Annual Reviews
www.annualreviews.org/aronline

Annu. Rev. Publ. Health 1991. 12:281–307
Copyright © 1991 by Annual Reviews Inc. All rights reserved

COMPETING RISKS IN MORTALITY ANALYSIS

Chin Long Chiang

School of Public Health, University of California, Berkeley, California 94720

KEY WORDS: potential lifetimes, crude and net probabilities, independence of risks, life table, multiple decrement table

INTRODUCTION

Every human is continuously exposed to many risks of death, such as cancer, heart disease, and tuberculosis. Because death is not a repetitive event and is usually attributed to a single cause, these risks compete with one another for the life of a person. Competing risks must be considered in any cause-specific mortality analysis. In a study of cancer as a risk of death, for example, some persons might die from other causes during the study period. These persons no longer could die from cancer, but neither would they survive to the end of the study period. What, then, would be the contribution of their survival experience to the study, and what adjustment would have to be made for the competing effect of other causes in the study of cancer? If cancer were eliminated as a risk of death, what would be a person's chance of surviving a given time period? How many years in life expectancy has one lost because cancer is risk of death? As another example, the AIDS epidemic is spreading over the world population. Does the presence the AIDS virus infection in a population increase the death rate from pneumonia, or from heart disease? Are AIDS patients more likely to die from cancer than persons without AIDS? A meaningful study of these questions requires the evaluation of AIDS as a competing risk.

The basic statistical quantities that measure the effect of a risk of death are survival probability, death probability, and life expectation. To evaluate a risk

0163-7525/91/0501-0281$02.00

Annual Reviews
www.annualreviews.org/aronline

of death, we would ideally have the risk in question operate alone in a population to determine the probabilities and the expectation of life. Alternatively, we would remove the risk in question from a population and evaluate the changes in the probabilities and the life expectation. Rarely is the ideal situation realized, but one can estimate these probabilities and the expectation of life by using the theory of competing risks in a cause-specific mortality study.

In this paper, we present a brief review of the concept of competing risks and the statistical methods of mortality analysis, including estimation of three types of probability of dying with respect to a particular cause of death. We will describe formulas of estimates for cohort studies, medical follow-up studies, and analyses of mortality data for a current population. To illustrate this method of analysis, we will use the major cardiovascular (CV) diseases and malignant neoplasms mortality data of the United States white male and female population in 1986.

BACKGROUND

Evolution of the Concept of Competing Risks

The concept of competing risks began in April 1760, when Daniel Bernoulli read his memoir on mortality due to smallpox and the advantage of inoculation for its prevention before the French Academy of Science (5). During the early eighteenth century, there were constant debates and discussions in England, France, and other European countries over the advantage of inoculation against smallpox, because deaths occurred among those who were inoculated. Data were collected and tables were prepared to show the results of some inoculation programs, without definitive conclusions regarding the advantage of inoculation. Karn (31) gave a detailed account of the events related to this controversy.

Bernoulli had proposed a mathematical approach to the problem. He wanted to compare the mean duration of life in two differently constituted populations: a real population who were subject to death from smallpox and from other causes versus another, hypothetical population for whom smallpox was not a cause of death. Assuming that during one year one in n persons acquires smallpox, and one in m persons who had smallpox dies, Bernoulli arrived at a formula for estimating the number of persons who will die from smallpox. He then used Edmond Halley's (27) life table of the city of Breslou to illustrate numerically the advantage of eliminating smallpox as a cause of death. Bernoulli set $n = 8$ and $m = 8$ and calculated that inoculation against smallpox would lengthen the average duration of life by about three years.

An important assumption in Bernoulli's solution to the smallpox controversy was, in present-day terminology, a constant incidence rate ($1/n$) and a constant case fatality rate ($1/m$) for smallpox. D'Alembert, Trembley, and

Annual Reviews
www.annualreviews.org/aronline

Laplace all had considered the problem when n and m both were functions of age. It was D'Alembert who was the most critical of Bernoulli's solution, and his criticism prompted Bernoulli to write an "Introduction apologétique" to preface his memoir. Although he, too, recognized the value of inoculation, D'Alembert felt that Bernoulli had overstated the epidemic and overestimated the advantage of inoculation. In response to the question, "Of all persons alive at a given epoch, what fractional part has not been attacked by the small pox?" D'Alembert estimated the number at one-fourth, whereas Bernoulli estimated two-thirteenths, which gave the estimate of the smallpox "prevalence rate" at 75% (= 1–1/4) by D'Alembert and 85% (= 1–2/13) by Bernoulli. D'Alembert also stressed the difference between the immediate danger of inoculation and the remote benefit in the additional years gained through inoculation. He also distinguished physical life from civil life, from which may have emerged the concept of quality of life. It was this exchange between D'Alembert and Bernoulli that brought out the notion of competing risks. Todhunter (49) gave mathematical details of their discussion.

Makeham (36) formulated the theory of multiple decrement forces and explored the practical applications. Actuarial mathematicians have applied Makeham's work to develop multiple decrement tables in the study of life contingencies. Spurgeon (48) described methods of analysis that involved two or more causes of decrement. In particular, Spurgeon included "withdrawal from observation" in his formula for the probability of dying during a year. The result was the now popular "actuarial method," which was promoted by Berkson & Gage (4). Bailey & Haycocks (2) discussed some theoretical aspects of multiple decrement life tables. Hooker & Longley-Cook (28) considered life and other contingencies. In the area of vital statistics, Greville (26) analyzed mortality tables by cause of death. The concept of competing risks also has applications in survival analysis (29, 40), reliability theory and life testing (3), and other fields.

Two papers on the problem of competing risks that have aroused much interest among researchers in public health were by Fix & Neyman (24) and by Cornfield (15). Fix & Neyman introduced a stochastic model to describe recovery, relapse, death, and loss of patients in medical follow-up studies. Cornfield described problems in the estimation of the probability of the development of a disease when there were competing risks. Chiang (9) considered causes of death as competing risks and formulated relations between three types of probability of death with respect to a specific cause as a basis for mortality analysis.

