CONFIDENTIAL

Estimation of loss of life in relation to a disease or to a factor causing it; with particular reference to smoking

Author : P.N. Lee

Date : 23.2.78

1. Introduction

Following the publication of the latest Royal College of Physicians Report on Smoking and Health (R.C.P., 1977), considerable attention was given in the press, both in the United Kingdom and abroad, to the claim contained in it that "on average the time by which a habitual cigarette smoker's life is shortened is about 5½ minutes for each cigarette smoked - which is not much less than the time he spends smoking it." This claim was first made by Diehl (1969), who based his calculations on tables provided by Hammond (1969) giving the loss of life expectancy of U.S. men of various ages smoking different numbers of cigarettes.

Such a claim refers to only one of many different ways in which the loss of life due to smoking can be quantified. The aim of this paper is to look generally at alternative methods for estimating loss of life in relation to a disease or to a factor causing it, and to apply the ones thought most useful to obtain estimates relevant to the population of England and Wales of the loss of life due to smoking and to diseases associated with it.

This paper starts, in Section 2, by looking at the theory behind estimation of loss of life in the experimental situation. The concept of the life table is introduced and the advantages and disadvantages of a number of alternative statistics describing differences in survival between exposed and non-exposed groups are discussed. In practice, human data on the relationship between smoking and mortality are collected observationally rather than experimentally. The problems involved in collecting relevant data are discussed in Section 3 along with the assumptions required in extrapolating results obtained to the current smoker in England and Wales. Following discussion of what data actually are available (Section 4), calculations of the loss of life due to smoking and some smoking-associated diseases are made in Section 5. The conclusions of the paper are discussed in Section 6 and summarized in Section 7.

2. Estimation of loss of life in the experimental situation

2.1 The two group experiment

Ideally, to determine the loss of life related to a particular factor, one would like to take a population and randomly allocate it into two groups. One group would then be exposed to the factor of interest while the other group would not. The mortality of each group would then be monitored for the rest of its life by noting the time at which each member died. Statistics describing the difference between the mortality experience of the two groups would then be computed and could be taken to be relevant to the effect the factor would have on the loss of life of other people typical of the original population.

In the real world, such an experimental approach is only usually possible with animals so that for humans inferences have to be made about the effects of particular factors from other types of data. The problems involved and assumptions required to make such inferences are discussed later (Section 3); for the moment our interest is centred on what are, and are not, useful statistics to describe the effect of a factor on mortality and for this it is convenient to stay with our ideal situation. To further simplify discussion of method we assume, firstly, that exposure to the factor of interest is at a regular rate throughout lifetime, and, secondly, that all the members of the original population are all of the same age at the beginning of the experiment.

2.2 Functions describing survival

Survival data measure time to death. The distribution of survival times can be characterized by three equivalent functions (Gross and Clark, 1975):

a) Death Density Function f(t)

$f(t)dt$ is the probability that a person will die in the time interval $(t, t \bullet dt)$. If we assume that the experiment starts at time zero it follows that

$$\int_0^\infty f(t)dt = 1$$

$f(t)$ is non-negative.

b) Survivorship Function S(t)

S(t) is the probability that a person will survive to at least time t (t>0). It follows that

$$S(t) = \int_t^{\infty} f(T)dT$$

and that

$$f(t) = -S'(t)$$

S(t) is a non-negative decreasing function starting from 1 at time zero.

c) Hazard Function λ(t)

λ(t)dt is the probability that a person will die in the time interval (t, t + dt) given he has survived to time t. This function, which is also known as the failure rate, the force of mortality or the incidence rate, satisfies the condition

$$\lambda(t) = f(t)/S(t)$$

λ(t) is non-negative but may be increasing (such as in the Weibull distribution $\lambda(t) = bt^k$ where b and k are constants), constant (such as in the exponential distribution λ(t) = a) or have other more erratic shapes.

2.3 Cohort life-table

For absolute precision one observes, as mentioned above, the actual time at which deaths occur. In practice, especially for human populations, it is usually convenient to group the data into certain defined time intervals rather than actual points of time. We shall assume, for our purposes that the data we have for each group consists of information relevant to n time intervals (i = 1, ... n) as follows:

t_i Age of population at beginning of interval i

A_i Number alive at beginning of interval i

D_i Number dying from all causes in interval i

L_i Number dying from a particular cause of interest in interval i

Y_i "Midpoint" of interval i.

Such information is known as a life-table. According to the nomenclature of Gross and Clark (1975), the particular type of life-table we are dealing with here is a cohort life-table, a "cohort" being a group of individuals born at about the same time. Later on (Section 3.5), we consider other forms of life table.

We note that, because we have assumed all people are followed until death, $A_i - D_i = A_{i+1}$ for all i and that $t_1 = 0$ and $A_n = D_n$. Y_i, the "midpoint" of interval i, can usually, if the interval is sufficiently small, be taken as the actual midpoint of the age-interval considered. If more accurate answers are required the actual average age at which deaths in the interval occur should be substituted. $P_i = A_i/A_1$ estimates $S(t_i)$ the survivorship function at age t_i.

Description of the mortality of a population by life-tables has a long history dating back to the pioneer work of Halley (1693 - sic). Mortality indicators in general have been reviewed in the literature on a number of occasions (e.g. Woolsey (1943), Haenszel (1950), Logan and Benjamin (1953), Kitagawa (1966), Benjamin and Haycocks (1970), Romeder and McWhinnie (1977)). In this, and the sections that follow, we discuss the merits of a number of statistics that have been suggested to summarize the main features both of the information contained in a life-table and of the differences between the two life-tables being compared. We start by looking at some of the more simple statistics that have been employed in the past.

2.4 Measures of the proportion dying

One obvious type of statistic to look at is the proportion dying. Clearly if one is looking at total mortality then the proportion dying over the whole experiment will be 100% and will offer no discrimination between the groups. However it can be useful to compare the proportion dying between two particular time points t_j and t_k, especially if the time period represents, in some sense, "premature" deaths. This statistic, Q_1, is defined by

$$Q_1 = 1 - \frac{S(t_k)}{S(t_j)} = \frac{A_j - A_k}{A_j}$$

If one is interested in mortality from a particular cause then the total proportion dying from this cause

$$Q_2 = \left(\sum_{i=1}^{n} L_i\right) / A_1$$

can give some indication of the magnitude of the problem caused by the disease. It is, however, limited in its usefulness by the fact that it gives no information as to when the deaths occur.

There are two other types of indicator which have the same objection that they concentrate on numbers of deaths and ignore when deaths occur. The first of these are standardised death rates. They can be calculated by two methods, the direct and the indirect method. In the direct method, the rate, Q_3, is a weighted sum of the individual crude death rates (R_i), with the weights representing the populations in each age-group in some standard population. Thus, if w_i are the weights, Q_3 is defined by

$$Q_3 = \sum_{i=1}^{n} w_i R_i$$

In the indirect method, the number of deaths from the cause of interest observed (O_i) in an interval is compared with that expected (E_i) if some standard death rates (R_i^*) from the cause had existed. The sum of deaths observed from all intervals is divided by the total expected to give a 'Standardized Mortality Ratio', Q_4. Q_4 is defined by

$$Q_4 = \frac{\sum_{i=1}^{n} O_i}{\sum_{i=1}^{n} E_i} = \frac{\sum_{i=1}^{n} L_i}{\sum_{i=1}^{n} A_i R_i^*}$$

A problem with both these indicators, as was pointed out by Yerushalmy (1951), is that they are markedly affected by relatively small differences in mortality in older ages when deaths are frequent and little affected by large proportional differences in early years which cause great loss of life.

The final type of statistic, quite popular in quantifying the effect of smoking (e.g. R.C.P. (1971)) is the number of deaths associated with the factor. To calculate this statistic, Q_5, the number of deaths that actually occur in an interval in the exposed group are compared with the number that would have occurred had the exposed group had the same number at risk in the interval but the death rates of the non-exposed group. In other words

$$Q_5 = \sum_{i=1}^{n} \left(D_{2i} - A_{2i} \cdot \frac{D_{1i}}{A_{1i}} \right)$$

where the first subscript refers to the group (1 = non-exposed, 2 = exposed) and the second to the time interval. Apart from the fact that Q_5 gives no indication at all of life-shortening, the main defect with this statistic is that it carries with it the implication that, had the factor not existed, this number of deaths would, in some sense, have been avoided. Clearly everyone dies once, so what is the real implication? As normally used, the number of deaths associated with a factor is attached to a time scale, e.g. "50,000 deaths a year are associated with smoking" but what does this mean? As calculated, if the time intervals were years, and if in fact this calculation was carried out on population data rather than our idealized cohort data, the statistic would be a reasonable estimate of the numbers of deaths that would not have occurred in the year following a universal giving up of smoking, assuming (and there is evidence to show that for some diseases, e.g. lung cancer (Doll (1971)), this is not the case) that on giving up smokers age-specific death rates reverted at once to those of never smokers. However, it would only be accurate for the first year and would be increasingly inaccurate for subsequent years. The reason being, of course, that in later years, due to the lower mortality immediately following mass giving-up (on the assumption quoted), there would be more survivors at higher ages and consequently more deaths than the current age-distribution would suggest. As shown in Appendix A, it can be estimated (under certain further assumptions) that, on mass giving up of smoking at the end of 1975, 72,000 less male deaths in England and Wales would have occurred the first year afterwards than had no giving-up occurred. However this number would be half as much by 1988 and down to 12,000 by the year 2,000.

2.5 Life expectation and average age at death

Another simple statistic that has been used to assess mortality is average age at death. Average age at death of the whole population from all causes is identical to expectation of life at birth; expectation of life at age t, Q_6, being defined by the expression

$$Q_6 = \int_t^{\infty} (T - t)\ f(T)dT/S(t)$$

and measures the average number of years of life still to be lived by those who have survived to age t.

Average age at death of those people dying only of the cause of interest, Q_7, is an alternative statistic which has been used by some workers. It is defined by

$$Q_7 = \sum_{i=1}^{n} (L_i Y_i) / \sum_{i=1}^{n} L_i$$

Though it can be of value in some circumstances to compare such an average for one cause of death with a similarly computed average for another cause, it only measures when the disease occurs and not how many people die of it. Furthermore it is not a very useful statistic to measure life shortening. It might be thought that, a cause of death resulting in an average age at death of x years less than the average age of death from all causes is in some sense an indication that the cause takes x years off life. That this reasoning is incorrect can be seen if one considers a cause of death such as stroke with an average age greater than the expectation of life at birth. On the implied line of reasoning this cause adds years onto life, which is, of course, nonsense.

Average age at death can also be a very misleading statistic to use when comparing groups exposed to different levels of a factor of interest. If, for example, the cause of death of interest is the only one affected by a factor and is relatively rare, and if the effect of the factor is simply to multiply the age-specific incidence rate from the cause by an age-independent constant, it can be easily seen that, though the proportion of cases of the cause of death in the group more exposed to the factor will be greater than in the group less exposed, the distribution of times of death from the cause, and hence the average age at death from the cause, will be virtually identical in the two groups. If, furthermore, the average ages at death from the cause are compared in cross-sectional data, where the age distribution of the more and less exposed groups are different, it is not surprising that fairly meaningless results can be obtained. For example, Passey (1962) studied successive hospital lung cancer patients and observed that the average age of death of the heavy smokers did not differ from that of the light smokers. He concluded that there was an anomaly to be explained, but as Pike and Doll (1965) pointed out, following out general line of argument above, if was only the poor choice of statistic that had led to the apparent anomaly.

2.6 Measures of loss of life expectation

The preceding sections demonstrate that any statistic not taking into account both the frequency and the time of occurrence of death is not an adequate description of the effect a factor has on loss of life. A better approach, and one that has been tried by various workers over the last 30 years, is to quantify the effect in terms of numbers of years lost. Some of these attempts have tried to take into account to at least some extent the belief that the loss of years of life at young ages may be of more importance to an individual, or to a society, than the loss of a similar number of years in old age.

Thus, a number of workers, e.g. Murray and Axtell (1974), Romeder and McWhinnie (1977), have estimated the number of years of "active life" lost. Though the critical age differs (usually between 60 and 70), the same essential method of calculation has been used; it has been assumed that any person dying before the critical age has lost a number of active years equal to the difference between the critical age and the year of death.

Other workers have used deaths occurring at all ages and have counted years lost to life expectancy, e.g. Dempsey (1947) who used life expectancy at birth and Dickinson and Welker (1948) who used life expectancy at the age of death.

