medical Statistics in the University of London; Past Honorary Director of the Statistical Research Unit of the Medical Research Council; Past President of the Royal Statistical Society; Civil Consultant in Medical Statistics to the Royal Navy and the Royal Air Force

HODDER AND STOUGHTON

LONDON SYDNEY AUCKLAND TORONTO

In association with The Lancet

1977

18 Life Tables and Survival after Treatment

In the assessment of the degree of success attending a particular treatment given to patients over a series of years the life-table method is sometimes an effective procedure. Before illustrating its application to such data, consideration of the national life table and its use in public health work will be of value. A life table, it must be realised, is only a particular way of expressing the death-rates experienced by some particular population during a chosen period of time. For instance, a recent life table constructed by the Registrar General of England and Wales is based upon the mortality experience of men in the three years 1972–74 (unpublished figures kindly provided by the Registrar General). It contains six columns as shown in the table below.

A LIFE TABLE BASED UPON THE DEATH-RATES OF MEN IN ENGLAND AND WALES IN 1972–74

Age x	l_x	d_x	p_x	q_x	$\mathring{e}_x$
0	100 000	1870	·98130	·01870	69·25
1	98 130	115	·99883	·00117	69·57
2	98 015	72	·99926	·00074	68·65
3	97 943	62	·99937	·00063	67·70
4	97 882	—	—	—	—
—	—	—	—	—	—
—	—	—	—	—	—
—	—	—	—	—	—
90	4 089	852	·79161	·20839	3·65
—	—	—	—	—	—
—	—	—	—	—	—
—	—	—	—	—	—

The essence of the table is this: suppose we observed 100 000 infants all born on the same day and dying as they passed through each year of life at the same rate as was experienced at each of these ages by the population of England and Wales in 1972–74, in what gradation would that population disappear? How many would be still left alive at age 25, at age 56, etc.? How many would die between age 20 and age 30? What would be the chance of an individual surviving from age 40 to age 45? What would be the average length of life enjoyed by the 100 000 infants? Such information can be obtained from these different columns. The basis of the table is the value known as q_x, which is the probability, or chance, of dying between age x and age $x + 1$, where x can have any value between 0 and the longest observed duration of life. For instance, q_{25} is the chance that a person who has reached a twenty-fifth birthday will die before reaching the twenty-sixth birthday. These probabilities, one for each year of age, are calculated from the mortality-rates experienced by the population in 1972–74. This probability of dying is the ratio of those who fail to survive a particular year of life to those who started that year of life; (to take an analogy, if 20 horses start in a steeplechase and 5 fail to survive the first round of the course the probability of 'dying' on that round is 5/20; 15 horses are left to start on the second round and if 3 fail to survive, the probability of 'dying' on the second round is 3/15).

The Probability of Dying

As pointed out above, the basic element of the life table is the probability of dying between one age and the next. Once those values are known throughout life the construction of the remainder of the table is a simple, though arithmetically laborious, process. To calculate these probabilities requires a knowledge, for the population concerned, of the numbers living and dying in each single year of life. Let us suppose that such detailed data are available and, to make it specific, that in the City of A there were 1500 persons enumerated (or estimated) at the middle of the year 1976 whose age was 22 years last birthday, i.e. their age was between 22 and 23. During the calendar year there were, say, 6 deaths between ages 22 and 23. Then the *death-rate* at ages 22–23, as customarily calculated, is the ratio of the deaths observed to the mid-year population, i.e. 6 in 1500, which equals 4 per 1000, or 0·004 per person. In symbols $m_x = D \div P$, which gives the death-rate per person. This mid-year population does not indicate precisely how many persons *started* the year of life 22 to 23, as it is an enumeration of those who were *still alive* at the middle of the calendar year. On the average they were at that point of time $22\frac{1}{2}$ years old, since some would have just passed their 22nd birthday, some would be just on the point of having their 23rd birthday, and all intermediate values would be represented. If we may reasonably presume that the deaths occurring between ages 22 and 23 are evenly spread over that year of life, then we may conclude that half of them would have occurred before the mid-year enumeration (or estimation) and half would have occurred after it. In other words, the population that *started* out from age 22 is the 1500 survivors plus a half of the recorded deaths, or 1500 + 3, this half of the deaths being presumed to have taken place before the mid-year point. The *probability of dying* is, by definition, the ratio of the deaths observed in a year of life to the number who set out on that year of life, i.e. 6 in 1503. In symbols, therefore, $q_x = D \div (P + \frac{1}{2}D)$.