Recent Developments

Recent developments in the subject of competing risks have been based mainly on the concept of potential lifetimes (41). Suppose that a system with r components fails as soon as one of the components fails. For example, a room

Annual Reviews
www.annualreviews.org/aronline

that has r electrical lights connected in series becomes dark as soon as one of the bulbs burns out. Each component is subject to a failure risk R_i and has a potential or net lifetime X_i, for i = 1, . . ., r. The lifetime of the system, denoted by Y, is the smallest of $(X_1, \ldots, X_r)$, or

$$Y = \min (X_1, \ldots, X_r). \quad 1.$$

Generally it is assumed that the r potential lifetimes, denoted by an r-dimensional random vector,

$$\mathbf{X} = (X_1, \ldots, X_r), \quad 2.$$

has a joint distribution function

$$F_{\mathbf{X}}(\mathbf{x}) = F_{\mathbf{X}}(x_1, \ldots, x_r). \quad 3.$$

The marginal distribution of X_i,

$$F_{X_i}(x) = \Pr(X_i \leq x) = F_{\mathbf{X}}(\infty, \ldots, \infty, x, \infty, \ldots, \infty), \quad 4.$$

is the net probability of failure of the i-th component before time x. Its complement, $1\text{-}F_{X_i}(x)$, is the net survival probability of the i-th component to time x. The ratio,

$$\frac{dF_{X_i}(x)/dx}{1 - F_{X_i}(x)} = \mu(x;i), \quad 5.$$

is the failure rate (force of mortality) of the i-th component. Thus, the connection between the potential lifetimes and competing risks is clear. Although the marginal distribution (Eq. 4) can be derived from the joint distribution (Eq. 3), the converse is not necessarily true. When the random components $(X_1, \ldots, X_r)$ are mutually dependent, the joint distribution (Eq. 3) cannot be derived from the marginal distribution (Eq. 4), for i = 1, . . ., r. In David & Moeschberger (19) and in Birnbaum (6), discussion on competing risks was given in sections that dealt separately with dependent lifetimes and with independent lifetimes. Elandt-Johnson & Johnson (22) offered a review of the theory of competing risks.

Using potential lifetimes, we can study competing risks within the framework of multivariate analysis (see, for example, 18, 34, 39). Most articles have used exponential distributions (38). Moeschberger & Klein (42) discussed consequences of departure from independence of exponential series systems. Boardman & Kendall (7) developed maximum likelihood estimators when there are only two causes of failure. Gail (25a) used the joint survival

Annual Reviews
www.annualreviews.org/aronline

function to compare the actuarial method with other models. Birnbaum (6) devised a situation to illustrate the difference between the net and crude lifetimes of a system. In his discussion on the nonidentifiability of competing risks, Tsiatis (50) showed that, when potential lifetimes $(X_1, \ldots, X_r)$ are not known to be mutually independent, the crude probabilities are not of much use for identifying the joint distribution of $(X_1, \ldots, X_r)$ or the net distribution of each X_i. But, when risks are dependent, there is no simple statistical method available for the analysis of competing mortality risks in the human population. The competing risks problem is difficult indeed.

REMARK There is a major conceptual difficulty in using the potential lifetime theory to study competing risks in the human population. The difficulty is in the definition of sample space. Generally, the sample space of a random vector $\mathbf{X}$ in formula 2 is an r-dimensional space. For every sample point $(x_1, \ldots, x_r)$, for $x_i \leq 0$, in the r-dimensional space, there is a density function $f(x_1, \ldots, x_r)$ and a distribution function

$$F_X(x) = \int_0^{x_1} \cdots \int_0^{x_r} f(t_1, \ldots, t_r) dt_1 \ldots dt_r. \qquad 6.$$

But what does the sample point $(x_1, \ldots, x_r)$ represent in a competing risks analysis? According to the concept of multivariate distribution, it represents the event that an individual dies from r different causes at r different times, which is an impossible event! A human being can die only once from a single cause, and no one dies more than once and at different times. Consequently, the corresponding density function $f(x_1, \ldots, x_r)$ has no meaning. The sample space of the random vector $\mathbf{X} = (X_1, \ldots, X_r)$ is not an r-dimensional space, and neither is the domain of the distribution function $F_{\mathbf{X}}(\mathbf{x})$ in formula 3. It is unclear what the sample space of the random vector $\mathbf{X}$ should be and how the distribution function $F_{\mathbf{X}}(\mathbf{x})$ should be determined. Perhaps we need to reevaluate some of the theoretical results regarding competing mortality risks that are derived from the distribution function $F_{\mathbf{X}}(\mathbf{x})$. This discussion of the conceptual difficulty also applies to the joint distribution of potential failure times in survival analysis.

In the following sections, competing risks of death will be discussed without the benefit of the potential lifetime concept. The lifetime of an individual will be represented by a single random variable that has a univariate distribution. The sample space is the positive real line, and competing risks affect the lifetime through the force of mortality.

Independence Assumption of Competing Risks

Competing risks of death are independent of one another if the force of mortality of each risk remains constant after one or more risks are eliminated

Annual Reviews
www.annualreviews.org/aronline

or altered. Because there is no simple statistical method available for cause-specific mortality analysis when risks are dependent, independence of risks is generally assumed. But some researchers have questioned the validity of the assumption (see, for example, Ref. 47), which has become the focal point in the discussion of analysis methods. Perhaps there is no unique answer to the question of risk independence. The answer probably depends on the risks involved and, possibly, on the population under study. The independence assumption may not hold among closely related causes of death, but it may be true between distant disease categories. A direct approach to the problem is to either physically remove the specific risk from the human population or introduce a new risk of death, and check the change in mortality from other causes. This seemingly drastic proposal is not always unrealistic, as we have seen in two events of the recent past. The first event occurred in 1955, when the Salk vaccine and subsequently the Sabin vaccine drastically reduced the incidence of poliomyelitis in the United States and elsewhere in the world. A thorough analysis of mortality data in the United States before and after the vaccine should help to determine the effect of poliomyelitis on other causes of death operating in the population, particularly among the very young.