Both these types of measure have objections. Measures of "active life" lost, as described above, are over-estimates as it is clear that some of those dying early would still not have reached the critical age had they not died when they did. Hakulinen and Teppo (1976) tried to get round this objection by using adjusted life-table procedures (see Section 2.10) to estimate the subsequent survival pattern of the "reincarnated" population, i.e. the survival of those who would not have died had the cause of death of interest been removed. An alternative method would be to compare the years of "active life" lost in the exposed and non-exposed groups. However these measures all have the disadvantage that an essentially arbitrary choice of critical age has to be made and that information on people of age greater than this is ignored.

The measures of loss of life expectation mentioned above also have the disadvantage that they do not take account of the fact that, had the cause of death been removed, the life expectancy itself would have been altered. Furthermore, the resulting statistic can be rather difficult to interpret especially if it is calculated on a per decedents rather than a per head of population at risk of basis. Thus, as we shall show later, the average years lost to life expectancy of lung cancer decedents for some populations is in fact less than the average years lost for all decedents. Is lung cancer a good thing therefore? It can also be shown that (considering cross-sectional data rather than life-table type data), even had smoking no effect on mortality at all, the average loss of life expectancy of smokers who die in a given time period would be greater than that of non-smokers who die in the same period, simply because smokers are younger than non-smokers. It is clear that such statistics are liable to misuse by the uninitiated.

2.7 Recommended methods for comparison of two life-tables

Probably the most informative method of quantifying the loss of life due to a factor is to compute the difference in life expectation of the exposed and non-exposed groups from the start of the experiment. If it is desired to place a different value, V(t), on life at different ages then one could calculate the difference between the two groups in their "expected value of life", Q_8, where Q_8 is defined by

$$Q_8 = \int_0^{\infty} T\ f(T)\ Z(T)dT$$

where Z(T) the total value of life up to time T is given by

$$Z(T) = \int_0^{T} V(U)dU$$

Another good method is to compare the proportions dying over some special age range of interest (Q_1). This method has, for example, been used by the R.C.P. (1971) to quantify the effect of smoking. They pointed out that, in the study of British Doctors (see Section 4.2) a male smoker of 25 cigarettes or more daily aged 35 had a 40% chance of dying by age 65 whereas, over the same period the chance of dying for a non-smoker was only 18%.

In particular circumstances, other comparisons of life-tables can be extremely informative. For example, if the effect of the factor is simply to transform the death density function so that either

$$f_2(t) = f_1(t + a)$$

or $$f_2(t) = f_1(bt)$$

holds for all t (where a and b are constants and the subscripts 2 and 1 refer to exposed and non-exposed respectively) then one could make statements such as "the probability of an exposed person dying is the same as that of a non-exposed person a years older" or "exposed people have a probability of dying the same as non-exposed people b times as old". Alternatively, if the factor multiplies the hazard function by a constant c, so that

$$\lambda_2(t) = c\lambda_1(t)$$

one might usefully make statements such as "if you are exposed you have c times the chance of dying at any instant as a non-exposed person of the same age." There is quite a lot of evidence that, for particular diseases the effect of some factors can be a multiplication of the hazard function (e.g. Peto and Lee (1973)) but such a simple relationship seems unlikely to hold for total death rates for more than at most a very few factors.

2.8 Quantifying loss of life expectation in terms of dose applied

None of the statistics which we have discussed so far to quantify the difference between life-tables representing exposed and not exposed groups have expressed the differences in terms of the dose applied to the exposed group. We now look at some that do.

If all we wish to say is something along the lines "exposure to X units a day for life results in a loss of life expectation of Y years" then, of course, the methods of the last section are directly applicable.

If we wish to generalize this statement so that we can make inferences about what would happen if lifetime exposure was at some other daily level we would need to have information on additional groups exposed at different dose levels in order to build up a dose-response relationship, but no new statistical treatment of this life-table would be needed.

It is not so straightforward, however, if from the results of an exposed group given X units a day for life we wish to infer the effect on life shortening of a single exposure to X. Such an inference has to be made to arrive at Diehl's (1969) claim "on average the time by which a habitual cigarette smoker's life if shortened is about 5½ minutes for each cigarette smoked."

The simple method to do this calculation, and the one used by Diehl (1969) is to compute the average total exposure during the life of the exposed group and to divide this into the estimated loss of life expectation. But is this correct? Normally, in working out such an average, one computes the life shortening per exposure for each individual and then averages the answers over the individuals; in other words if LS_j is the life shortening for individual j and E_j his exposure up to time of death one would calculate

$$Q_9 = \frac{1}{N} \sum_{j=1}^{N} (LS_j/E_j)$$

where N is the number of individuals, and not as Diehl did,

$$Q_{10} = \left(\sum_{j=1}^{N} LS_j \right) / \left(\sum_{j=1}^{N} E_j \right)$$

As is well known, unless LS_j/E_j is constant for every individual, these expressions will differ. The reason why Q_9 has not been calculated, of course, is that it is not possible, on an individual basis to measure life shortening.

Another worrying thing about the whole concept is the fact that one might be averaging effects which are very different at different times of life. In fact, as we show below, the assumption that each exposure has an equal effect, implies a particular relationship between the death density functions of the exposed and non-exposed groups.

The death density function, $f(t)$, in the non-exposed group can be seen as the proportion of people who have "tickets to die" at time t. (Or more precisely $f(t)dt$ represents the proportion with tickets to die in $(t, t + dt)$). Suppose that every unit of time that elapses the exposed group take R off their original life expectation; in other words they take R off the value of their ticket. It follows that, at time T, the $f(T + RT)$ people who originally had tickets to die at time $T + RT$ then hold tickets to die at time T. It follows that the survivors at time T are only those who originally had tickets to die at least at time $T + RT$. The survivorship function in the exposed group at time T, $S_2(T)$, is therefore given by

$$S_2(T) = \int_{T(1+R)}^{\infty} f(u)du = S_1(T(1 + R))$$

In other words, at any time the surviving proportion in the exposed group is the same as the surviving proportion in the non-exposed group at a time a factor of R + 1 as great. It follows that, unless such a condition holds at least approximately for some R, expression of the effect of a single exposure as an average tends to be rather uninformative.

2.9 The one group experiment

On some occasions information is available on the survival of a group exposed to a factor or subject to a cause of death, but no such data are available on a comparable non-exposed control group. Can inferences then be made about the effect the factor or cause of death has had on loss on life from this "one group experiment"? Provided the cause of death of each member of the group is recorded, then under certain assumptions which we shall discuss in the next section it is possible to estimate the loss of life caused by a certain cause of death. Inferences cannot in general be made about the effect of a factor unless, from other knowledge, it is possible to label certain deaths as having been caused by the factor.

2.10 Adjusted life-tables

When information is only available on a single group the recommended procedure is to calculate the life-table that would have existed had the particular deaths (from the cause of interest or classed as due to the factor) not occurred. This life-table is known as the "adjusted life-table" and inferences about loss of life can then be made by comparison of the actual and adjusted life-tables in exactly the same way as have been described for the comparison of life-tables for the exposed and non-exposed groups in the two group experiment.

How is the adjusted life-table calculated? The most common approach to this problem is to assume that those people who die of the cause of death of interest would have had the same chance of dying from the other causes of death had the cause of interest not existed. Under this <u>assumption of independence</u> the adjusted life-table is estimated as follows.

Consider the ith time interval. Let the length of this interval be T_i and let us assume it to be small enough for the force of mortality from causes other than the one of interest to be taken as constant (= α_i), and that from the cause of interest also to be taken as constant (= β_i).

Now the total survival from all causes

$\exp(-(\alpha_i + \beta_i) T_i)$ is estimated by $(A_i - D_i)/A_i$

and the relative mortality

$$\alpha_i/\beta_i \quad \text{by} \quad (D_i - L_i)/L_i$$

(L_i are the deaths from the cause(s) to be adjusted for)

It follows that

$$\exp(-\alpha_i T_i) = \left(1 - \frac{D_i}{A_i}\right)^{\frac{D_i - L_i}{D_i}}$$

and

$$\exp(-\beta_i T_i) = \left(1 - \frac{D_i}{A_i}\right)^{\frac{L_i}{D_i}}$$

$\exp(-\alpha_i T_i)$ is the proportion that would survive the interval if the cause of death of interest had not existed. The survivorship function of the adjusted life-table is thus built up by starting with 100% survivors (i.e. $S(0) = 1$) at the beginning of the first interval and successively multiplying S by the estimate of $\exp(-\alpha_i T_i)$ in each consecutive interval.

The formula for $\exp(-\alpha_i T_i)$ has been derived previously by Chiang (1961). An alternative formula

$$\exp(-\alpha_i T_i) = 1 - \frac{D_i - L_i}{A_i - L_i/2}$$

has been attributed to Berkson by Schwartz and Lazar (1961). As Lee (1970) shows these two formulae give virtually identical answers provided the proportions dying in the interval are reasonably small.

Although most work done in this area makes the assumption we made at the beginning of the previous section, it is clear that it is only likely to hold under special circumstances. It may well be, for example, that the sort of person who dies early from one cause may be generally "weak" and, were this cause to be removed, he would in fact have a greater chance than average of dying from other causes subsequently. In a recent paper, Wong (1977), considered this problem. His approach was, instead of assuming the relative susceptibility of those dying in an interval to be the same as those surviving, to assume that it was different by a general proportional factor F. In Appendix B, we discuss this alternative not assuming independence in more detail. There we show that there are some objections to the way in which Wong carried out his calculations and derive a somewhat different scheme along Wong's basic idea.

The method described in Appendix B could be generalized to the situation where there are more than 2 groups with differing susceptibility. Indeed the population could be given some defined continuous susceptibility distribution initially. Such generalizations are not investigated in this paper, the simpler situation dealt with in Appendix B being sufficient to allow illustration of the sort of effect that variations in susceptibility of the population have on estimates of loss of life expectation due to elimination of a particular cause of death.

2.11 Estimation of loss of life in a multifactorial situation

All the estimates of loss of life described in this section deal with quantifying the effect a single factor has on loss of life. In practice many diseases are multifactorial in origin. It follows that any conclusions made about the factor studied only apply to a population with the same levels of other relevant factors as in the groups studied. For example the true loss of life related to smoking may be much higher in a group of asbestos-exposed workers than that estimated from a study of the normal population. By studying more than 2 groups it would be possible in a single experiment to obtain inferences on the loss of life related to more than 1 factor.

For example, by carrying out a 4 group experiment with groups exposed to both, one only or neither of 2 factors, inferences can be made about the loss of life related to each factor. In these circumstances it can be convenient, if it is possible to do so, to choose a statistic to measure loss of life which allows independent expression of the effect of each factor. For example, if factor A causes 3 years loss of life expectation, factor B 4 years and the two together 7 years it would be convenient to use loss of life expectation as a measure as the conclusions about the two factors individually hold whether or not the other factor is present. Of course, if in this example, there was no loss of life in the group exposed to both factors it would be misleading to say simply that factor A causes 3 years loss of life expectation and factor B 4 years as this only holds if the other factor is not present.

Although we do not intend to pursue the multifactorial situation in detail in this paper, we note, finally that clear thinking is needed in the interpretation of the findings. Rose (1977), carrying out an analysis on data relating coronary heart disease mortality to a number of factors previously described by Reid et al (1976), found that the "number of deaths associated" with each factor, when added together, exceeded the total number of deaths occurring. Todd (1977), considering these results, felt that there was something wrong with the "number of deaths associated" approach. So there may be, as we showed in section 2.4, but it is not the choice of statistic that caused the apparent problem. The problem lay in adding together a number of results, each of which represented the effect of removing a particular factor in the presence of all the others. If the individual deaths associated with each factor are to add to the total, one should add together

the number of deaths associated with factor A in the presence of B, C .... N +

the number associated with B in the presence of C ... N (and not A) +

the number associated with C in the presence of D ... N (and not A or B)

+ etc. etc.

Of course, if it came to attribution of deaths, or rather claims related to them in a legal case, then one should be aware of the fact that the order in which the associations with the factors are calculated (i.e. the order in which the factors are successively eliminated) may affect the answers. For example, if a disease only occurs if two conditions are met, then whichever factor is considered first in the analysis will appear to be wholly responsible. But this is a problem for lawyers and not statisticians, whose job is to present the facts in an unbiassed and meaningful way.....

3. Estimation of loss of life in the human smoking situation - problems involved and assumptions required

3.1 Randomization and causality

In the real life situation smokers and non-smokers are not randomly allocated from a single group, but themselves choose whether or not they smoke. For this reason excess death rates from various diseases observed in smokers may, in theory, to at least some extent, represent differences due to the smoker rather than to smoking. In applying the methods described in the previous section to smoking we choose to ignore this problem, however, and make the simplifying assumption that any differences in death rates observed between smokers and non-smokers of the same age and sex are due to smoking itself.