If, then, we know the mid-year population at each age and the deaths taking place between each age and the next, the probabilities of dying can be readily calculated from this formula. (It is, however, not very accurate in the first 2 or 3 years of life and particularly in the first year. In the first year in countries with a low infant mortality rate the deaths occur more frequently in the first 6 months of life than in the second 6 months and a more appropriate fraction would be $D \div (P + \frac{3}{4}D)$.)

It is clear that there must be a simple relationship between m_x, the death-rate, and q_x, the probability of dying. It may be demonstrated as follows:—

$m_x = D/P$, so that $Pm_x = D$.
$q_x = D/(P + \frac{1}{2}D)$. Substituting Pm_x for D gives $q_x = Pm_x/(P + \frac{1}{2}Pm_x)$

But P occurs in both numerator and denominator, so it may be removed to give $q_x = m_x/(1 + \frac{1}{2}m_x)$; and, finally multiplying top and bottom by 2 to get rid of the half, gives $q_x = 2m_x/(2 + m_x)$.

In other words, the probability of dying may be calculated from the formula (twice the death-rate) ÷ (the death-rate plus 2), where the death-rate is calculated not as usual per 1000 persons but as per person; or from the formula (deaths) ÷ (population plus half the deaths).

The Construction of the Life Table

Having thus calculated, by one or other formula, these values of q_x for each year of life, the life table is started with an arbitrary number at age 0, e.g. 1000, 100 000, or 1 000 000. By relating the probability of a new-born infant dying before its first birthday (q_0) to this starting number, we find the number who will die in the first year of life. By subtracting these deaths from the starters we have the number of survivors that there will be at age 1. But for these survivors at age 1 we similarly know the probability of dying between age 1 and age 2; by relating this probability to the survivors we can calculate how many deaths there will be between age 1 and age 2. By simple subtraction of these deaths we must reach the survivors at age 2. And so on throughout the table till all are dead. Thus the figures on p. 199 show that for males the probability of dying in the first year of life is 0·01870, or in other words, according to the infant mortality-rate of 1972–74, 1·870 per cent of our 100 000 infants will die before they reach their first birthday. The actual number of deaths between age 0 and age 1 will therefore be 1·870 per cent of 100 000, or 1870. Those who *survive* to age 1 must be 100 000 less 1870 = 98 130. According to this table, the probability of dying between age 1 and age 2 is 0·00117, or in other words 0·117 per cent of these 98 130 children aged 1 year old will die before reaching their second birthday. The actual number of deaths between age 1 and age 2 will therefore be 0·117 per cent of 98 130 = 115; those who survive to age 2 must therefore be 98 130 less 115 = 98 015. From these q_x values the l_x and d_x columns can thus be easily constructed, l_x showing the number of individuals out of the original 100 000 who are still alive at each age, and d_x giving the number of deaths that take place between each two adjacent ages. p_x is the probability of living from one age to the next. $p_x + q_x$ must equal 1, since the individuals must either live or die in that year of life. To return to our analogy, if 5 out of 20 horses do not complete the round, clearly 15 out of 20 do survive the round. p_x, therefore, equals $1 - q_x$; for example, of the 98 015 children aged 2, 0·074 per cent die before reaching age 3, and it follows that 99·926 per cent must live to be 3 years old.

Finally, we have the column headed $\mathring{e}_x$ which is the 'expectation of life' at each age. This value is not, in a sense, an 'expectation' at all, for it is the *average* duration of life lived beyond each age. For example, if we added up all the ages at death of the 100 000 male infants and took the average of these durations of life we should reach the figure 69·25 years. If, alternatively, we took the 98 130 infants who had lived to be 1 year old, calculated the *further* duration of life that they enjoyed, and then found the average of those durations, we should reach the figure 69·57 years. At age 90 there are 4089 persons still surviving, and the average duration of life that they will enjoy after that age is only 3·65 years. The so-called expectation of life is thus only the average length of life experienced after each age. We thus have the complete life table.