The second event was the AIDS epidemic in 1981, which was a completely new risk. The epidemic started rather suddenly and spread swiftly in the human population. Tens of thousands of persons have died of AIDS and millions of others are thought to be infected with the human immunodeficiency virus. This disease and the changes in mortality also provide us with an opportunity to verify the independence assumption, at least between major disease categories. Does the appearance of AIDS affect the force of mortality of other risks of death, such as cancer?

The National Cancer Institute (NCI) has published data that may help to determine if cancer is independent of AIDS. Table 1 was reproduced from *Cancer Statistics Reviews,* 1973–1986, published by the NCI (Ref. 43, especially Table IV-5). This table summarizes 14-year trends of cancer mortality from 1973 to 1986 in the United States among white males and females. For our purpose, the years 1973–1974 represent a period *before* the AIDS epidemic, and the years 1985–1986 represent a period *after* the outbreak. In addition to the age-specific cancer death rates during the *before* and *after* years, Table 1 contains percentage changes and the estimated annual percent changes from *before* to *after* for each age group among males and females. As most AIDS victims were young males and very few were females in 1985–1986, the white males may be considered "cases," and the white females, "controls." The changes in cancer mortality from *before* to *after* among males (cases) can be compared with the changes among females (controls) for each age group. If the changes in cancer mortality from *before*

Annual Reviews
www.annualreviews.org/aronline

Table 1 Summary of 14-year trends. Age specific cancer death rates by sex and age, US white population, 1973–1986[a]

Sex/Age	Average rate 1973–74	Average rate 85–86	Percent change	EAPC[b]
Males	202.7	212.6	4.9	0.4
0–54	37.3	33.1	−11.2	−1.0
0–14	6.2	4.0	−35.8	−3.4
15–34	11.3	9.2	−18.5	−1.8
35–44	47.5	40.7	−14.4	−1.3
45–54	172.3	160.4	−6.9	−0.6
55–64	492.5	494.9	0.5	0.1
65+	1290.0	1423.0	10.3	0.8
65–74	1018.0	1085.0	6.5	0.5
75+	1734.0	1975.0	13.9	1.1
Females	130.0	138.2	6.3	0.5
0–54	37.3	33.2	−11.1	−1.0
0–14	4.8	3.2	−33.7	−3.1
15–34	9.4	7.5	−20.3	−1.8
35–44	58.7	49.5	−15.6	−1.4
45–54	170.1	158.7	−6.7	−0.7
55–64	345.1	363.6	5.4	0.5
65+	690.8	790.3	14.4	1.2
65–74	553.8	651.1	17.6	1.4
75+	914.2	1017.0	11.3	0.9

[a]All sites combined. Rates per 100,000 and age-adjusted to the 1970 US standard population.
[b]EAPC: Estimated Annual Percent Change over the 14-year interval.
Source: Natl. Cancer Inst. May 1989. *Cancer Statistics Review*, Section IV, Table IV-5

to *after* among males are quite different from the changes among females, then cancer may be dependent on AIDS. If the changes in cancer mortality among males are similar to those among females, then cancer probably is independent of AIDS.

In the four age groups less than 55 years of age, the percentage changes in cancer death rates from *before* to *after* were very close: (−35.8, −18.5, −14.4, −6.9) for white males and (−33.7, −20.3, −15.6, −6.7) for white females. Thus, the NCI cancer mortality trends analysis seems to suggest that cancer is independent of AIDS. Although more data and statistical analysis are needed to establish, or to repudiate, the independence assumption, we use the assumption to proceed with our discussion.

Annual Reviews
www.annualreviews.org/aronline

COMPETING RISKS

Three Types of Probability

The concept of competing risks has been expressed in terms of probability of dying (9). In a mortality analysis without specification of cause of death, the meaning of the probability of dying (in a time interval) is clear. When competing risks are considered, the probability of dying is subject to various interpretations. Each interpretation leads to a different probability, and each probability serves a different purpose. One can select a particular type of probability to suit the needs of a mortality study. To understand the concept of competing risks, one needs to understand various types of probability. For a person alive at the exact age x_i, three types of probability are possible:

THE CRUDE PROBABILITY: The probability of dying from a specific cause in the presence of all other competing risks. In reference to age interval (x_i, x_{i+1}), the probability is:

$Q_{i\delta}$ = Pr (of dying in the interval (x_i, x_{i+1}) from cause R_δ in the presence of all other risks in the population).

THE NET PROBABILITY: The probability of dying if a specific risk is the only risk in effect in the population or, conversely, the probability of dying if a specific risk is eliminated from the population. For age interval (x_i, x_{i+1}), the probabilities are:

$q_{i\delta}$ = Pr (of dying in the interval (x_i, x_{i+1}) if risk R_δ is the only risk in effect in the population);
$q_{i.\delta}$ = Pr (of dying in the interval (x_i, x_{i+1}) if risk R_δ is eliminated as a risk of death).

THE PARTIAL CRUDE PROBABILITY: The probability of dying from a specific cause when another risk (or risks) is eliminated as a risk of death from the population. Or

$Q_{i\delta.1}$ = Pr (of dying in the interval (x_i, x_{i+1}) from R_δ when risk R_1 is eliminated as a risk of death);

and

$Q_{i\delta.12}$ = Pr (of dying in the interval (x_i, x_{i+1}) from R_δ when R_1 and R_2 are eliminated as risks of death).

When cause of death is not specified, the probabilities are:

q_i = Pr (a person alive at age x_i will die in the interval (x_i, x_{i+1}))

Annual Reviews
www.annualreviews.org/aronline

and

p_i = Pr (a person alive at age x_i will survive to the end of the interval (x_i, x_{i+1})),

with $q_i + p_i = 1$.