While medical authorities (R.C.P., 1977) generally believes that this assumption is approximately true for some causes of death, there are particular cases of death where at least some opinion holds that this assumption is false. For example, the excess death rate of cirrhosis of the liver amongst smokers is generally held to be attributable to the fact that cirrhosis of the liver is caused by drinking and that virtually all heavy drinkers smoke (Doll and Hill (1964)). Furthermore, though this view is generally disavowed by the medical authorities (R.C.P., 1977), some workers (Fisher (1959), Burch 1976)) believe that the excess lung cancer death rate of smokers is not caused by their smoking at all, but by a common genetic tendency to smoke and to get lung cancer. To determine the true proportion of the excess deaths among smokers that are actually caused by smoking is beyond the scope of this paper. Indeed, the R.C.P. (1971) when considering this problem, said it was not possible to give a precise estimate, saying only "there can be little doubt that at least half of the estimated 31,000 excess deaths among male smokers aged 35-64 in the United Kingdom (in 1968) were due to smoking". Rather, we simply note that there is no practical difficulty in adjusting the estimates we make, if different assumptions are used.

3.2 Problems of cohort data

Even assuming that cohort data on self-selected groups of smokers and non-smokers can be used in the same way as data on randomly selected there still remain a number of reasons why the interpretation of results from such data is not straightforward.

Firstly, for information over the whole of its lifetime to be available, the cohort must have been born before about 1875 and much of the information on cause of death will of necessity refer to times when diagnostic practices may have differed substantially from those used today. It is generally believed that, for lung cancer, standards of diagnosis have changed dramatically since the beginning of the century, accounting for a considerable proportion of the observed rise in incidence (R.C.P., 1977).

Even if death rates from the cause of interest could be adjusted in some suitable way for changes in diagnosis (and attempts have been made to do this for lung cancer (ibid)), there would still remain a second practical objection. This is that, if there have been marked trends in time with respect to the level and age-distribution of the cause of death, statistics summarizing the effect a cause has had on a past cohort may have little relevance to people in cohorts still alive today. Although, for particular purposes it may be useful to quantify what effect a particular cause of death used to have, our main interest is really in trying to produce statistics that quantify the effect it has on present-day man.

A third problem in studying the effect of smoking is that cigarettes themselves have changed over the years. In particular there has been a general switch from smoking plain to smoking filter cigarettes. Since 1955 when 98% of cigarettes smoked in the U.K. were plain, the percentage has dropped to 13% in 1975 (Lee, 1976). It seems likely that the health risks associated with filter cigarettes are significantly less than those associated with plain cigarettes (Hammond <u>et al</u>, 1976; Dean <u>et al</u>, 1977; Bross and Gibson, 1968; Wynder <u>et al</u>, 1970) and consequently inferences about the effect of smoking on loss of life based on old data may markedly overestimate the true effect on current populations of smokers.

A fourth difficulty with cohort data lies in the fact that in practice an appreciable number of smokers modify the amount they smoke during their lives or even give up. At first sight, provided regular information on the smoking habits of the cohort is recorded, this does not seem much of a problem as one can exclude from analysis any people whose habits deviate from those that one originally intended to study. However, if "modifiers" or "givers-up" are not representative of "continuers" bias can occur in estimating loss of life. In particular, if some diseases related to smoking prompt people to cut down or give up, and if such diseases are related to the subsequent probability of death (a situation where "smoking causes disease" interacts with "disease causes smoking") great care should be taken in the interpretation of the findings. This is discussed again later in this paper.

Finally, as we have mentioned previously, one should take care before assuming that estimates of loss of life obtained from cohort studies of special populations are necessarily relevant to other populations. The relevance to current smokers in England and Wales of some of the data that is available on smoking and mortality is discussed in Section 4.

3.3 Population and prospective life-tables

Theoretically, it is possible to construct a more up-to-date life-table by using available data from a more recent cohort and, by extrapolation from the trend of age-specific mortality rates in previous cohorts, filling in estimates of death rates in ages not yet reached by the cohort. In view of the unreliability of extrapolation for more than a short time ahead, such an approach is unlikely to be very useful. It seems better to look for another method more referable to current experience and not involving extrapolation or adjustment for changes in diagnosis.

The alternative approach, which, though not without objections, is perhaps the best available to illustrate the current effect of a cause of death, is to construct a life-table using up-to-date cross-sectional data. Data are said to be collected cross-sectionally if, for each age-group, it is collected over the same small period of time. For our purposes we would need to collect, for each age-group of interest, estimates of the population alive and of the number of deaths occurring, both in total and due to the cause of interest. Such data does not, of course, refer to one cohort (it can be seen as a snapshot of the current status of a number of consecutive cohorts) and does not form a life-table directly (the population alive in one age group may, for example, be higher than that in a younger age group due to variation in the birth rate over time). However, a life-table can be constructed from it by starting with some arbitrary number alive (A_1, usually taken as 100,000) and reducing it, for each successive age group, by multiplying it by $(1 - d_i)^{t_i}$ where d_i is the cross-sectional death rate per year from all causes in age group i and t_i is the length of the interval in years.

In the terminology used by Gross and Clark (1975), the life-table so constructed is a population life-table. It is the life-table of a hypothetical cohort having at each age of its life the mortality rates of the whole population observed cross-sectionally at that time. This should be borne in mind in considering statistics summarising this experience. Thus, the fact that the expectation of life calculated for such a cohort may be 72 years does not imply that a child born at the time the rates were observed will, on average, live 72 years. To obtain an estimate of how long children born then will live on average requires more than just knowledge of the age-specific death rates then. It also requires assumptions to be made as to how the rates are going to change in the future. And this one cannot reliably predict for more than a short distance ahead.

In the case of some human data, as we shall see in the next section, a population (consisting of a number of different birth cohorts) may be followed for a number of years in a "prospective" study. It can be useful, on occasion, to combine all the survival experience from the study into a single life-table. To do this, the average death rate for a particular age interval is calculated by dividing the total number of people in the study who died in the age interval by the total number of "man-years at risk". Man-years at risk is the total length of time spent by the study population in the age interval during the study. Given these death rate estimates a life-table can be constructed in the same way as for the population life-table. For clarity, we christen this a prospective life-table.

An objection to a prospective life-table is that, in forming age-specific death rate estimates, one is combining together survival experience from a number of years. Such an averaging will tend to conceal any trends in survival over the period of the prospective study.

4. Data available on smoking and mortality

4.1 Introduction

Having, in the last 2 Sections, discussed the theoretical merits and demerits of some statistical techniques to measure life shortening and of applying such techniques to human populations, we now turn our attention to using the methods discussed to quantify the loss of life due to smoking and to diseases associated with it. As noted before our main interest is centred on obtaining estimates relevant (under certain assumptions) to the current population of England and Wales. Before embarking on the calculations (Section 5) it is convenient to consider what data is actually available to us on smoking and mortality.

4.2 Data from prospective studies

It was not until the early 1950's that evidence started to accumulate from retrospective studies (e.g. Doll and Hill (1952)) to suggest a strong association between smoking and lung cancer mortality. Before that time there would have been no real reason to carry out a prospective study and it is not surprising therefore that no actual study carried out has followed the mortality of cohorts of smokers and non-smokers for more than about 20 years. Thus no cohort life-tables can be calculated directly.

It would not be appropriate to consider here all the prospective studies on smoking and health that have been started in the last 25 years. Some are only on small populations, others have been running for only a relatively short period of time and others have not reported results in a form suitable for our analyses. We will restrict ourselves to mentioning three, one British and two American which have followed large populations for a long period of time and have reported results in detail.

One of the most important of the prospective studies is the British Doctor's Study started in 1951 by Doll and Hill, for which 20 year follow-up results for males have been given by Doll and Peto (1976). In this study all doctors were asked to complete a questionnaire on smoking habits in 1951, 69% responding. Subsequently participating doctors were requestionned in 1957, 1966 and 1972 and their mortality continuously monitored. Only a few, 103 out of 34,440 were lost to follow-up of the male doctors who completed the original questionnaire.

There are a number of reasons why doctors are not typical of the general population. Doctors are financially better off than average and, as members of social class 1 are known to have levels of mortality for many diseases less than the national average. And as Doll and Peto (1976) pointed out, many more doctors than would have been expected from national smoking figures gave up smoking in the last 20 years and, as the authors claim, their total mortality in consequence decreased markedly relative to national figures. However, despite these differences, most medical opinion holds that the relative mortality of smokers to non-smokers found in the British Doctor's Study from the various causes of death tabulated is a reasonable indicator of the relative mortality of smokers to non-smokers in the British population at large. It would seem reasonable, however, that even if this opinion is correct, relative mortalities from the British Doctor's Study are only likely to be applicable to the British population in about 1961 (the centre-point of the 20 year follow-up period) and not to the British population in 1978, because of the large switch to filter cigarettes that has taken place since 1961 (see Section 3.2).

The second prospective study worth considering here is Hammond's million person study for which the latest results have been reported by Hammond et al in 1976. In this study, 1,078,894 people in 25 American States were interviewed by American Cancer Society volunteers between October 1959 and February 1960 and subsequently requestionned at 2 yearly intervals until 1965 to obtain details of smoking habits. Mortality has been followed-up for 12 years. The study population is not typical of the U.S. population, being of markedly higher than average social class. However, as for the British Doctor's Study, it has been used as a vehicle for providing estimates for the population at large of the relative risk of smoking for many diseases or groups of diseases.

The third large prospective study is the Dorn Study of U.S. Veterans. In this case the population studied, 198,834 in total, was drawn from policy holders of U.S. Government Life Insurance, an insurance available to those who served in the U.S. Armed Forces between 1917 and 1940. Like Hammond's million person study there were very few non-whites studied and the higher social classes were over represented, the reason for this being that poorer members of society tend less to be able to afford insurance policies. In addition, Dorn's Study, results for which were reported in Kahn (1966), consisted almost completely of males.

4.3 National mortality data

Although information is not nationally available on mortality for smokers and non-smokers separately, many countries publish regularly information on mortality for the population as a whole. For England and Wales, for example, the Registrar General publishes annually tables giving estimates of the number living and the number dying from each main cause of death by 5 year age group.

4.4 National smoking data

Imperial Tobacco Limited have carried out annual surveys of smoking habits in the United Kingdom since 1948. The results of these surveys have been published at regular intervals by the Tobacco Research Council in their Research Paper 1, "Statistics of Smoking in the United Kingdom", the latest edition Lee (1976) giving figures up to 1975. Currently the surveys obtain information on smoking habits from about 10,000 people a year and the published results give details, inter alia, on the percentage of people who are non-smokers, ex-smokers and smokers of varying numbers of cigarettes a day by sex and age group. In recent years, estimates of smoking habits in the population have also been published in the tabulations of the General Household Survey (O.P.C.S. (1975)). Where comparable figures are presented, the degree of agreement between the surveys is, on the whole, very close.

4.5 Data used in the estimations

In the examples considered in Section 5 estimations are made of loss of life to males in three situations:

a) due to smoking in Hammond's study
b) due to lung cancer in England and Wales in 1971-75
c) due to smoking in England and Wales in 1971-75.

The data for the first situation was drawn from Hammond (1969) who gave a life-table showing the survivorship of men aged 25 in relation to current number of cigarettes smoked per day which is reproduced here as Table 1. The life-table (a population life-table - see Section 3.3) was constructed from the deaths occurring in the 5 year period July 1, 1960 to June 30, 1965 of 447,196 male subjects born between 1868 and 1927 classified into 5 year age-groups 33-37, 38-42, ... 88-92. As described in the paper the life-table has been adjusted to the 1959-61 U.S. life-table rates for all males and extrapolations have been made to cover the age range 25-34.

Table 2 gives estimates of the population, total deaths per year and deaths from lung cancer per year for males in England and Wales over the period 1971 to 1975. These estimates were calculated by averaging figures provided for individual years in the Registrar General's Statistical Review of England and Wales (Tables 1, 17 and 2). For the years 1971 to 1973, where the living population data by age group ended with an 85+ group, the distribution into 85-89 and 90+ was estimated by assuming the same ratio of population in the two age groups for the years 1974 and 1975, where the data was given.

The data required for estimation of the effect of smoking in England and Wales are more difficult to obtain. As cohort data on the mortality of smokers and non-smokers in this population is not available, it has to be estimated indirectly using data on the relative risk of mortality associated with smoking obtained from the prospective studies and on the distribution of smoking habits of the England and Wales population from national smoking data. To be useful such mortality and smoking data should be given by similarly defined age and smoking habit group.