Calculation of the Expectation of Life

The expectation of life, or average length of life given by a life table, can be calculated from the column of deaths (d_x) or from the column of survivors (l_x). Using the former, we have an ordinary frequency distribution. Thus of the males in this 1972–74 table there were 1870 who died between age 0 and 1. We may presume they lived on the average half a year (some exaggeration, in fact). Their contribution to the total years lived by the original 100 000 is therefore $(1870 \times \frac{1}{2})$. Between ages 1 and 2 there were 115 deaths, and we may presume for these that the average length of life was $1\frac{1}{2}$ years. Their contribution to the total years lived by the original 100 000 is therefore $(115 \times 1\frac{1}{2})$. Between ages 2 and 3 there were 72 deaths, and we may presume for these that the average length of life was $2\frac{1}{2}$ years. Their contribution to the total years lived by the original 100 000 is therefore $(72 \times 2\frac{1}{2})$. And so on until at the far end of the table when there are no more survivors. Proceeding to the final point, we then have the total years of life lived by the whole population of 100 000:—

$$(1870 \times \tfrac{1}{2}) + (115 \times 1\tfrac{1}{2}) + (72 \times 2\tfrac{1}{2}) + \ldots + (852 \times 90\tfrac{1}{2}) + \ldots$$

The expectation of life is the mean number of years lived, and so this sum is divided by the 100 000 persons starting at age 0 to whom it relates. In short, $\mathring{e}_x$ at birth equals the sum of all values of $d_x \times (x + \frac{1}{2})$ divided by 100 000, and in this table for males is, as previously stated, 69·25 years.

Using the survivor column (l_x), we may proceed as follows. The 98 130 survivors at age 1 have all lived a whole year of life, between age 0 and age 1. They give a contribution therefore of 98 130 whole years of life to the total years lived. But the survivors at age 2, who were 98 015 in number, have all lived another whole year of life – from age 1 to age 2. They therefore give a further 98 015 whole years of life to the total years lived. Similarly the survivors at age 3, who are 97 943, are contributing that extra number of whole years of life lived. Thus by summing the l_x column from age 1 to the final entry we have the number of whole years of life lived by the 100 000 (clearly the 100 000 must not be included in the sum, for they are the starters at 0 and at that point have lived no duration).

There is, however, a small error involved if we stop at that sum. In it we have made no allowance for the period each person lives in the year of his, or her death. We have considered only the survivors at ages 1, 2, 3, etc. But we may presume that the 1870 who died between 0 and 1 had half a year's life before their death, that the 115 who died between 1 and 2 also had half a year's life in the year 1–2 in which they died, that the 72 who died between 2 and 3 also had half a year's life in the year 2–3 in which they died (it will be noted that their whole years of life up to age 2 have already been counted in the survivorship column and it is only the half-year in the year in which they died that has been omitted). We must therefore add to the sum of whole years lived, derived from the l_x column, these half-years lived by those dying in the precise year in which they died. In other words, we have to add in $(1870 \times \frac{1}{2}) + (115 \times \frac{1}{2}) + (72 \times \frac{1}{2}) + \ldots$ to the end of the table. But this implies, as everyone is dead by the end of the table, adding in $(100\,000 \times \frac{1}{2})$. The final sum of years required is, therefore, given by the sum of the l_x column from 1 to the end of the table plus 100 000 times a half. The average, or expectation of life, is then this sum divided by the 100 000 at the start, which may be expressed as

$$\frac{\text{Sum of } l_x \text{ column (excluding } l_0)}{100\,000} + \frac{1}{2}$$

As stated previously, the expectation of life at a later age than 0, say age 25, is the average length of life lived beyond that age by the survivors at age 25. It can be calculated by either of the above methods (the use of the l_x column being the simpler); the sum of years lived relating only to the entries beyond age 25, and the denominator, to give the average, being, of course, the survivors at age 25.

Practical Aspects of the Life Table

In using the life table as a method of comparison of the mortality experience between place and place, or epoch and epoch, various values may be chosen. For example, we may take:—

(*a*) The probability of dying between any two selected ages (the ratio of the total deaths between the two ages to the number alive at the first age).

(*b*) The number of survivors at any given age out of those starting at age 0 (the l_x column).

(*c*) The probability of surviving from one age to another (the ratio of the survivors at a given age to the survivors at a previous age).

(*d*) The expectation of life (which suffers from the limitation inherent in any average that it takes no cognisance of the variability around it).