For example, if R_1 represents the risk of dying from cancer and the age interval is (40, 45), then the crude probability Q_{i1} is the probability that a person 40 years of age will die from cancer before reaching age 45. The net probability q_{i1} is the probability of the person dying in the interval (40, 45) if cancer were the only cause of death operating in a population, and $q_{i.1}$ is the probability that the person will die in interval (40, 45) if cancer were eliminated as a risk of death. If R_2 represents the risk of death from heart disease, then the partial crude probability $Q_{i2.1}$ is the probability of dying in the age interval (40, 45) from heart disease, if cancer were eliminated as a risk of death.

The probabilities p_i, q_i, and $Q_{i\delta}$ are real and can be estimated directly from a cause-specific mortality analysis. The net probabilities $q_{i\delta}$ and $q_{i.\delta}$ and the partial crude probabilities $Q_{i\delta.1}$ and $Q_{i\delta.12}$ are probabilities in a hypothetical situation. They cannot be estimated directly, but only through their relations with p_i, q_i and $Q_{i\delta}$. Generally, the net probability of dying $q_{i.\delta}$ and the partial crude probability $Q_{i\delta.1}$ are of particular interest in a mortality analysis, and we will use them in the following section.

The terms "risk" and "cause" need clarification, as both may refer to the same condition, but are distinguished by their position in time relative to the occurrence of death. Before death, a condition is a risk; after death the same condition is a cause. For example, cancer is risk of death to which a person is exposed, but cancer also is the cause of death if a person eventually dies from it.

Relations Between Crude, Net, and Partial Crude Probabilities

Suppose that r risks of death are acting simultaneously on each person in a population, and let these risks be denoted by $R_1, \ldots, R_r$. For each risk, R_δ, there is a corresponding force of mortality $\mu(t;\delta)$ such that

$$\mu(t;\delta)dt = \text{Pr (a person alive at time t will die in time element } (t,\ t+dt) \text{ from risk } R_\delta), \qquad 7.$$

for $\delta = 1, \ldots, r$. The sum

$$\mu(t;1) + \ldots + \mu(t;r) = \mu(t) \qquad 8.$$

is the total force of mortality so that,

Annual Reviews
www.annualreviews.org/aronline

$$\mu(t)dt = \Pr \text{ (a person alive at time t will die in time element } (t,\ t+dt)). \qquad 9.$$

The probability of dying q_i is a function of the force of mortality $\mu(t)$:

$$q_i = 1 - \exp\left\{ - \int_{x_1}^{x1+1} \mu(t)dt \right\}, \qquad 10.$$

where the limits of the integral are the limits of the interval (x_i, x_{i+1}). When the force of mortality $\mu(t) = \mu$ is constant in the interval, the formula of the probability q_i reduces to

$$q_i = 1 - e^{-n_i \mu}, \qquad 11.$$

where $n_i = x_{i+1} - x_i$ is the length of the interval.

PROPORTIONALITY ASSUMPTION The theory of competing risks requires two assumptions: the above-mentioned independence assumption and the proportionality assumption described below. For each risk R_δ, the force of mortality $\mu(t;\delta)$ is a function of time t and of risk R_δ. Under the proportionality assumption, within the time interval (x_i, x_{i+1}) and ratio of $\mu(t;\delta)$ to the total force of mortality $\mu(t)$,

$$\frac{\mu(t;\delta)}{\mu(t)} = c_{i\delta}, \qquad 12.$$

is independent of t, but is a function of the interval (x_i, x_{i+1}) and of risk R_δ. This assumption permits the risk-specific force of mortality $\mu(t;\delta)$ to vary in absolute magnitude, but requires that it remain a constant proportion of the total force of mortality in the interval (x_i, x_{i+1}). David (17) has shown that the proportionality assumption in formula 12 can be satisfied whenever the underlying distribution of lifetime has one of three possible forms of the extreme-value distribution of the minimum. Thus, the assumption also is satisfied in the exponential and Weibull distributions.

Formula 12 can be extended immediately to the probability of dying. When the ratio of the risk-specific force of mortality to the total force of mortality is constant throughout a time interval, this constant must be equal to the ratio of the corresponding probabilities of dying over the entire interval. That is,

$$\frac{\mu(t;\delta)}{\mu(t)} = \frac{Q_{i\delta}}{q_i}, \qquad 13.$$

Annual Reviews
www.annualreviews.org/aronline

and hence

$$Q_{i\delta} = \frac{\mu(t;\delta)}{\mu(t)} q_i, \qquad 14.$$

for $\delta = 1, \ldots, r$. Thus, the (crude) probability of dying in an interval from risk R_δ is the proportion $\mu(t;\delta)/\mu(t)$ of the (total) probability of dying in the interval, q_i. The larger this proportion is, the greater is the probability of dying from the corresponding risk R_δ.

Taking the summation of both sides in equation 14, for $\delta = 1, \ldots, r$, yields the equation

$$Q_{i1} + \ldots + Q_{ir} = q_i \,. \qquad 15.$$

The sum on the left hand side of the equality in formula 15 is the probability of dying from one of the risks $(R_1, \ldots, R_r)$, and hence is equal to the probability of dying, q_i.

Using the proportionality assumption in formula 12, we find formulas that express the net probability $q_{i.\delta}$ and the partial crude probability $Q_{i\delta.1}$ in terms of the probabilities p_i, q_i and $Q_{i\delta}$. For example, when R_1 is eliminated as a risk of death, the net probability of dying in age interval (x_i, x_{i+1}) is

$$q_{i.1} = 1 - p_i^{(q_i - Q_{i1})/q_i}, \qquad 16.$$

and the partial crude probability of dying from R_δ is

$$Q_{i\delta.1} = \frac{Q_{i\delta}}{q_i - Q_{i1}} [1 - p_i^{(q_i - Q_{i1})/q_i}], \qquad 17.$$

for $\delta = 1, \ldots, r$. Formulas 16 and 17 are basic for estimating the net and the partial crude probabilities in practical applications of the theory of competing risks.