Considering the mortality data first, we have to decide firstly which prospective study data to use. Table 3 gives estimates of the relative mortality of cigarette smokers only and never smokers by five 10 year age groups taken from Doll and Peto (1976), Hammond (1966) and Kahn (1966). It shows that, though relative risk estimates are consistently lower in the latter study, the differences between the studies are small. This suggests that using data from any one of the studies should give fairly similar answers.

In practice, information on cigarette only smokers and never smokers is not enough as we have to consider the total population in some of our analyses. For this reason we must look for a subdivision of the total population by smoking habit for which information by age group is available not only on relative risk from one of the prospective studies but also on distribution in the England and Wales population in the national smoking data. After looking at the data available, it was decided that it was convenient to divide the population into four smoking habit groups:

a) never smoked
b) current smokers of cigarettes (with or without pipes or cigars)
c) current smokers of pipes and/or cigars (but not cigarettes)
d) ex-smokers.

Data on relative mortality by five 10 year age groups for these four smoking habit groups, taken from the Dorn Study, the only one presenting data in the form required, is given in Table 4.

As can be seen Table 4 gives information only for men in the age-range 35 to 84. For our purposes some assumption will have to be made as to relative risk of mortality for men outside this range. In our calculations we have made two somewhat arbitrary assumptions. The first is that above the age of 84 the same relative risks apply as in the age group 75-84. Inaccuracies in this assumption will not have a great effect as deaths of people age 85 or more form only about 9% of total deaths (see Table 2). The second assumption is that below the age of 35 the four smoking habit groups have equal relative risk of mortality. Of total deaths in the 25-34 age range almost half are due to accidents, while only about 7% (unlike the 30% or so in the 35-44 age range) are due to ischaemic heart disease, the cause of death contributing most strongly to premature deaths associated with smoking. In view of this, and the fact that deaths in the age range 10-34 (it seems inconceivable that any earlier deaths are due to smoking) form only about 2% of total deaths, any error associated with this approximation is bound to be small compared with errors associated with other assumptions made.

The figures given in Table 4 are based on deaths occurring over a period 1954 to 1962, and represent the effect of smoking for a period when many of the population were smoking high-tar plain cigarettes. To make inferences from such data and apply it to people living today is to assume that the risks associated with smoking are the same as they were 20 years ago. We call this assumption <u>Assumption A</u>.

As noted before (Section 3.2) some evidence has recently accumulated all indicating that the health risks associated with filter cigarettes are significantly less than those associated with plain cigarettes so it is very likely that inferences under Assumption A would over-estimate the loss of life smoking modern cigarettes causes a modern smoker. Unfortunately much of the evidence on the relative effects of plain and filter cigarette smoking considers lung cancer mortality only (Bross and Gibson (1968), Wynder <u>et al</u> (1970) or only selected diseases (Dean <u>et al</u> (1977)) and the one paper that considers total mortality (Hammond (1976)) bases information on smoking habits at latest in 1966, defining "low" T/N as less than 1.2 mg of nicotine and normally less than 17.6 mg of tar, a borderline which includes many cigarettes classed today as "low to middle tar" or "middle tar" in the Government Chemist's League Tables. Hammond's estimated relative risk for total mortality for male deaths in 1966-72 of "low" T/N smokers as against "high" T/N smokers is 0.81, a ratio much higher than the relative risks Dean <u>et al</u> (1977) found when comparing retrospectively the lung cancer, coronary heart disease, chronic bronchitis and stroke death rates of people who had smoked filters in 1954, 1964 and 1969 compared with continuing plain smokers (0.39, 0.49, 0.58 and 0.53 respectively). As Dean's study is open to doubt for a number of reasons discussed in it, and as the latest death considered is only in 1972, a year when the sale of "low tar" cigarettes had hardly started in England, it is clear that there is no reliable basis for calculating an alternative assumption.

However to illustrate the effect alternative assumptions to Assumption A have on the statistics derived, and to give some feel of what might be the true number of deaths associated with smoking nowadays, we have used a somewhat arbitrary second assumption, Assumption B, in some calculations. In this assumption we have assumed that only 60% of the excess relative risk of current cigarette smokers to never smokers indicated by Kahn's figures applies today. We assume (in the absence of knowledge either way) that the relative risk of current pipe and/or cigar smokers is the same as under Assumption A. Also that ex-smokers, most of whom will have smoked wholly or a majority of plain cigarettes when they smoked, have the same relative risks as under Assumption A. The relative risks involved in Assumption B are given in Table 5.

Having obtained the mortality data required for part 3 of our example we now consider the data needed to give the distribution of the male population of England and Wales by smoking habit group and age. Lee (1976) does not give figures in precisely the form required. However, by methods described in detail in Appendix C, it was possible to derive a sufficiently good approximation to the data required for our purposes indirectly. The resultant data are given in Table 6. Figures for men aged less than 35 are omitted as they are not required since it has been assumed that smoking does not affect mortality rates below 35.

5. Calculations of loss of life due to smoking and some smoking-associated diseases

5.1 Loss of life due to smoking in Hammond's study

The analyses described in this section are based on Table 1, which, as noted in section 4.5, shows the survivorship of men aged 25 in relation to current number of cigarettes smoked per day based on the results of Hammond's study. It is effectively the same as a life-table differing only in the fact that the entries are multiplied by 100, so that it measures percentage rather than proportion surviving. It also only refers to ages 25 or over taking 100% as the percentage surviving at age 25.

To calculate life expectation, Q_{11}, for a particular group of men the formula

$$Q_{11} = \sum_{i=1}^{n} Y_i (Z_i - Z_{i+1})/100$$

is used (which is equivalent to that for Q_6 for continuous data) where Y_i is the midpoint of interval i and Z_i is the entry in the survivorship table. If, as Hammond did, one assumes that, within each 5 year period, those men who die do so at the centre of the period, the successive mid-points are 5 years apart and the formula reduces to

$$Q_{11} = 27.5 + \frac{1}{20} \sum_{i=2}^{n} Z_i$$

The expectations of life for men aged 25 for all men and the various smoking groups were calculated by Hammond (1969) and are presented in Table 1. As can be seen it is estimated that a smoker of 40 or more cigarettes a day who continues to smoke can expect to live for 8.3 years less than someone who has never smoked regularly. The loss of life expectation reduces as the number of cigarettes smoked is smaller.

Another useful way of looking at the data is to consider the proportions surviving to some age of interest. For example, taking age 65 as of interest as being the customary age of retirement it can be seen that slightly over three quarters (77.7%) of those who have never smoked regularly can expect to live to age 65 whereas only just over a half (54.0%) of 40+ a day cigarette smokers can. Putting it another way the probability of death by 65 for 40+ a day smokers (46.0%) is almost double that for never smokers (22.3%).

To calculate the proportion of those alive at one given age who survive to another one simply divides the survivors in the given column at the two relevant ages. R.C.P. (1971) presented figures, based on Doll's study, showing the proportion of men aged 35 who will survive to age 65. These are reproduced in Table 7, together with comparable figures based on Hammond's data in Table 1. As can be seen the figures in the two studies are in the main reasonably comparable, though Hammond's percentages in general are slightly higher.

We next consider loss of life expectation per cigarette smoked. To calculate this from the losses of life expectation given in Table 1 two additional pieces of data are required, the average age of starting to smoke of cigarette smokers and the average number of cigarettes smoked per day within the broad categories 1-9, 10-19, 20-39 and 40+. From Table 33M of Lee (1976) it can be seen that, in every survey year considered (1965 to 1975) the median age of starting to smoke of those who have ever smoked and know when they started lies in the age group 16-17. Though this figure refers to all smokers in the U.K. we will assume that it applies to Hammond's cigarette smokers and take the centre of the age group 17.0 years as the average. Similarly, using data in Table 22M from the same source, one can estimate that the average number of cigarettes smoked in the 1-9 a day group is about 4 and in the 10-19 group about 13. Information on the other groups is less reliable, as Lee (1976) only goes up to 30+ a day in his tables, and we have somewhat arbitrarily assumed the midpoints are 25 of the 20-39 group (it is dominated by 20 a day smokers) and 50 of the 40+ group.

Based on these assumptions (and the further one that deaths between the ages of 16 and 25 are not affected by smoking) Table 8 shows the calculations involved and the results obtained in estimating the loss of life per cigarette smoked. The figures show that, for a 40+ a day smoker, the loss of life expectation is 4.8 minutes, somewhat different from the figure of "almost 6 minutes per cigarette" given by Diehl (1969) who, presumably used slightly different assumptions of age at starting or average number smoked within the 40+ a day group. It is noteworthy that the estimates of loss of life per cigarette smoked increase continuously with decreasing amount smoked. Particularly noteworthy is the very high figure of 31.2 minutes for 1-9 a day smokers. This may partly be due to an artefact related to the fact that Hammond's life-table was not based on a study of men who actually smoked this number all their lives but on the mortality experience over a limited period of men who happened to be smoking 1-9 a day at the time. It could well be that a proportion of the 1-9 a day smokers contained men who had previously smoked more but had cut down due to health problems. Such people would presumably be at a higher risk than actual continuing 1-9 a day smokers. It should be pointed out, however, looking back at Table 7, that Doll's data does not show as large an increase in risk of death at the lowest level of smoking as Hammond's. Dolls' 1-14 a day doctors had a 39% increase in probability of dying before 65 (given survival to 35) as compared with never smokers whereas Hammond's 1-9 a day men had a 48% increase despite presumably smoking about half as much as the doctors.

We have shown that loss of life expectation per cigarette smoked depends on the number of cigarettes smoked. We can also see from Table 9 that the implicit assumption required for this measure to be useful, as discussed in section 2.8, does not hold. As we showed there, the assumption that each cigarette has an equal effect implies that the ratio of the ages at which smokers and never smokers reach any given proportion surviving is constant. As can be seen, this is not so, the ratio increasing steadily over the period for any of the smoking groups considered. In fact what is more nearly constant is the difference in ages to reach a given proportion surviving.

5.2 Loss of life due to lung cancer in England and Wales in 1971-75

Table 2, given previously, shows the population, total deaths and deaths from lung cancer by age for males in England and Wales based on 1971-75 data. To estimate loss of life due to lung cancer we first construct a population life-table (see section 3.3) so that we can see what the effect of lung cancer would be on a population having at each age of its life the average mortality rates present in England and Wales in 1971-75. This life-table is given in Table 10.

As can be seen from Table 10, lung cancer would be responsible for 8.4% of total deaths in such a population, slightly less than the proportion, 8.8%, in the actual population. From the population life-table it is possible to calculate, using adjusted life tables (see section 2.10) the expectation of life of the population if lung cancer did not exist and to compare it with that in the adjusted life-table.

The results, given in Table 11, are based on calculations made under more than one assumption. Firstly we assumed that the population was homogeneous, i.e. that those dying of lung cancer would have the same death rates from other causes as the rest of the population had lung cancer not existed. Secondly, we assumed (see Appendix B) that the population consisted of two groups (proportions P, 1-P) with the first group having F times the general level of susceptibility as the first.

As Table 11 shows, under the first assumption, lung cancer is responsible for a loss of life expectation equal to just about a year per head of the population or equal to almost 12 years in those actually dying of it. Under the second assumption, the loss is smaller, as would have been expected. However unless one postulates both a fairly large relative susceptibility factor (F) and also a susceptible proportion (P) which is neither very small nor very large, the difference is fairly marginal. Even when the factor is 10-fold and the proportions equally divided lung cancer is still responsible for just over 2/3 of a year of loss of life. Although a different sort of assumption, in which lung cancer decedents in particular are assumed to have been virtually certain to have died of some other cause very soon after their age of death had they survived, would have reduced the estimate of loss of life due to lung cancer this does not seem very plausible. It thus seems reasonable to assume that if deaths from lung cancer ceased to exist, the population at large would live the best part of a year longer.

Based on Table 12, which gives numbers of deaths from ischaemic heart disease (ICD 410-414) and from all neoplasms (ICD 140-239) for the same population as for Table 10, we calculated, for comparison with lung cancer, the effect these two major cause of death categories have on loss of life. The results are summarised in Table 13. It can be seen that, per head of population, loss of life expectation due to lung cancer is about a third of that due to all neoplasms and just over a quarter of that due to ischaemic heart disease.