In practice, however, it is not often possible to construct a life table by the methods described above since they require, it was seen, a knowledge of the population and deaths in single years of life. More often than not the numbers available relate to 5- or 10-yearly age-groups, and some device has then to be adopted for estimating from these grouped data the appropriate numbers at single years of life. Alternatively there

are available for public-health work excellent short methods of making a life table from the actual death-rates observed in different age-groups. To describe these methods fully is outside the scope of this chapter, the object of which is not to show how best to construct a life table in practice, but to clarify the general principles underlying its construction so that the values given by it may be understood. Taking the example given above, it shows, to reiterate, how a male population would die out if it experienced as it passed through life the same death-rates as were prevailing in England and Wales in 1972–74. It does not follow, therefore, that of 100 000 male children born in those years in reality only 97 943 would be alive at age 3; if the death-rate were, in fact, declining below its 1972–74 level, then more than that number would survive; if it were rising, less than that number would survive. As the Government Actuary points out in relation to English Life Table No. 12, 'the l_x columns could only be interpreted as showing the survivors of 100 000 children born in the period 1960–62 if the improbable assumption were made that the 1960–62 rates of mortality will remain unaltered throughout their life-times, that is until at least the year 2070. The same applies to the expectations of life; if, in line with past experience, rates of mortality decline in future, babies born in 1961 have an expectation of life greater than $\mathring{e}_x$ as shown by English Life Table No. 12.' In other words the life table can show only what would happen under *current conditions of mortality*, but it puts those current conditions in a useful form for various comparative purposes and for estimating such things as life insurance risks (inherent in the questions that were propounded above).

A Cohort Life Table

A *cohort* life table, on the other hand, can sometimes be constructed to show the actual dying out of a defined group of persons, or cohort, all born at about the same time. Thus all persons born in the year 1920 were subject to the infant mortality rate of 1920–21 and then, as the years passed, to the already recorded childhood, adolescent and adult mortality of those later years. And so in 1977 we can summarise in life-table form the true mortality of the cohort at different ages for the first 57 years of its life. We cannot, of course, go beyond that point until further years have elapsed. The cohort born in, say, 1870 could be traced to extinction.

The Measurement of Survival-rates after Treatment

We turn now to the application of the life table method to groups of patients treated over a period of calendar years whose subsequent after-history is known. Let us suppose that treatment was started in 1971, that patients were treated in each subsequent calendar year and were followed up to the end of 1976 on each yearly anniversary after their treatment had been started (none being lost sight of). Of those treated in 1971 we shall know how many died during the first year after treatment, how many died during the second year after treatment, and so on to the fifth year after treatment. Of those treated in 1972 we shall know the subsequent history up to only the fourth year after treatment, in 1973 up to only the third year after treatment, and so on. Our tabulated results will be, let us suppose, as in Table 36.

TABLE 36

RESULTS OF TREATMENT (HYPOTHETICAL FIGURES)

Year of Treat- ment	Number of Patients Treated	Number Alive on Anniversary of Treatment in				
		1972	1973	1974	1975	1976
1971	62	58	51	46	45	42
1972	39	—	36	33	31	28
1973	47	—	—	45	41	38
1974	58	—	—	—	53	48
1975	42	—	—	—	—	40

Separate calculation of the survival-rates of patients treated in each calendar year becomes somewhat laborious if the number of years is extensive and has also to be based upon rather small numbers. If the constitution of the samples treated yearly and their fatality-rates are not changing with the passage of time there is no reason why the data should not be amalgamated in life-table form. Indeed the great advantage of the life-table method is that we can utilise *all* the information to hand at some moment of time. In computing, say, a 5-year survival-rate we make *all* the patients contribute to the picture and do not restrict ourselves to only those who have been observed for the full five years. For clarity we can write Table 36 in the form given in Table 37.