SOME OBSERVATIONS Table 2 represents a hypothetical situation in which an individual is exposed to $r = 3$ risks of death (R_1, R_2, R_3) in two time intervals. The forces of mortality of risks R_1 and R_3 are constant with $\mu(t;1) = .10$ and $\mu(t;3) = .30$, respectively, in both intervals, but the force of mortality of R_2 changes, from $\mu(t;2) = .20$ in the first interval to $\mu(t;2) = .25$ in the second. These values and the total force of mortality $\mu(t)$ are recorded in columns 2 through 5. The probabilities of dying q_i, $Q_{i\delta}$, $q_{i.\delta}$, and $Q_{i\delta.1}$, computed from formulas 11, 14, 16, and 17 are shown in columns 6 through 14. The following points deserve some attention when studying competing risks of death, as illustrated with the numerical example in Table 2.

1. A risk that has a low force of mortality has a small crude probability of

Annual Reviews
www.annualreviews.org/aronline

Table 2 Force of mortality and probability of dying when three risks acting in a population

Time	Force of mortality				Crude probability				Net probability			Partial crude probability	
interval	R_1	R_2	R_3	total	R_1	R_2	R_3	total	R_1	R_2	R_3	R_2	R_3
i	$\mu(t;1)$	$\mu(t;2)$	$\mu(t;3)$	$\mu(t)$	Q_{i1}	Q_{i2}	Q_{i3}	q_i	$q_{i.1}$	$q_{i.2}$	$q_{i.3}$	$Q_{i2.1}$	$Q_{i3.1}$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1	.10	.20	.30	.60	.075	.150	.226	.451	.393	.330	.259	.157	.236
2	.10	.25	.30	.65	.073	.184	.221	.478	.424	.330	.295	.193	.231

Annual Reviews
www.annualreviews.org/aronline

dying. In Table 2, the force of mortality of R_1, R_2, and R_3 are in the order of magnitude: $\mu(t;1)<\mu(t;2)<\mu(t;3)$. The corresponding crude probabilities of dying are in the same order, $Q_{i1}<Q_{i2}<Q_{i3}$, in both time intervals. More precisely, the proportion of the probability of dying, q_i, attributable to a risk R_δ is equal to the proportion of the corresponding forces of mortality, $Q_{i\delta}/q_i = \mu(t;\delta)/\mu(t)$, as shown in formula 13.

In the first time interval, $Q_{i1}/q_i = \mu(t;1)/\mu(t) = .10/.60$ for risk R_1, $Q_{i2}/q_i = \mu(t;2)/\mu(t) = .20/.60$ for risk R_2, and $Q_{i3}/q_i = \mu(t;3)/\mu(t) = .30/.60$ for risk R_3.

2. Elimination of a risk that has a low force of mortality will cause a small reduction in the probability of dying. Therefore, when risk R_δ is eliminated as a risk of death, the net probability of dying $q_{i.\delta}$ has a reverse order of magnitude as that of the force of mortality.

In the first interval, the forces of mortality are: $[\mu(t;1)<\mu(t;2)<\mu(t;3)] = [.10<.20<.30]$, whereas the net probabilities are: $[q_{i.1}>q_{i.2}>q_{i.3}] = [.393>.330>.259]$.

3. When R_1 is eliminated as a risk of death, the net probability of dying $q_{i.1}$ equals the sum of the partial crude probabilities: $q_{i.1} = Q_{i2.1}+Q_{i3.1}$, because when R_1 is eliminated, an individual either dies from R_2 with a probability $Q_{i2.1}$ or dies from R_3 with a probability $Q_{i3.1}$. Therefore their sum equals $q_{i.1}$.

From columns 10, 13 and 14, we find the equality $.393 = .157+.236$ in the first interval and the equality $.424 = .193+.231$ in the second.

4. Although survival and death of an individual are determined by the force of mortality $\mu(t)$, the chance of dying from a specific cause is influenced by competing risks. For example, the crude probability of dying from risk R_1, Q_{i1}, is a function of the force of mortality $\mu(t;1)$, as well as the forces of mortality of R_2 and R_3, $\mu(t;2)$ and $\mu(t;3)$. The probability Q_{i1} decreases as the sum $\mu(t;2) + \mu(t;3)$ increases, even when $\mu(t;1)$ remains unchanged.

In Table 2, the crude probability of dying from R_1 decreases from $Q_{i1} = .075$ in the first interval to $Q_{i1} = .073$ in the second when $\mu(t;2)$ increases from $\mu(t;2) = .20$ to $\mu(t;2) = .25$, even though $\mu(t;1) = .10$ in both intervals.

5. Independence of competing risks is judged by the force of mortality, not by the (crude) probability of dying. In this example, risk R_1 is independent of risk R_2, because the force of mortality of R_1 $\mu(t;1)$ remains constant when $\mu(t;2)$ changes in the two intervals; however, the crude probability of dying from R_1, Q_{i1}, changes with $\mu(t;2)$. Similarly, risk R_3 also is independent of risk R_2, as $\mu(t;3) = .30$ in both intervals, although Q_{i3} changes with $\mu(t;2)$.

6. When R_1 is eliminated, only R_2 and R_3 remain as competing risks. The (partial crude) probability of dying from R_3 is affected by the magnitude of the force of mortality of R_2 $\mu(t;2)$.

When $\mu(t;2)$ increases from .20 to .25, the probability of dying from R_3

Annual Reviews
www.annualreviews.org/aronline

decreases from $Q_{i3.1} = .236$ in the first interval to $Q_{i3.1} = .231$ in the second.

The example in Table 2 was taken, with changes, from a table in Kimball (32). Kimball suggested a conditional probability of dying from a risk, say R_2, given not dying from another risk, R_1, or $Q_{i2}/(1-Q_{i1})$, as a substitute for the partial crude probability $Q_{i2.1}$. These two probabilities, however, are different in concept. Kimball's article has caused much discussion from Mantel & Bailer (37), Pike (46), and Chiang (13). Another substitute for the partial crude probability $Q_{i2.1}$ was proposed by Wong (51), who uses multiple causes of death information to estimate the additional number of deaths from R_2 if risk R_1 is eliminated as a risk of death.