In the previous paragraphs we have been computing average loss of expectation of life due to cause of death by comparing the expectations of life in unadjusted and adjusted life tables. This loss can be related to the people dying from the cause by dividing the total years lost by the total number dying from the cause. As noted in section 2.6, other workers have computed average years lost to life expectancy of people dying from a cause by averaging the remaining life expectancy of those people at the time they died. Table 14 gives, based on the population life-table of Table 10, estimates of remaining life expectancy by age and illustrates the method of calculation used. Applying the figures in column 3 to the numbers of lung cancer deaths in Table 10 one can then show that the average loss of life expectancy of lung cancer decedents is 11.91 years. Applying them to the numbers of total deaths in Table 10, the average loss, as noted previously, is lower - 11.42 years.

5.3 Loss of life due to smoking in England and Wales in 1971-75

For the purposes of the analyses in this section we wish to compare the estimated life table of men who have never smoked with that for the total 1971-75 England and Wales male population. The life table of never smokers is not available directly, but can be estimated from the information available by age group on population and total deaths (Table 2), distribution of smoking habits (Table 6) and relative total death rates by smoking habit group (Table 4 - Assumption A; Table 5 - Assumption B). The first stage in the calculation is to estimate, for each age group, death rates within each smoking habit group. The method used to do this is illustrated in Table 15 taking the 60-64 age group as an example and using assumption A. The results obtained are given in Table 16 (Assumption A) and Table 17 (Assumption B).

From the estimates of death rates of never smokers, their survivorship functions can easily be calculated. The functions under both assumptions, and that for the total population (taken from Table 10) are given in Table 18. From Table 19, which gives certain characteristics of these survivorship functions, it can be seen that, assuming cigarettes have as much relation with mortality in 1971-75 as they did some 15 years earlier in Dorn's study (Assumption A), if the whole male population of England and Wales were never smokers then their life expectation (72.3 years) would be 3.2 years greater than that of the current population (69.1 years). On the alternative assumption, that the excess risk related to smoking cigarettes in 1971-75 is only 60% of what it was in Dorn's study, the difference is 2.4 years.

It is also of interest to compare the estimated life table of men who have never smoked with that of continuing cigarette smokers and continuing pipe and/or cigar only smokers. These three life tables are given in Table 20 (Assumption A) and Table 21 (Assumption B) based on the death rate estimates given previously in Tables 16 and 17, and their mortality is compared in Table 22. It can be seen that under Assumption A the expectation of life of continuing cigarette smokers is 4.96 years less than that of never smokers while that of pipe and/or cigar smokers is 0.69 years less. Under Assumption B these differences are 3.39 years and 0.80 years respectively.

Table 23 gives estimates of the average annual consumption of manufactured cigarettes per male smoker by 5 year age groups. These figures are averages for data for the years 1971-75 given in Lee (1976), with the figures for the highest 2 age groups estimated by extrapolation. Taking these estimates as being applicable to the average annual consumption of all cigarette smokers (data for hand-rolled smokers is not available) one can then apply them to the estimated life tables for cigarette smokers given in Table 20 (Assumption A) and Table 21 (Assumption B) to produce estimates of lifetime cigarette consumption. Thus, under Assumption A, the loss of expectation of life of cigarette smokers of 4.96 years can be related to an estimated lifetime average consumption of 398,000 cigarettes to yield an estimate of loss of life per cigarette of 6.6 minutes. Similarly, under Assumption B, the estimated lifetime average consumption of 402,000 cigarettes gives an estimate of loss of life per cigarette of 4.4 minutes.

6. Discussion

From the preceding sections it can be clearly seen that there are right ways and wrong ways to quantify the effect of a disease or a factor causing it on loss of life. That this is not generally recognised can be seen from Miller's recent (1977) observation that people who smoke filter cigarettes die two to four years sooner than smokers of plain cigarettes, and his subsequent attempt to explain this finding in terms of higher blood carbon monoxide levels in the blood of people who smoke filter cigarettes. His study was based on a cross-sectional observation of deaths occurring between 1972 and 1974 in Pennsylvania's Erie County, and the difference in average age at death observed of filter and plain smokers can be explained by the fact that the average age of living filter smokers (due to the switch to filters occurring more in younger smokers) is markedly lower than that of plain smokers. Had Miller compared relative death rates or expectations of life based on life-tables (both of which would have required observation of filter and plain smoking habits in the living population also) he would have doubtless found, in line with the studies we quoted in section 3.2, that filter cigarettes were in fact safer, and not the reverse.

However, though it is easy to see in some cases that particular statistical approaches are wrong it is not so easy, in the case of smoking, to arrive at estimates of loss of life which are right. Death is a fairly rare event, and, therefore, to get information on adequate numbers of deaths of people of different smoking habits, it has been necessary in all the major prospective studies to study deaths over a fairly long period of time. When one additionally takes into account the time required to report the results of these studies, the fact that the tar yield of the average cigarette smoked has reduced rapidly and the evidence that the lowering of average tar yields has been beneficial to health (see section 3.2) it is clear that estimates of the relative death rates of smokers and non-smokers derived from the major prospective studies are virtually certain to be markedly too high, though by how much one cannot know precisely.

It is also possible that the fact that all three of the major prospective studies considered here (see section 4.2) studied nationally unrepresentative populations might have caused a bias in relative risk estimates. However, though the actual levels of mortality in these studies are known to be different from national levels, the consistency of the estimates of relative levels of mortality of never smokers and cigarette only smokers in the three studies (Table 3), makes it seem unlikely that the relative levels of mortality in the U.S. and U.K. populations around the time of the three studies of these two groups differed much from those found in the studies. Unrepresentativeness does not seem a major problem therefore.

A more difficult problem in quantifying the relationship of smoking to mortality is the fact that many smokers modify the amount they smoke during their lives. If, as is often the case, this modification is due to symptoms associated with smoking, it can be seen that construction of life-tables using prospective study data on risk of smoking by age may be somewhat misleading inasmuch as the data used on men in younger age groups is likely to include some information on men who fail to reach older age groups, not because they die, but because they give up. For this reason alone if one took a random sample of the population and forced them to smoke throughout their lives (the analogue of animal experiments) the actual loss of life expectation observed in this group as compared with a control group of continuing non-smokers may well be different to the estimates we calculate. It would of course also be likely to be different because smokers are clearly not a random sample of the whole population in so many ways.

Bearing all these reservations in mind, one can calculate that, assuming the relative risks of smokers and never smokers are as found by Kahn (1966), (Assumption A), and assuming total death rates are those observed in England and Wales in 1971-75, the average male cigarette smoker can expect to live almost 5 years less than the average male never smoker. Put another way, one can show that 23% of male cigarette smokers are likely to die before age 60, if they continue to smoke, as compared with 13½% of male never smokers. If the relative risks of cigarette smokers to never smokers are assumed to be 60% of those given by Kahn (Assumption B), as may be the case nowadays due to the switch to lower tar cigarettes, then the loss of life expectation related to smoking falls to 3.39 years.

This loss of life expectation can be expressed on a per cigarette smoked basis by dividing by the estimated average total lifetime cigarette consumption of cigarette smokers (about 400,000 cigarettes) to give figures of 6.6 minutes (Assumption A) and 4.4 minutes (Assumption B). However, apart from the fact that it assumes, which is unlikely to be the case (see section 3.1), that excess death rates of smokers are wholly due to their smoking, this estimate is open to criticism on two grounds. Firstly, as noted in section 2.8, the theoretically proper way to compute average loss per cigarette smoked is to calculate, for each person, his loss per cigarette and then to average it (which is not practical as one cannot estimate life expectation on an individual basis) and not as we have done to divide average loss by average number of cigarettes smoked. Secondly, and more importantly, the argument of sections 2.8 and 5.1, show clearly that in fact each cigarette does not take an equal amount off life expectation. The use of this statistic to quantify loss of life related to smoking cannot be recommended.

7. Summary

The merits of a number of different ways of estimating loss of life in relation to a disease or to a factor causing it are discussed. It is concluded that those based on life-tables, such as "expectation of life" or "percentage dying between given ages", are likely to be the most useful. Among statistics to be avoided are "average age at death" based on cross-sectional observations, which can lead to gross bias, and "number of deaths attributable to a factor" which is difficult to interpret and gives little useful information. The problems involved in inferring loss of life expectation due to a single exposure to a factor from observations on a population exposed continuously or at regular intervals throughout their lives are stressed, and it is shown that the assumption that each exposure leads to an identical loss of life expectation implies the existence of a particular mathematical relationship between the life-tables of exposed and non-exposed populations.

The choice of appropriate data to use in constructing life-tables, where full lifetime survival history of a population is not available, is discussed and it is concluded that life-tables constructed from up-to-date cross-sectional data are most useful as a basis for quantifying the current situation. The problems of constructing such life-tables for assessing loss of life due to smoking are noted, particularly important being the fact that the major sources of reliable information on the relative risk of death by age of smokers and non-smokers all refer to a time period when the average cigarette smoked had a much higher tar yield than it does today.

A number of applications of the methodology are made to estimating loss of life related to smoking. For example, it is shown that, assuming the whole of the excess of the death rates of smokers over non-smokers is due to their smoking, and assuming that their relative risks of death by age are as found in the U.S. Veteran's study, the life expectation of the male population of England and Wales would be 3.4 years greater if none of them had ever smoked. The difference in life expectation of never smokers and continuing smokers under these same assumptions is almost 5 years which, related to the estimated lifetime cigarette consumption of smokers of about 400,000 cigarettes, means an estimated average loss of life per cigarette of 6.6 minutes. However comparison of the life-tables of smokers and non-smokers makes it quite clear that each cigarette smoked does not have the same effect on loss of life.

APPENDIX A - "Deaths associated with smoking" - an estimate of the reduction in deaths that would occur over 25 years if the population had never smokers death rates

As we point out in Section 2.4, the "number of deaths associated with smoking" as calculated by, e.g. R.C.P. (1971), is, in fact, an estimate of the number of deaths that would not have occurred in the year following a universal giving up of smoking assuming that on giving up, smokers age-specific death rates reverted at once to those of never smokers. This statistic ignores completely the fact that everyone must die at some time and only measures a short-term effect. The purpose of this Appendix is to estimate, under certain assumptions, the number of male deaths that would occur in England and Wales each year for 25 years given the population at large had non-smokers death rates and to compare these numbers with those that would occur if the population kept to present death rates.

The present (1971 - 1975) average distribution of the England and Wales population and current numbers of deaths by 5 year age groups has been given previously in Table 2. Estimated death rates of non-smokers under two assumptions A and B (see Section 4.5) have also been given in Tables 16 and 17. To compute, for a particular set of death rates, expected numbers of deaths occurring in each of the following 25 years, calculations were made for single year cohorts. For each cohort the number alive at the beginning of the 25 year period was calculated from the population figures for the corresponding age group in Table 2 assuming that the age-distribution within the population was uniform. (Thus the cohort aged 32 years old initially was assumed to number 301,404 initially, one fifth of the 1,507,020 people aged 30-34 in Table 2). The numbers in the cohort were then successively reduced by applying the death rate for that 5 year age group, taking into account that in each successive year the cohort would be one degree older. Cohorts born after the beginning of the period were started off with 356,300 alive (i.e. assuming the birth rate of Table 2 remained constant). Cohorts reaching ages of 90 or over had the death rate of the 90+ group applied to them thereafter.

Table A1 shows the numbers of deaths that would be expected to occur assuming

i) death rates stayed constant at present rates

ii) death rates stayed constant at never smokers rates estimated under Assumption A

iii) death rates stayed constant at never smokers rates estimated under Assumption B.

It also shows the differences between (ii) and (i), and between (iii) and (i), i.e. in the "deaths associated with smoking" under the two assumptions.

The estimates show that, assuming historical figures of the relative risk of smoking derived from past prospective studies apply nowadays (Assumption A), 72,000 less deaths of males would occur in England and Wales in 1976 if the population at large had never smokers death rates. However if the population continued to have non-smokers death rates, these estimates of deaths associated with smoking decrease steadily with time, reaching 36,000 by 1988 and 12,000 by the year 2,000. Under Assumption B, where one is assuming that the relative risk of smoking applying nowadays is rather less than that under Assumption A due to the switch to lower tar filter cigarettes, the estimated deaths associated with smoking are consistently about three quarters of those under Assumption B.

It should be noted that, in fact, as time passes over the period 1976 - 2000 an increasing number of smokers (and ex-smokers) will have smoked predominantly lower tar filter cigarettes during their lifetime and that it is somewhat unrealistic therefore to assume both constancy of death rates assuming that no smokers give up and as high relative risks as those used for Assumption B. We have not made calculations taking this fact into account as the data is not available to do this at all accurately. However it is clear that the effect of such a calculation would be to decrease the deaths associated by a proportion that would increase over the time period considered. In other words the drop-off in "deaths associated" would be even steeper than we suggest in Table A1.