All the patients have been observed for at least one year and their number is 62 + 39 + 47 + 58 + 42 = 248. Of these there were alive at the end of that first year of observation 58 + 36 + 45

TABLE 37

RESULTS OF TREATMENT (HYPOTHETICAL FIGURES)

Year of Treat- ment	Number of Patients Treated	Number Alive on Each Anniversary (none lost sight of)				
		1st	2nd	3rd	4th	5th
1971	62	58	51	46	45	42
1972	39	36	33	31	28	—
1973	47	45	41	38	—	—
1974	58	53	48	—	—	—
1975	42	40	—	—	—	—

+ 53 + 40 = 232. The probability of surviving the first year after treatment is, therefore, 232/248 = 0·94, or, in other words, 94 per cent of these patients survived the first year after treatment. Of the 40 patients who were treated during 1975 and were still surviving a year later, no further history is yet known. (If the year's history happens to be known for some of them these data cannot be used, for the history would tend to be complete more often for the dead than for the living, and thus give a bias to the results.) As the exposed to risk of dying during the second year we therefore have the 232 survivors at the end of the first year minus these 40 of whom we know no more – viz. 192. Of these there were alive at the end of the second year of observation 51 + 33 + 41 + 48 = 173. The probability of surviving throughout the second year is therefore 173/192 = 0·90. Of the 48 patients who were treated in 1974 and were still surviving in 1976 no later history is yet known. As the exposed to risk of dying in the third year we therefore have the 173 survivors at the end of the second year minus these 48 of whom we know no more – viz. 125. Of these there were alive at the end of the third year of observation 46 + 31 + 38 = 115. The probability of surviving the third year is therefore 115/125 = 0·92. We know no further history of the 38 patients first treated in 1973 and still surviving on the third anniversary. The number exposed to risk in the fourth year becomes 46 + 31 = 77, and of these 45 + 28 = 73 are alive at the end of it. The probability of surviving the fourth year is therefore 73/77 = 0·95. Finally, during the fifth year we know the history only of those patients who were treated in 1971 and still survived at the end of the fourth year – namely, 45 persons. Of these 42 were alive on the fifth anniversary, so that the probability of surviving the fifth year is 42/45, or 0·93.

Construction of the Life, or Survivorship, Table. Anniversary Data

Tabulating these probabilities of surviving each successive year, we have the values denoted by p_x in column (2) of Table 38. The probability of not surviving in each year after treatment, q_x, is immediately obtained by subtracting p_x from 1. The number of patients with which we start the

TABLE 38

RESULTS OF TREATMENT IN LIFE-TABLE FORM

Year after Treatment	Probability of Surviving Each Year	Probability of Dying in Each Year	Number Alive on each Anniversary out of 1000 Patients	Number Dying during Each Year
x (1)	p_x (2)	q_x (3)	l_x (4)	d_x (5)
0	·94	·06	1000	60
1	·90	·10	940	94
2	·92	·08	846	68
3	·95	·05	778	39
4	·93	·07	739	52
5	—	—	687	—

l_x column is immaterial, but 100 or 1000, or some such number is convenient. Starting with 1000, our observed fatality-rate shows that 94 per cent would survive the first year and 6 per cent would die during that year. The number alive, l_x, at the end of the first year must therefore be 940 and the number of deaths, d_x, during that year must be 60. For these 940 alive on the first anniversary the probability of living another year is 0·90, or in other words there will be 90 per cent alive at the end of the second year, i.e. 846, while 10 per cent will die during the second year, i.e. 94. Subsequent entries are derived in the same way. (The order of the columns in the table is immaterial. The order given in Table 38 is the most logical while the table is being constructed, because p_x is the value first calculated and the others are built up from it. In the final form the order given in the 1972–74 life table on page 199 is more usual.)

By these means we have combined all the material we possess for calculating the fatality in each year of observation after treatment, and have found that according to those fatality-rates approximately 69 per

cent of treated patients would be alive at the end of 5 years. Having found from the available material the probability of surviving each of the separate years 1 to 5 we are, in effect, finding the probability of surviving the whole 5 years by multiplying together these probabilities, viz. $p_1 \times p_2 \times p_3 \times p_4 \times p_5$.

If we want the average duration of life so far lived by the patients, it is easily obtained. 687 patients of our imaginary 1000 live the whole 5 years. If we presume that those who died lived half a year in the year in which they died (some will have lived less, some more, and we can take, usually without serious error, the average as a half), then 60 lived only half a year after treatment, 94 lived a year and a half, 68 lived two years and a half, 39 lived three years and a half, and 52 lived four years and a half. The average length of subsequent life is, therefore, so far as the experience extends, $(687 \times 5 + 60 \times 0{\cdot}5 + 94 \times 1{\cdot}5 + 68 \times 2{\cdot}5 + 39 \times 3{\cdot}5 + 52 \times 4{\cdot}5) \div 1000 = 4{\cdot}15$ years.