Estimation of Probabilities

Chiang (9, 12) and David & Moeschberger (19) have reported methods of analysis and statistical inference in competing risks studies. This section briefly describes probability estimates in three types of studies: cohort, medical follow-up, and current population mortality analysis.

COHORT STUDIES Let a cohort of l_0 newborn infants be observed from birth until the death of the last member of the cohort. For age interval (x_i, x_{i+1}), let l_i be the number of persons (out of l_0) alive at x_i, l_{i+1} who survive to age x_{i+1}, and $d_{i\delta}$ die from cause R_δ, for $\delta = 1, \ldots, r$, so that

$$d_{i1} + \ldots + d_{ir} + l_{i+1} = l_i \,. \qquad 18.$$

Each of the l_i individuals is subject to the probability $Q_{i\delta}$ of dying from R_δ in (x_i, x_{i+1}) and p_i of surviving to x_{i+1}, with

$$Q_{i1} + \ldots + Q_{ir} + p_i = 1. \qquad 19.$$

Estimates of the probabilities in formula 19 are the corresponding proportions in formula 18. Namely,

$$\hat{Q}_{i\delta} = d_{i\delta}/l_i, \; \hat{q}_i = d_i/l_i, \text{ and } \hat{p}_i = l_{i+1}/l_i, \qquad 20.$$

for $\delta = 1, \ldots, r$; where $d_i = d_{i1} + \ldots + d_{ir}$ is the total number of deaths in the interval (x_i, x_{i+1}).

Substituting the estimates $\hat{Q}_{i\delta}$, $\hat{q}_i$ and $\hat{p}_i$ in formulas 16 and 17 yields the estimates of the net and the partial crude probabilities:

$$\hat{q}_{i.\delta} = 1 - [l_{i+1}/l_i]^{(d_i - d_{i\delta})/d_i}, \qquad 21.$$

and

Annual Reviews
www.annualreviews.org/aronline

$$\hat{Q}_{i\delta.1} = \frac{d_{i\delta}}{d_i - d_{i1}}\{1 - [l_{i+1}/l_i]^{(d_i - d_{i1})/d_i}\}. \qquad 22.$$

Birnbaum (6) proved that, under the proportionality assumption 12, $\hat{q}_{i.\delta}$ in formula 21 is a consistent estimate of the probability $q_{i.\delta}$. Using Birnbaum's approach, we can show that $\hat{Q}_{i\delta.1}$ in formula 22 is a consistent estimate of the probability $Q_{i\delta.1}$.

MEDICAL FOLLOW-UP STUDIES Consider a medical follow-up study conducted over a period of y years. A total of N_0 patients are admitted to the study at various times during the study period and observed until either their deaths or the end of the observation period (such as termination of the study), whichever comes first. The time of admission is taken as the common point of origin for all N_0 patients. For a given patient, time zero is the date of admission. Thus, if Patient A is admitted to the study on January 1, 1978, and Patient B is admitted on July 1, 1981, their points of origin are January 1, 1978 and July 1, 1981, respectively. The first anniversary of follow-up is January 1, 1979, for Patient A and July 1, 1982, for Patient B. It is customary in medical follow-up studies to use the anniversary year (the number of years since admission) as the time scale. The typical interval will be denoted by $(x, x+1)$, for $x = 0, 1, \ldots, y-1$, so that x is the exact number of years of follow-up. The symbol p_x will denote the probability that a patient alive at time x will survive to the end of the interval $(x, x+1)$; q_x, the probability of dying during the interval; and $Q_{x\delta}$, the probability of dying during the interval from cause R_δ, with $Q_{x1} + \ldots + Q_{xr} = q_x$, and $p_x + q_x = 1$. At time x, there are N_x patients alive and to be observed over the interval $(x, x+1)$. Of these patients, S_x will survive to time $x+1$ to become N_{x+1}; $D_{x\delta}$ will die from R_δ in $(x, x+1)$, for $\delta = 1, \ldots, r$. Finally, the sum $D_{x1} + \ldots + D_{xr} = D_x$ is the total number of deaths in $(x, x+1)$.

Up to this point, the follow-up study is similar to the cohort study. In a follow-up study, however, there are two categories of patients for whom the survival and mortality information will be incomplete. First, there will be patients who are admitted to the study between x and $x+1$ years before termination of study. These patients cannot be observed for the entire interval $(x, x+1)$. They are subject to withdrawal from the study during the interval. Second, there will be patients who are lost to the study because of follow-up failure in the interval $(x, x+1)$. Survival or death of these patients will be unknown to the researchers. These two groups of patients are different from a statistical viewpoint, simply because every one of the N_x patients is subject to the risk of getting lost, but only those who are admitted between x and $x+1$ years before termination of study are subject to withdrawal in $(x, x+1)$. Loss to follow-up can be treated as a competing risk, whereas withdrawal should

Annual Reviews
www.annualreviews.org/aronline

not be. However, when the end of observation of each patient is known, the distinction between the two groups has little effect on the estimates of the probabilities.

The two sources of incomplete information have created interesting statistical problems. Many have contributed to the method of analysis of follow-up data. Spurgeon (48) proposed formulas to deal with withdrawals; Frost (25) introduced the concept of "person years." Others include Fix & Neyman (24), Armitage (1), Dorn (20), and Littel (35). Berkson & Gage (4) and Culter & Ederer (16) promoted the actuarial method to compute estimates of the probability q_x. Kaplan & Meier (30) introduced a nonparametric formula, and Elvebeck (23), Chiang (10), and Drolette (21) each proposed formulas for estimating q_x. Kuzma (33) provided a review of some of these methods (see also 25a). Others have extended the follow-up concept in survival analysis and introduced several types of censorship (see, for example, 29, 40).

With the current easy access to computer facilities, one should collect more information so that statistical formulas will be simple in concept and require fewer assumptions. The most useful information in a follow-up study is the time of each death, the time of every withdrawal, and the time when a patient is lost for each lost case. With such information in mind, we can proceed to derive estimates of the probabilities $p_x, Q_{x1}, \ldots, Q_{xr}$. For convenience, being lost is considered as a competing risk denoted by R_0 with the "force" $\mu(t;0)$.