APPENDIX B - Calculation of adjusted life-tables not assuming independence

Given a life-table for a single group in which, for each time interval i (i = 1 ... n) we observe the number alive at the beginning of the interval A_i, the total number dying D_i and the number dying from a particular cause of interest L_i. We wish to estimate the life-table that would have existed had deaths from the particular cause of interest not occurred, i.e. the adjusted life-table. We assume that the population consists of two groups of individuals, one groups ("susceptibles") having F times the susceptibility to death from all causes of the other group ("normals"). We also assume that both F and the proportion of susceptibles at the beginning of the first interval p_i are known.

Wong (1977) derived the following formula for the probability, p_i, of dying in the ith interval given the cause of death of interest is removed.

$$p_i = \frac{(D_i - L_i)}{A_i} + F \frac{L_i}{A_i} \frac{(D_i - L_i)}{A_i}$$

Now $(D_i-L_i)/A_i$ is simply the crude probability of dying from causes other than the one of interest, and L_i/A_i is the crude probability of dying from the cause of interest. It follows his formula "reincarnates" the whole of the L_i people dying from the cause to be removed and applies the former crude probability multiplied by F to them, ignoring the fact that these decedents are already known to have survived the other causes of death for a part (on average, about half) of the interval. Furthermore, Wong does not take into account in his formulae the fact that in subsequent intervals the reincarnated survivors will have a greater risk of dying assuming their extra susceptibility continues. In the following paragraphs, therefore, we consider the estimation afresh.

Consider the ith time interval and let the proportion of the population who are susceptibles be p_i. Let the length of this interval be T_i and let us assume it to be small enough for the force of mortality in the normal population from causes other than the one of interest to be taken as constant (= α_i), and that from the cause of interest also to be taken as constant (= β_i). In the susceptibles the forces of mortality will therefore be $F\alpha_i$ and $F\beta_i$ respectively.

Now the total survival from all causes

$$P_i \exp(-F(\alpha_i+\beta_i)T_i) + (1-P_i) \exp(-(\alpha_i+\beta_i)T_i)$$

is estimated by

$$(A_i-D_i)/A_i \qquad (1)$$

and the relative mortality

$$\alpha_i/\beta_i \text{ by } \qquad (D_i-L_i)/L_i \qquad (2)$$

Substituting from (1) into (2) and writing

$$Q = \exp(-\alpha_i T_i D_i/(D_i-L_i)) \qquad (3)$$

this gives the equation

$$P_i Q^F + (1-P_i)Q = (A_i-D_i)/A_i \qquad (4)$$

It can easily be seen that this equation has one real root in the range (0,1) which can be obtained iteratively without difficulty.

The estimated survival rate in the normal population due to all causes except the one to be eliminated is thus given by

$$\exp(-\alpha_i T_i) = Q^{(D_i-L_i)/D_i} \qquad (5)$$

Similarly

$$\exp(-\beta_i T_i) = Q^{L_i/D_i} \qquad (6)$$

It is now possible to define a sequence of recursive equations to show how the life-table estimate is built up. Let, at the beginning of the interval, values of P_i, A_i, D_i, L_i, H_i and J_i be known, where the first four terms have been defined previously, and H_i and J_i are defined as follows.

H_i number of living "reincarnated" susceptibles, i.e. the additional number of susceptibles that would have been alive had the cause of death of interest been eliminated.

J_i number of living "reincarnated" normals.

It follows that

A_iP_i is the true number of susceptibles alive.

$A_i(1-P_i)$ is the true number of normals alive

$\frac{A_i+H_i+J_i}{A_1}$ is the adjusted life-table estimate.

We now estimate $\exp(-\alpha_iT_i)$ and $\exp(-\beta_iT_i)$ from equations (5) and (6). The total number of susceptibles dying in the interval (DS_i) is given by the expression

$$DS_i = A_iP_i(1-Q^F)$$

and the total number of normals dying (DN_i) by

$$DN_i = A_i(1-P_i)(1-Q)$$

It follows that P_{i+1}, the proportion susceptible at the beginning of the next interval (or end of this one), is given by

$$P_{i+1} = (A_iP_i-DS_i)/(A_i-D_i) = (A_iP_i-DS_i)/A_{i+1} \qquad (7)$$

Of the living reincarnated susceptibles H_i, a proportion $1-\exp(-F\alpha_iT_i)$ will "die" in the interval due to causes other than the one eliminated. However, there will also be an increase in their numbers due to the difference between the actual force of mortality occurring and that that would have occurred had the cause of death been eliminated. It follows that

$$\begin{aligned} H_{i+1} = {} & (H_i+A_iP_i)\ \exp(-F\alpha_iT_i) \\ & - A_iP_i\ \exp(-F(\alpha_i+\beta_i)T_i) \end{aligned} \qquad (8)$$

Similarly

$$\begin{aligned} J_{i=1} = {} & (J_i+A_i(1-P_i))\ \exp\ (-\alpha_iT_i) \\ & - A_i(1-P_i)\ \exp(-(\alpha_i+\beta_i)T_i) \end{aligned} \qquad (9)$$

Given that, at the beginning of the first interval H_1 and J_1 are defined as 0, and given, as noted above the user supplies estimates of P_1 and F, the above equations supply all that is needed for calculation of the adjusted life-table.

APPENDIX C - Estimation of distribution of England and Wales male population by smoking habit and age

Table 6 gives the estimated distribution of the England and Wales male population by smoking habit and age. This was derived from Lee (1976), assuming such data for the United Kingdom as a whole could be applied to England and Wales, as follows.

Firstly, Table 11M was used to extract the following data on percentage of smokers of each type of product by age for 1971 to 1975 by averaging the figures given for 1971 and 1975.

		Age group	
		35-59	60+
1.	All smokers of manufactured cigarettes	49.5%	42.0%
2.	Hand-rolled cigarettes only	5.0%	5.5%
3.	Pipe only	4.5%	9.5%
4.	Cigars only	4.0%	2.5%
5.	Ex-smokers	19.5%	26.0%
6.	Have never smoked	15.5%	11.0%
	Total	98.0%	96.5%

The reason the figures do not add up to 100% is that no data is available for two smoking habit groups; 7. hand-rolled cigarettes and (pipe and/or cigars) and 8. pipe and cigars (but no cigarettes). Arbitrarily assigning the residue equally (1% each) to both categories for 35-59 year olds and 2% to category 7 and 1½% to category 8 for 60+ year olds, one could then compute the distribution by the four smoking habit groups of interest as follows.

	Age group	
	35-59	60+
Current smokers of cigarettes (categories 1, 2 and 7)	55.5%	49.5%
Current smokers of pipes and/or cigars only (categories 3, 4 and 8)	9.5%	13.5%
Ex-smokers	19.5%	26.0%
Never smoked	15.5%	11.0%
Total	100.0%	100.0%

It should be noted that the smoking habit group, current smokers of cigarettes, was defined so as to include both manufactured and hand-rolled cigarette smokers. This was to bring the data in line with that from the three prospective studies, in all of which such a definition had been used.

Next we wished to check whether the percentages derived could be taken as validly applying to the whole of the age range they contain. Lee does not give information directly on the distribution of the population by smoking habits and five year age ranges but some information can be gained on this indirectly from Table 14M which gives the annual consumption of manufactured cigarettes per adult by 5 year age groups and Table 17M which gives the annual consumption per smoker. Averaging the figures given for each year from 1971-75 in Table 14M and dividing by the corresponding averages for Table 17M gives estimates of the proportion of the male population smoking manufactured cigarettes as follows.

Age group	Percentage of manufactured cigarette smokers
35-39	50.6%
40-44	50.5%
45-49	54.0%
50-54	50.7%
55-59	45.4%
60-64	45.4%
65-69	44.5%
70-74	37.9%
75-79	30.8%
80+	25.3%

It can be seen that, within the 35-59 age range, the percentage of manufactured cigarette smokers is fairly constant within each 5 year age group. For this age range therefore we propose to assume that the distribution of the population by the 4 smoking habit groups given above applies to each 5 year age group within it, ignoring variations which are essentially minor.

For the 60+ age range, however, it is quite clear that the percentage of manufactured cigarette smokers, and hence presumably that of all cigarette smokers, decreases steadily with age. Within this age range we therefore estimated the percentage of all cigarette smokers separately for each 5 year age group by multiplying the above figures by the ratio 49.5/42.0 observed for the ratio % of all cigarette smokers/ % of manufactured cigarette smokers for the 60+ age group as a whole.

Rounding the answers to the nearest ½% this gives

Age group	Estimated percentage of current cigarette smokers
60-64	53.5%
65-69	52.5%
70-74	44.5%
75-79	36.5%
80+	30.0%

Next, as 66% of the 80+ age group are aged 80-84, and as it is likely that a much higher proportion than this were actually interviewed, it was decided to take the estimate of 30.0% given in the table above as applying to the 80-84 age group. Estimates of the percentage of current cigarette smokers for ages 85-89 and 90+ were made by extrapolation, assuming the proportional reduction in percentages of current cigarette smokers observed over the age range 65-69 to 80-84 persisted. This gave figures of 25.0% for 85-89 and 21.0% for 90-94 year olds.

Finally it was assumed that the reductions observed in the percentage of cigarette smokers were balanced by corresponding increases in the percentage of ex-smokers, with percentages of never smokers and pipe and/or cigars only remaining constant throughout the 60+ age range. This resulted in the figures presented in Table 6.

It should be noted that, for our purposes, quite marked inaccuracies in the estimates of the relative proportion of ex-smokers, never smokers and pipe and/or cigar smokers at older ages will make little difference to our conclusions as the relative risk observed in these three groups is not very different (see Table 4). The essential is that the proportion of current cigarette smokers should be reasonably accurate, and it seems likely this is so.

8. References

Benjamin, B. and Haycocks, H.W. (1970). The analysis of mortality and other actuarial statistics. Cambridge University Press, London.

Bross, I.D.J. and Gibson, R. (1968). Risks of lung cancer in smokers who switch to filter cigarettes. American Journal of Public Health, 58, 1396.

Burch, P.R.J. (1976). The biology of cancer, a new approach. MTP Press Ltd., Lancaster.

Chiang, C.L. (1961). On the probability of death from specific causes in the presence of competing risks. Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability, 169.

Dean, G., Lee, P.N., Todd, G.F. and Wicken, A.J. (1977). Report on a second retrospective mortality study in North-East England. Tobacco Research Council Research Paper 14, Part I.

Dempsey, M. (1947). Decline in tuberculosis; the death rate fails to tell the entire story. American Review of Tuberculosis, 86, 157.

Dickinson, F.G. and Welker, E.L. (1948). What is the leading cause of death? Two new measures. Bureau of Medical Economic Research, American Medical Association, Bulletin 64, Chicago.

Diehl, H.S. (1969). Tobacco and your health. McGraw-Hill, New York.

Doll, R. and Hill, A.B. (1952). A study of the aetiology of carcinoma of the lung. British Medical Journal, 2, 1271.

Doll, R. and Hill, A.B. (1964). Mortality in relation to smoking: 10 years' observations of British doctors. British Medical Journal, Part 1, 5395, 1399, Part 2.

Doll, R. (1971). The age distribution of cancer: implications for models of carcinogenesis. Journal of the Royal Statistical Society, Series A, 134, 2, 133.

Doll, R. and Peto, R. (1976). Mortality in relation to smoking: 20 years' observations on male British doctors. British Medical Journal, 25th December, 1525.

Fisher, R.A. (1959). Smoking, the cancer controversy. Oliver and Boyd, Edinburgh.

Gross, A.J. and Clark, V.A. (1975). Survival distributions: reliability applications in the biomedical sciences. John Wiley and Sons, New York.

Haenszel, W. (1950). A standardised rate for mortality defined in units of lost years of life. American Journal of Public Health, 40, 17.

Hakulinen, T. and Teppo, L. (1976). The increase in working years due to elimination of cancer as a cause of death. International Journal of Cancer, 17, 429.

Halley, E. (1693). An estimate of the degrees of the mortality of mankind, drawn from curious tables of the births and funerals of the city of Breslau. Philosophical Transactions of the Royal Society of London, 17, 596.

Hammond, E.C. (1966). Smoking in relation to the death rates of one million men and women. National Cancer Institute Monograph 19, 127.

Hammond, E.C. (1969). Life expectancy of American men in relation to their smoking habits. Journal of the National Cancer Institute, 43, 4, 951.

Hammond, E.C., Garfinkel, L., Seidman, H. and Lew, E.A. (1976). Tar and nicotine content of cigarette smoke in relation to death rates. Environmental Research, 12, 263.