The percentage alive at different points of time makes a useful form of comparison. For instance, studying the treatment of that era we find for patients with cancer of the cervix treated between 1925 and 1934, the following number of survivors out of 100 in each stage of disease:—

	Stage 1	2	3	4
At end of 5 years .	86	70	33	11
At end of 9 years .	78	57	27	0

Exclusion of Patients

If some of the patients have been lost sight of, or have in a few instances died from causes which we do not wish to include in the calculation (accidents, for example), these must be taken out of the exposed to risk at the appropriate time –e.g. an individual lost sight of in the fourth year is included in the exposed to risk for the first three years but cannot be included for the complete fourth year. If he is taken out of the observations from the very beginning, the fatality in the first three years is rather overstated, for we have ignored an individual who was exposed to risk in those years and did not, in fact, die. If patients are being lost sight of at different times during the year or dying of excluded causes during the year, it is usual to count each of them as a half in the exposed to risk for that year. In other words, they were, on the average, exposed to risk of dying of the treated disease for half a year in that particular year of observation.

Construction of the Life, or Survivorship, Table. Data at a Specified Date

Sometimes in putting data into the life-table form we have patients observed not on anniversaries, as above, but to a specified date. Let us suppose, for instance, that patients have been treated in the years 1971 to 1975 and have been followed up to December 31st, 1976 (and thus not, as in the previous example, to the yearly anniversaries of their treatment). The data may be put in life-table form according to the technique set out in Table 39.

The total number of patients treated in the 5 years and to be followed up was 194. During the first year of the follow-up 4 were lost sight of and 2 died of violence and these deaths it is proposed to exclude as irrelevant. Since the last patient was treated in 1975 and the follow-up was to December 31st, 1976, all the patients had been observed for at least one full year. The exposure to risk of dying during the first year will, therefore, be computed as 194 less half the number lost sight of and less half the number of deaths from violence, i.e. 191. In other words, we give only half a year's exposure to those who passed out of observation for these reasons during the year. During these 191 person-years of exposure there were 24 deaths, giving a probability of dying of $24/191 = 0{\cdot}126$. The probability of surviving the first year is, therefore, 0·874.

The number of patients entering the second year of follow-up is the original 194 less all those who have died or who have been lost sight of during the first year, i.e. $194 - 24 - 2 - 4 = 164$. Of these 164 exposed during the second year there were 3 lost sight of during the year who must be allowed only a half-year's exposure, and there were also 35 *who were still alive at December 31st, 1976*, but who had not been exposed to risk over the *whole* of that second year. These are the patients who were treated in 1975 and who, therefore, by the end of 1976, have been exposed for one year and some fraction of a year. As usual we shall take the fraction to be, on the average, one-half. The number exposed to risk of dying during the full second year is, accordingly, $164 - \frac{1}{2}$ of 3 lost sight of $- \frac{1}{2}$ of 35 alive at December 31st, 1976, and not exposed for the whole of the second year $= 145$. During that second year the number of patients dying was 12 so that the probability of dying was 0·083 and the probability of surviving 0·917.

The number entering the third year of exposure is 164 minus the 3 lost sight of, the 12 who died and the 35 who passed out of observation alive at December 31st, 1976, which equals 114. Of these 114 patients, 1 died of violence during the third year and 42 were seen for only part of that year (i.e. those who were treated in 1974 and by December 31st,

TABLE 39

CONSTRUCTION OF LIFE TABLE OF PATIENTS UNDERGOING A CERTAIN TREATMENT (HYPOTHETICAL FIGURES)

Year after Treatment (1)	Number Alive at Beginning of the Year (2)	Number lost sight of During the Year (3)	Number Dying of Violence During the Year (4)	Number Alive Observed for only Part of the Year (5)	Number Exposed to Risk of Dying During the Year (6)*	Number Dying During the Year (7)	Proportion Dying During the Year q_x (8)	Proportion Surviving the Year p_x (9)	Proportion Surviving from Start of Treatment to End of each Year (10)†
0–	194	4	2	0	191	24	0·126	0·874	0·874
1–	164	3	0	35	145	12	0·083	0·917	0·801
2–	114	0	1	42	92·5	6	0·065	0·935	0·749
3–	65	1	0	23	53	3	0·057	0·943	0·706
4–	38	2	1	21	26	2	0·077	0·923	0·652
etc.									