For time interval $(x,x+1)$, let

N_x = number of patients alive at time x;
S_x = number of patients who survive to $x+1$;
W_x = number of withdrawals;
τ_i = the time of i-th withdrawal, $i = 1, \ldots, W_x$;
D_{x0} = number of lost patients;
t_{0j} = the time at which j-th patients is lost, $j = 1, \ldots, D_{x0}$;
$D_{x\delta}$ = number died from R_δ, $\delta=1, \ldots, r$;
$t_{\delta j}$ = the time of j-th death from R_δ, $j = 1, \ldots, D_{x\delta}$.

The total length of time that the N_x patients are under observation in the interval $(x, x + 1)$ is

$$T_x = S_x + \sum_{i=1}^{W_x} \tau_i + \sum_{\delta=0}^{r} \sum_{j=1}^{D_{x\delta}} t_{\delta j}. \qquad 23.$$

The estimates of p_x, q_x, and $Q_{x\delta}$ are function of the number of deaths and the total length of observation T_x (12). Namely,

$$\hat{p}_x = \exp\{-D_x/T_x\}, \quad \hat{Q}_{x\delta} = \frac{D_{x\delta}}{D_x}[1 - \exp\{-D_x/T_x\}], \qquad 24.$$

Annual Reviews
www.annualreviews.org/aronline

and $\hat{q}_x = 1-\hat{p}_x$. Using formulas 16 and 17 once again, we can find the estimates of the net and the partial crude probabilities. For example, the estimate of the net probability $q_{x.\delta}$ is

$$\hat{q}_{x.\delta} = 1 - \exp\{-(D_x - D_{x1})/T_x\} . \qquad 25.$$

CURRENT POPULATION MORTALITY ANALYSES Mortality data of a current population, such as the United States 1989 population, are of the form of age-specific and age-cause-specific death rates. The National Center for Health Statistics publishes annual vital statistics that contains tables of age-specific death rate M_i and age-cause-specific death rate $M_{i\delta}$, for each cause R_δ and for each interval (x_i, x_{i+1}), by race and sex for the US population and for many geographical areas in the country. The rates also can be computed from $M_i = D_i/P_i$ and $M_{i\delta} = D_{i\delta}/P_i$, for $\delta = 1, \ldots, r$, and $i = 0,1, \ldots, w$. Here, $D_{i\delta}$ is the number of deaths from cause R_δ, $D_i = D_{i1} + \ldots + D_{ir}$ is the total number of deaths, and P_i is the midyear population for age interval (x_i, x_{i+1}) during the current year. The midyear population P_i can be found in the Bureau of the Census publications (8).

Tables 3 and 4 show age-specific death rates for all causes (M_i), for malignant neoplasms (ICD# 140–209) (M_{i1}) and for major cardiovascular diseases (ICD# 390–448) (M_{i2}), for white males and white females in the United States in 1986. The last age interval is an open interval, x_w and above. In this case, $x_w = 85$ years.

These rates are used to derive estimates of the probabilities q_i and $Q_{i\delta}$ by means of formulas of conversion. Several conversion formulas from death rate M_i to the probability q_i (known as methods of life table construction) have appeared during the development of the life table. King (32a) used a graduation process to derive q_i from M_i. Reed & Merrell (47a) proposed an exponential function of M_i for $1\text{-}q_i$. Greville (25b) used Euler-Maclaurin summation formula to obtain a formula for q_i. The formulas proposed by Chiang (11), Sirken (47b), and Keyfitz (31a) all use the concept of the fraction of the last age interval of life a_i. When a person dies in an age interval (x_i, x_{i+1}), he or she has lived a fraction of the interval before death. This fraction varies from one person to another; the mean (expected) value is the fraction a_i. Generally, the fraction a_i is invariant with respect to cause of death and is subject to little variation over time. For a discussion of the fraction see Chiang (12, p. 204). The average length of time lived in the interval (x_i, x_{i+1}) by those who die during the interval is $a_i n_i$, where $n_i = x_{i+1} - x_i$ is the length of the interval. We discuss one of the formulas below.

This formula of converting the age-specific death rate M_i to the corresponding age-specific probability of dying q_i is based on the following definition of the death rate M_i:

Annual Reviews
www.annualreviews.org/aronline

$$M_i = \frac{\text{expected number of deaths occurring in } (x_i, x_{i+1})}{\text{expected length of exposure to the risk of dying in } (x_i, x_{i+1})} . \quad 26.$$

The definition in formula 26 is independent of the number of persons involved. For a person alive at age x_i, the number of deaths is either one or zero. If the person dies in the interval (with a probability q_i), the number of deaths is one. If the person survives the interval (with a probability $1-q_i$), the number of deaths is zero. Therefore, the expected number of deaths is q_i, which is the numerator in formula 26. For the denominator, we realize that the person is exposed to the risk of dying in the entire interval (x_i, x_{i+1}). But this exposure to death ends as soon as death occurs. If the person dies in the interval (with a

Table 3 Population, death rate from all causes, from malignant neoplasms and from cardiovascular diseases by age group, US white males, 1986

Age interval (in years)	Midyear population[a] (in 1000s)	Death rate per 100,000[b]		
		All causes	Malignant neoplasms (140–208)	Cardio-vascular diseases (390–448)
x_i to x_{i+1}	P_i	M_i	M_{i1}	M_{i2}
0– 1	1,565	976.6	3.0	29.0
1– 5	5,973	52.2	4.7	2.4
5–10	7,171	25.3	4.1	1.0
10–15	6,849	34.7	3.6	1.3
15–20	7,757	124.2	5.5	3.1
20–25	8,532	165.6	7.9	4.6
25–30	9,347	157.6	10.8	7.5
30–35	8,846	180.6	16.4	16.3
35–40	8,028	212.3	25.9	38.0
40–45	6,144	295.6	53.0	85.8
45–50	5,060	452.0	107.7	162.9
50–55	4,603	746.3	214.4	306.0
55–60	4,742	1,221.5	390.8	521.7
60–65	4,548	1,939.8	622.6	864.4
65–70	3,928	2,908.7	900.9	1,338.2
70–75	2,948	4,602.1	1,279.6	2,200.9
75–80	1,982	6,988.1	1,661.6	3,485.7
80–85	1,080	10,825.7	2,130.6	5,723.5
85+	706	18,576.1	2,462.3	10,555.7