Kahn, H.A. (1966). The study of smoking and mortality among U.S. veterans: report on eight and one-half years observation. National Cancer Institute Monograph 19, 1.

Kitagawa, E.M. (1966). Theoretical considerations in the selection of a mortality index, and some empirical comparisons. Human Biology, 38, 3, 293.

Lee, P.N. (1970). The simulated population method of analysis of animal painting experiments in cancer research. Biometrics, 26, 4, 777.

Lee, P.N. (1976). Statistics of smoking in the United Kingdom. Tobacco Research Council Research Paper 1, 7th Edition.

Logan, W.P.D. and Benjamin, B. (1953). Loss of expected years of life - a perspective view of changes between 1848-72 and 1952. Monthly Bulletin of the Ministry of Health and Public Health Laboratory, 12, 244.

Miller, G.H. (1977). Study finds it's safer to smoke nonfiltered cigarettes. New York Daily News, 10th February.

Murray, J.L. and Axtell, L.M. (1974). Impact of cancer: years of life lost due to cancer mortality. Journal of the National Cancer Institute, 52, 3.

O.P.C.S. (1975). The General Household Survey 1972. Her Majesty's Stationery Office, London.

Passey, R.D. (1962). Some problems of lung cancer. Lancet, July 21st, 107.

Peto, R. and Lee, P.N. (1973). Weibull distributions for continuous-carcinogenesis experiments. Biometrics, 29, 457.

Pike, M.C. and Doll, R. (1965). Age at onset of lung cancer: significance in relation to effect of smoking. Lancet, March 27th, 665.

Reid, D.D., Hamilton, P.J.S., McCartney, P., Rose, G.A., Jarrett, R.J. and Keen, H. (1976). Smoking and other risk factors for coronary heart disease in British civil servants. Lancet, November 6th, 979.

Romeder, J-M. and McWhinnie, J.R. (1977). Potential years of life lost between ages 1 and 70: an indicator of premature mortality for health planning. International Journal of Epidemiology, 6, 2, 143.

Rose, G.A. (1977). Private communication.

R.C.P. (1971). Smoking and health now. The second report from the Royal College of Physicians of London. Pitman, London.

R.C.P. (1977). Smoking or health. The third report from the Royal College of Physicians of London. Pitman, London.

Schwartz, D. et Lazar, Ph. (1961). Taux de mortalité par une cause donnée de décès en tenant compte des autres causes de décès ou de disparition. Revue Institute International de Statistique, 29, 44.

Todd, G.F. (1977). Private communication.

Wong, O. (1977). A competing-risk model based on the life table procedure in epidemiological studies. International Journal of Epidemiology, 6, 2, 153.

Woolsey, T.D. (1943). Adjusted death rates and other indices of mortality. Chapter IV of Techniques of Vital Statistics, Reprint of Chapters I-IV of Vital Statistics Rates in the United States 1900-1940, by Leider, F.E. and Grove, P. National Office of Vital Statistics, 91. U.S. Government Printing Office, Washington, D.C., 1959 (original 1943).

Wynder, E.L., Mabuchi, K. and Beattie, E.J. (1970). The epidemiology of lung cancer, recent trends. Journal of the American Medical Association, 213, 2221.

Yerushalmy, J. (1951). A mortality index for use in place of the age-adjusted death rate. American Journal of Public Health, 41, 8, 907.

TABLE 1

Survivorship of men aged 25 in relation to current number of cigarettes smoked per day: based on rates adjusted to the 1959-61 U.S. life-table for all males (reproduced from Hammond (1969), Table 6)

Age (yr)	All men (%)	Never smoked regularly (%)	Smokers: Current No. of cigarettes per day			
			1-9 (%)	10-19 (%)	20-39 (%)	40+ (%)
25	100.0	100.0	100.0	100.0	100.0	100.0
30	99.1	99.4	99.1	99.1	99.1	98.8
35	98.2	98.7	98.1	98.1	98.0	97.3
40	96.8	97.8	96.6	96.5	96.5	95.1
45	94.6	96.4	94.2	94.0	93.8	91.0
50	91.1	94.4	90.6	90.0	89.3	85.6
55	85.6	90.9	85.9	83.8	82.5	77.7
60	78.1	85.5	77.8	75.3	73.5	67.1
65	67.8	77.7	67.3	63.4	61.1	54.0
70	55.2	66.7	52.4	47.7	45.9	40.0
75	41.2	52.3	36.2	33.3	30.3	25.7
80	26.7	35.6	20.6	18.6	18.1	14.3
85	13.6	19.2	7.3	8.5	7.2	6.5
90	4.9	7.0	2.2	2.2	2.2	2.1
95	1.0	1.5	0.5	0.5	0.5	0.4
Life expectancy						
i) as age	70.2	73.6	69.0	68.1	67.4	65.3
ii) in years	45.2	48.6	44.0	43.1	42.4	40.3
Difference from never smoked regularly rate in years	3.4	0	4.6	5.5	6.2	8.3

TABLE 2

Population, total deaths and deaths from lung cancer by age

Males - England and Wales (1971 - 75)

Age group	Population	Total deaths per year	Deaths from lung cancer per year
0	356,300	6,670	0
1-4	1,544,760	1,149	1
5-9	2,061,000	795	0
10-14	1,972,380	680	0
15-19	1,779,120	1,539	1
20-24	1,790,980	1,745	6
25-29	1,811,160	1,587	17
30-34	1,507,020	1,604	35
35-39	1,439,340	2,269	83
40-44	1,439,580	4,092	251
45-49	1,489,740	7,891	743
50-54	1,518,080	13,815	1,616
55-59	1,332,360	20,402	2,756
60-64	1,324,300	33,368	4,689
65-69	1,092,120	45,517	5,672
70-74	745,460	49,682	5,121
75-79	415,180	42,231	3,072
80-84	218,130	32,653	1,341
85-89	85,320	18,671	411
90+	26,370	8,594	83
Total	23,948,700	294,954	25,898

TABLE 3

Total and relative death rates of never smokers and cigarette only smokers by age taken from three prospective studies

Age group	Study	Death rates (per 100,000 per year) Never smokers	Cigarette only smokers	Relative rates
35-44	Doll[1]	142	266[4]	1.87
	Dorn[2]	127	222[5]	1.75
	Hammond[3]	211	384[6]	1.82
45-54	Doll	360	819	2.28
	Dorn	264	758	2.87
	Hammond	402	886	2.20
55-64	Doll	1086	2134	1.97
	Dorn	1056	1942	1.84
	Hammond	1187	2207	1.86
65-74	Doll	2907	5327	1.83
	Dorn	2411	4313	1.79
	Hammond	3118	4918	1.58
75-84	Doll	8145	11150	1.37
	Dorn	6214	9581	1.54
	Hammond	7897	10635	1.35

Key

1. Calculated from Table XIII, Doll and Peto (1976)
2. From Appendix Table A, pp. 30 and 46, Kahn (1966)
3. From Table 4, Hammond (1966)
4. Started smoking before 25 years of age and continuing to smoke and no history of ever smoking anything but cigarettes
5. Current smokers of cigarettes only
6. Cigarettes only (lifetime history).

TABLE 4

Total and relative death rates of four smoking habits by age from the Dorn Study (Kahn (1966)) - Assumption A

Age group	Smoking group	Death rate (per 100,000 per year)	Relative rates to never smokers
35-44	Never smokers	127	1.00
	Current cigarette smokers	232	1.83
	Current pipe and/or cigars only	180	1.42
	Ex-smokers	134	1.06
45-54	Never smokers	264	1.00
	Current cigarette smokers	728	2.76
	Current pipe and/or cigars only	377	1.43
	Ex-smokers	359	1.36
55-64	Never smokers	1056	1.00
	Current cigarette smokers	1819	1.72
	Current pipe and/or cigars only	1100	1.04
	Ex-smokers	1328	1.26
65-74	Never smokers	2411	1.00
	Current cigarette smokers	4032	1.67
	Current pipe and/or cigars only	2633	1.09
	Ex-smokers	3077	1.28
75+	Never smokers	6214	1.00
	Current cigarette smokers	8471	1.36
	Current pipe and/or cigars only	5782	0.93
	Ex-smokers	7013	1.13

TABLE 5

Relative total death rates of current cigarette smokers to never smokers by age under Assumption B (see text)

Age group	Smoking group	Relative rates to never smokers
35-44	Never smokers	1.00
	Current cigarette smokers	1.50
45-54	Never smokers	1.00
	Current cigarette smokers	2.06
55-64	Never smokers	1.00
	Current cigarette smokers	1.43
65-74	Never smokers	1.00
	Current cigarette smokers	1.40
75+	Never smokers	1.00
	Current cigarette smokers	1.22

Note: Relative total death rates of current pipe and/or cigars only smokers and ex-smokers assumed the same as in Table 4.

TABLE 6

Estimated distribution of England and Wales male population by smoking habit group and age

Age group	Never smokers	Current cigarette smokers	Current pipe and/or cigar only smokers	Ex-smokers	Total
35-59	15.5%	55.5%	9.5%	19.5%	100.0%
60-64	10.5%	53.5%	12.5%	23.5%	100.0%
65-69	10.5%	52.5%	12.5%	24.5%	100.0%
70-74	10.5%	44.5%	12.5%	32.5%	100.0%
75-79	10.5%	36.5%	12.5%	40.5%	100.0%
80-84	10.5%	30.0%	12.5%	47.0%	100.0%
85-89	10.5%	25.0%	12.5%	52.0%	100.0%
90-94	10.5%	21.0%	12.5%	56.0%	100.0%

TABLE 7

Percentage of men aged 35 who will die before they reach the age of 65 according to their smoking habits

Study	Never smokers	Cigarette Smokers						
		1-9	1-14	10-19	15-24	20-39	25+	40+
Doll	18%	-	25%	-	31%	-	40%	-
Hammond	21%	31%	-	35%	-	38%	-	45%

Source: Doll given in R.C.P. (1971)
Hammond calculated from Table 1.

TABLE 8

Loss of life expectation per cigarette smoked based on Hammond's data

	Smokers : Current No. of cigarettes per day			
	1-9	10-19	20-39	40+
Loss of life expectation in years (from Table 1)	4.6	5.5	6.2	8.3
Loss of life expectation in minutes (millions)	2.42	2.89	3.26	4.36
Average age at death	69.0	68.1	67.4	65.3
Average number of years smoked (assuming starting at age 16)	53.0	52.1	51.4	49.3
Assumed average number of cigarettes per day	4	13	25	50
Total number of cigarettes smoked to average age at death (thousands)	77.4	247.2	469.0	899.7
Loss of life expectation in minutes per cigarette smoked	31.2	11.7	6.9	4.8

TABLE 9

Test that each cigarette takes an equal amount off life expectation based on Hammond's data (see text)

Age at which survivorship proportion is	Never smokers	Smokers : Current No. of cigarettes per day							
		1-9		10-19		20-39		40+	
		Age	Ratio	Age	Ratio	Age	Ratio	Age	Ratio
95%	48.5	43.3	0.89	43.0	0.89	42.8	0.88	40.1	0.83
90%	55.8	50.6	0.91	50.0	0.90	49.2	0.88	45.9	0.82
75%	66.2	61.4	0.93	60.1	0.91	59.2	0.89	56.3	0.85
50%	75.7	70.7	0.93	69.3	0.92	68.7	0.91	66.4	0.88
25%	83.2	78.6	0.94	77.8	0.94	77.2	0.93	75.3	0.91
10%	88.8	84.0	0.95	84.3	0.95	83.7	0.94	82.8	0.93
5%	91.8	87.2	0.95	87.8	0.96	87.2	0.95	86.7	0.94

TABLE 10

Population life-table based on 1971-75 mortality experience of men in England and Wales

Age	Number alive at beginning (A_i)	Total deaths in period (D_i)	Lung cancer deaths in period (L_i)
0	1,000,000	18,720	0
1-4	981,280	2,916	3
5-9	978,364	1,886	0
10-14	976,478	1,682	0
15-19	974,796	4,209	3
20-24	970,587	4,719	16
25-29	965,868	4,224	45
30-34	961,644	5,107	111
35-39	956,537	7,516	275
40-44	949,021	13,411	823
45-49	935,610	24,518	2,309
50-54	911,092	40,709	4,762
55-59	870,383	64,629	8,730
60-64	805,754	96,524	13,564
65-69	709,230	135,978	16,945
70-74	573,252	167,204	17,235
75-79	406,048	168,560	12,262
80-84	237,488	131,924	5,418
85-89	105,564	74,858	1,648
90+	30,706	30,706	297
Total		1,000,000	84,446