* Col. 6 = Col. 2 minus half Cols. 3, 4, and 5.

† Col. 10 = the products of the values of Col. 9, i.e. $p_1 \times p_2 \times p_3$ . . . .

1976, had been observed for 2 full years and some fraction of a year). The number exposed to risk of dying is, accordingly, 114 less $\frac{1}{2}$ of 1 and $\frac{1}{2}$ of $42 = 92{\cdot}5$. With 6 deaths during the year the probability of dying is 0·065 and the probability of surviving is 0·935. And so on.

Taking the probabilities of dying and applying them to the customary hypothetical 1000 patients at start of treatment we can calculate the number alive at the end of each year as in Table 38. Alternatively if we require only the percentage of the total who will be surviving at the end of each year (i.e. the l_x column), the answer can be obtained by multiplying together the p_x values in column 9 of Table 39. Thus the probability of surviving one year is, according to these data, 0·874; the probability of surviving two years is $0{\cdot}874 \times 0{\cdot}917 = 0{\cdot}801$; of surviving three years $0{\cdot}874 \times 0{\cdot}917 \times 0{\cdot}935 = 0{\cdot}749$; of surviving four years $0{\cdot}874 \times 0{\cdot}917 \times 0{\cdot}935 \times 0{\cdot}943 = 0{\cdot}706$; and of surviving five years $0{\cdot}874 \times 0{\cdot}917 \times 0{\cdot}935 \times 0{\cdot}943 \times 0{\cdot}923 = 0{\cdot}652$. In other words, these data give a 5-year survival-rate of 65 per cent.

As a general rule the exclusion of deaths regarded as irrelevant is undesirable, e.g. from violence in a follow-up of patients operated upon for some form of cancer. If the number of such deaths is few their inclusion (or omission) can have little effect upon the results. If the number is large it may be difficult to interpret the results when they are omitted. It is probably better to compute the survival-rate with such deaths included and to compare this rate with the figure normally to be expected amongst persons at those ages.

Patients Lost to Sight

Finally, if a relatively large number of patients is lost sight of we may be making a serious error in calculating the fatality-rates from the remainder, since those lost sight of may be more, or less, likely to be dead than those who continue under observation.

For instance, if 1000 patients were observed, 300 are dead at the end of 5 years, 690 are alive, and 10 have been lost sight of, this lack of knowledge cannot appreciably affect the survival-rate. At the best, presuming the 10 are all alive, 70 per cent survive; at the worst, presuming the 10 are all dead, 69 per cent survive. But if 300 are dead, 550 are alive, and 150 have been lost sight of, the corresponding upper and lower limits are 70 per cent surviving and 55 per cent surviving, an appreciable difference. To measure the survival-rate on those patients whose history is known, or, what comes to the same thing, to divide the 150 into alive and dead according to the proportions of alive and dead in the 850 followed up, is certainly dangerous. The characteristic 'lost sight of' *may* be correlated with the characteristics 'alive or dead'; a patient who cannot be traced may be more likely to be dead than a patient who can be traced (or *vice versa*), in which case the ratio of alive to dead in the untraced cases cannot be the same as the ratio in the traced cases. Calculation of the possible upper rate shows at least the margin of error.

The Life Table in General

The use of the life table has been illustrated above in relation to death and survival, which is, indeed, its customary use. However, the method can be more generally applied to any defined feature in the follow-up of persons or patients, i.e. so long as such a feature occurs at one point in the course of time and can at that time be clearly defined as present or not present.

For instance, we might be concerned with a treatment newly introduced for patients suffering from multiple sclerosis. The question at issue is how long will the patient remain free from all, or certain defined, symptoms (in comparison, say, with a similar group of patients not so treated). The decisive end-point in the follow-up, then, is not death but the appearance of symptoms as previously defined and which, of course, need to be clear-cut.

In the same way one might follow up patients after an operation for a form of cancer and note the time at which recurrence occurred. In these examples one would have, in place of death, rates of appearance or recurrence in given intervals of time, and thus the probability of their occurring within so many months or years.

Summary

Life tables are convenient methods of comparing the mortality-rates experienced at different times and places. The same methods may be usefully applied to the statistics of patients treated and followed up over a number of years.