TABLE 11

Effect of lung cancer on life expectation in men in England and Wales, 1971-75

Life expectation of men as at present	69.06 years
Life expectation of men if lung cancer did not exist, assuming independence (see section 2.10)	70.05 years
Loss of life expectation due to lung cancer :	
per head of population	0.99 years
per person dying of lung cancer	11.73 years
Life expectation of men if lung cancer did not exist, not assuming independence (see text)	

Proportion susceptible (P)	Ratio of death rates of susceptibles to normals (F) 2	5	10	50
0.05	70.04	70.02	70.00	70.01
0.10	70.03	69.98	69.95	69.93
0.25	70.01	69.90	69.82	69.72
0.50	70.00	69.84	69.74	69.61
0.75	70.01	69.89	69.81	69.71
0.90	70.03	69.97	69.93	69.88
0.95	70.04	70.00	69.98	69.96

TABLE 12

Population life-table based on 1971-75 mortality experience of men in England and Wales : further data

Age	Ischaemic Heart Disease deaths in period	All neoplasms deaths in period
0	3	79
1-4	1	341
5-9	1	379
10-14	2	320
15-19	7	393
20-24	29	548
25-29	127	721
30-34	518	970
35-39	1,628	1,542
40-44	4,466	2,988
45-49	9,561	6,124
50-54	16,251	11,063
55-59	23,927	19,173
60-64	33,962	28,682
65-69	45,499	37,601
70-74	51,372	40,602
75-79	47,512	33,994
80-84	33,659	20,332
85-89	17,481	8,541
90+	6,472	2,123
	292,478	216,516

TABLE 13

Comparison of effects of three causes of death on loss of life in men in England and Wales, 1971-75

	Lung Cancer	Ischaemic Heart Disease	All Neoplasms
Percentage dying from cause	8.4%	29.2%	21.7%
Loss of life expectation due to cause (assuming independence)			
per head of population	0.99 years	3.57 years	2.76 years
per person dying of cause	11.73 years	12.22 years	12.75 years
Loss of life expectation due to cause (not assuming independence P = 0.5, F = 10)			
per head of population	0.68 years	2.51 years	1.95 years
per person dying of cause	8.06 years	8.59 years	9.01 years

<u>TABLE 14</u>

Remaining life expectancy by age based on 1971-75 mortality experience of men in England and Wales

Age	Expected age of death of survivors to this age (X_i) *	Expected years remaining of survivors to this age	Expected years remaining of decedents in this age group +
0	69.06	69.06	69.21
1	70.36	69.36	67.46
5	70.56	65.56	63.13
10	70.69	60.69	58.24
15	70.79	55.79	53.41
20	71.02	51.02	48.64
25	71.26	46.26	43.86
30	71.45	41.45	39.06
35	71.66	36.66	34.30
40	71.93	31.93	29.64
45	72.35	27.35	25.19
50	73.02	23.02	21.00
55	73.98	18.98	17.14
60	75.30	15.30	13.67
65	77.04	12.04	10.67
70	79.30	9.30	8.20
75	82.10	7.10	6.24
80	85.37	5.37	4.66
85	88.95	3.95	3.23
90	92.50	2.50	1.25

* X_i calculated from formula $X_i = (\sum_{j=i}^{20} D_i Y_i)/A_i$ where D_i and A_i are given in Table 10 and Y_i is midpoint of age interval

\+ By interpolation from second column

TABLE 15

Estimation of death rates by smoking habit group for age group 60-64 under Assumption A

From Table 2 :

Population 1,324,300 Total deaths per year 33,368

∴ Death rate per 100,000 per year = 2,519.67

	Never smokers	Current cigarette smokers	Current pipe and/or cigars only smokers	Ex-smokers
Population distribution (from Table 6)	10.5%	53.5%	12.5%	23.5%
∴ Numbers in groups	139,051.5	708,500.5	165,537.5	311,210.5
Death rates (from Table 4)	1.00X	1.72X	1.04X	1.26X
	where X has to be estimated			
Number of deaths	139,051.5X	1,218,620.9X	172,159.0X	392,125.2X
∴ Total number of deaths	= 1,921,956.6X = 33,368			
∴ X	= 1,736.15 per 100,000 per year			
∴ Death rates per 100,000 per year	1,736.15	2,986.17	1,805.59	2,187.55

TABLE 16

Estimated death rates (per 100,000 per year) by smoking habit group and age - Assumption A

Age group	Never smokers	Current cigarette smokers	Current pipe and/or cigar only smokers	Ex-smokers
0	1,872.02			
1-4	74.38			
5-9	38.57		Rates in other groups at	
10-14	34.48	....	ages less than 35 assumed	....
15-19	86.50		same as in never smokers	
20-24	97.43			
25-29	87.62			
30-34	106.44			
35-39	104.24	190.76	148.03	110.50
40-44	187.96	343.98	266.91	199.24
45-49	253.70	700.21	362.79	345.03
50-54	435.87	1,203.00	623.29	592.78
55-59	1,053.07	1,811.28	1,095.19	1,326.87
60-64	1,736.15	2,986.17	1,805.59	2,187.55
65-69	2,911.26	4,861.81	3,173.28	3,726.42
70-74	4,759.08	7,947.66	5,187.39	6,091.62
75-79	8,654.58	11,770.23	8,048.76	9,779.68
80-84	12,900.86	17,545.17	11,997.80	14,577.97
85-89	19,048.18	25,905.52	17,714.80	21,524.44
90+	28,596.56	38,891.32	26,594.80	32,314.11

TABLE 17

Estimated death rates (per 100,000 per year) by smoking habit group and age - Assumption B

Age group	Never smokers	Current cigarette smokers	Current pipe and/or cigar only smokers	Ex-smoker
<35	....	Rates same as under Assumption A:	see Table 16	....
35-39	118.61	177.91	168.42	125.72
40-44	213.87	320.80	303.69	226.70
45-49	311.70	642.10	445.73	423.91
50-54	535.52	1,103.17	765.79	728.30
55-59	1,184.14	1,693.32	1,231.50	1,492.01
60-64	1,943.97	2,779.87	2,021.72	2,449.40
65-69	3,231.20	4,523.68	3,522.01	4,135.94
70-74	5,205.71	7,287.99	5,674.22	6,663.31
75-79	9,047.97	11,038.53	8,414.62	10,224.21
80-84	13,385.36	16,330.14	12,448.38	15,125.45
85-89	19,646.72	23,969.00	18,271.45	22,200.79
90+	29,353.81	35,811.64	27,299.04	33,169.80

TABLE 18

Survivorship of a million men in England and Wales based on their 1971-75 mortality experience compared with that of a million never smokers under two assumptions

Age	Number alive at beginning: Total population	Never smokers Assumption A	Never smokers Assumption B
0	1,000,000	1,000,000	1,000,000
1-4	981,280	Same as in total population as smoking assumed to have no effect before age 35	......
5-9	978,364		
10-14	976,478		
15-19	974,796		
20-24	970,587		
25-29	965,868		
30-34	961,644		
35-39	956,537	956,537	956,537
40-44	949,021	951,562	950,878
45-49	935,610	942,653	940,753
50-54	911,092	930,756	926,183
55-59	870,383	910,648	901,634
60-64	805,754	863,698	849,500
65-69	709,230	791,281	770,078
70-74	573,252	682,614	653,449
75-79	406,048	534,925	500,176
80-84	237,488	340,194	311,304
85-89	105,564	170,528	151,754
90+	30,706	59,283	50,835

TABLE 19

Comparison of mortality of total population with that of a population of never smokers under two assumptions based on 1971-75 male England and Wales rates

	Total population	Never smokers Assumption A	Never smokers Assumption B
Expectation of life at birth	69.06 years	72.28 years	71.42 years
Difference from total population	-	+3.22 years	+2.36 years
Percentage dying by age 50	8.89%	6.92%	7.38%
Ratios compared with total population	1.00	0.78	0.83
Percentage dying by age 60	19.42%	13.63%	15.05%
Ratios compared with total population	1.00	0.70	0.77
Percentage dying by age 70	42.67%	31.74%	34.66%
Ratios compared with total population	1.00	0.74	0.81
Percentage dying by age 80	76.25%	65.98%	68.87%
Ratios compared with total population	1.00	0.87	0.90

TABLE 20

Survivorship of a million men in each of three different smoking groups based on 1971-75 England and Wales mortality - Assumption A

Age group	Never smokers	Current cigarette smokers	Current pipe and/or cigar only smokers
	Number alive at beginning		
0	1,000,000	1,000,000	1,000,000
1-35	.... Numbers as in total population life-table: see Table 10		
35-39	956,537	956,537	956,537
40-44	951,562	947,448	949,478
45-49	942,653	931,265	936,875
50-54	930,756	899,114	920,003
55-59	910,648	846,318	891,687
60-64	863,698	772,399	843,916
65-69	791,281	663,758	770,430
70-74	682,614	517,350	655,706
75-79	534,925	341,947	502,389
80-84	340,194	182,824	330,239
85-89	170,528	69,681	174,299
90+	59,282	15,561	65,751

TABLE 21

Survivorship of a million men in each of three different smoking groups based on 1971-75 England and Wales mortality - Assumption B

	Number alive at beginning		
Age group	Never smokers	Current cigarette smokers	Current pipe and/or cigar only smokers
0	1,000,000	1,000,000	1,000,000
1-35	.... Numbers as in total population life-table: see Table 10		
35-39	956,537	956,537	956,537
40-44	950,878	948,059	948,509
45-49	940,753	933,949	933,194
50-54	926,183	903,379	913,559
55-59	901,634	854,637	879,111
60-64	849,500	784,688	826,296
65-69	770,078	681,519	746,079
70-74	653,449	540,700	623,629
75-79	500,176	370,370	465,670
80-84	311,304	206,370	300,061
85-89	151,754	84,624	154,358
90+	50,835	21,500	56,286

TABLE 22

Comparison of mortality of never smokers, continuing cigarette smokers and continuing pipe and/or cigar only smokers under two assumptions based on 1971-75 male England and Wales rates

	Assumption	Never smokers	Current cigarette smokers	Current pipe and/or cigar only smokers
Expectation of life at birth	A	72.28	67.32	71.59
	B	71.42	68.03	70.62
Difference from never smokers	A	-	4.96	0.69
	B	-	3.39	0.80
Percentage dying by age 50	A	6.92%	10.09%	8.00%
	B	7.38%	9.67%	8.64%
Percentage dying by age 60	A	13.63%	22.76%	15.61%
	B	15.05%	21.53%	17.37%
Percentage dying by age 70	A	31.74%	48.27%	34.43%
	B	34.66%	45.93%	37.64%
Percentage dying by age 80	A	65.98%	81.72%	66.98%
	B	68.87%	79.36%	69.99%

TABLE 23

Annual consumption of manufactured cigarettes per male smoker by age (based on 1971-75 data of Lee (1976))

Age group	Annual consumption
16-19	6,640
20-24	7,900
25-29	8,000
30-34	8,250
35-39	8,770
40-44	8,320
45-49	8,460
50-54	8,410
55-59	7,870
60-64	7,560
65-69	6,270
70-74	5,460
75-79	4,930
80-84	4,340*
85-89	3,840*
90+	3,400*

* Estimated assuming Lee (1976)'s 80+ figure applies to 80-84 year olds and extrapolating to older age groups using average proportionate reduction per 5 year group observed from age group 65-69 onwards.

TABLE A1

Estimated future deaths in England and Wales males under three assumptions

	Deaths occurring (thousands) if death rates				
	Stay as in 1971-75	Drop to never smokers' rates under assumption		"Deaths associated" under assumption	
Year		A	B	A	B
1976	295	223	240	72	55
1977	301	232	249	69	52
1978	305	238	255	67	50
1979	307	243	259	64	48
1980	307	246	262	61	45
1976-1980	1515	1182	1265	333	250
1981	306	249	264	57	42
1982	312	257	272	55	40
1983	315	264	277	51	38
1984	317	269	281	48	36
1985	317	272	284	45	33
1981-1985	1567	1311	1378	256	189
1986	317	275	286	42	31
1987	321	282	293	39	28
1988	323	287	297	36	26
1989	324	291	300	33	24
1990	325	294	302	31	23
1986-1990	1610	1429	1478	181	132
1991	324	296	304	28	20
1992	326	300	308	26	18
1993	327	304	310	23	17
1994	328	306	312	22	16
1995	327	307	313	20	14
1991-1995	1632	1513	1547	119	85
1996	327	308	314	19	13
1997	328	311	316	17	12
1998	328	313	317	15	11
1999	328	314	318	14	10
2000	327	315	319	12	8
1996-2000	1638	1561	1584	77	54
1976-2000	7962	6996	7252	966	710