[a] Bur. of the Census. 1988. *Current Population Reports*, Ser. P-25, No. 1022
[b] Natl. Cent. for Health Stat. 1986. *Vital Statistics of the US*, Vol. II, Part A

Annual Reviews
www.annualreviews.org/aronline

Table 4 Population, death rate from all causes, from malignant neoplasms and from cardiovascular diseases by age group, US white females, 1986

Age interval (in years)	Midyear population[a] (in 1000s)	Death rate per 100,000[b]		
		All causes	Malignant neoplasms (140–208)	Cardio-vascular diseases (390–448)
x_i to x_{i+1}	P_i	M_i	M_{i1}	M_{i2}
0– 1	1,486	759.1	2.4	21.5
1– 5	5,674	40.7	3.4	2.3
5–10	6,803	17.4	3.2	0.9
10–15	6,493	19.9	3.0	1.1
15–20	7,448	49.1	3.8	1.7
20–25	8,413	51.6	4.6	2.9
25–30	9,150	54.1	8.2	4.8
30–35	8,702	67.0	16.2	7.3
35–40	8,031	94.2	33.0	12.8
40–45	6,266	156.0	65.8	28.3
45–50	5,213	250.1	116.7	54.7
50–55	4,826	416.9	197.7	103.5
55–60	5,161	650.4	299.7	187.0
60–65	5,190	1,055.0	438.7	357.0
65–70	4,707	1,608.2	580.0	630.3
70–75	3,950	2,536.7	752.4	1,155.8
75–80	3,111	3,995.0	879.0	2,101.2
80–85	2,055	6,794.5	1,073.5	4,002.8
85+	1,825	14,502.9	1,283.6	9,509.7

[a] Bur. of the Census. 1988. *Current Population Reports*, Ser. P-25, No. 1022
[b] Natl. Cent. for Health Stat. 1986. *Vital Statistics of the US*, Vol. II, Part A

probability q_i), the length of exposure is $a_i n_i$; if the person survives the interval (with a probability $1-q_i$), the length of exposure is the entire length of the interval n_i. Therefore the expected length of exposure is $q_i a_i n_i + (1-q_i)n_i$, and the analytic expression of the definition in formula 26 is

$$M_i = \frac{q_i}{q_i a_i n_i + (1 - q_i)n_i}. \qquad 27.$$

Solving equation 27 for q_i yields the desired formula for the probability q_i:

$$q_i = \frac{n_i M_i}{1 + (1 - a_i)n_i M_i}, \qquad 28.$$

Annual Reviews
www.annualreviews.org/aronline

which was given in Chiang (11). For a theoretical derivation see Chiang (14) and Elandt-Johnson & Johnson (22).

The age-cause-specific death rate $M_{i\delta}$ from cause R_δ is defined in a similar manner as M_i. For a person alive at age x_i, the death rate $M_{i\delta}$ is defined as:

$$M_{i\delta} = \frac{\text{expected number of deaths from cause } R_\delta \text{ in } (x_i, x_{i+1})}{\text{expected length of exposure to the risk of dying from cause } R_\delta \text{ in } (x_i, x_{i+1})}. \quad 29.$$

The corresponding analytic formula is

$$M_{i\delta} \quad \frac{Q_{i\delta}}{q_i a_i n_i + (1 - q_i) n_i}. \quad 30.$$

The probability $Q_{i\delta}$ in the numerator is the expected number of deaths from R_δ in (x_i, x_{i+1}). The denominator is the expected length of exposure to the risk of dying from R_δ, which of course is the same as the denominator in formula 27. Equations 27 and 30 imply that the estimate of $Q_{i\delta}$ is given by

$$\hat{Q}_{1\delta} = \frac{M_{i\delta}}{M_i} \hat{q}_i. \quad 31.$$

Note that formula 31 is a logical extension of formula 14, as the ratio of two death rates is equal to the ratio of the corresponding forces of mortality.

In summary, formulas 28 and 31 are used to estimate the probabilities q_i and $Q_{i\delta}$, for each age interval. Substituting these formulas in 16 and 17 gives formulas for estimating the net probability $q_{i.\delta}$ and the partial crude probability $Q_{i\delta.1}$. For example, estimate of $q_{i.\delta}$ is

$$\hat{q}_{i.\delta} = 1 - \hat{p}_i^{(M_i - M_{i\delta})/M_i}. \quad 32.$$

where $\hat{p}_i = 1 - \hat{q}_i$.

An Application to Current Mortality Analysis

We have chosen the net probability of dying ($q_{i.\delta}$) as an example, and use the current mortality data from malignant neoplasms (ICD # 140–209) and major CV diseases (ICD# 390–448) of US white male and female populations in 1986 for illustration. Cardiovascular diseases and malignant neoplasms have been the major causes of death in the US for many years. These diseases accounted for nearly 70% of all deaths in the entire white population in 1986, including 75% of deaths among persons aged 55 or older. Cardiovascular diseases alone were responsible for 50% of all deaths among persons older than age 75. Tables 5, 6, and 7 show the impact of these diseases on the probability of dying and the expectation of life in numerical figures.

Annual Reviews
www.annualreviews.org/aronline

Table 5 Multiple decrement: Probability of dying, q_i, and crude probabilities of dying, Q_{i1}, Q_{i2}, and Q_{i3}, United States white male population and white female population, 1986[a]

Age Interval (in years)	White Males				White Females			
	All causes	Neoplasms neoplasms	Major CV diseases	Other causes	All causes	Neoplasms neoplasms	Major CV diseases	Other causes
x_i to x_{i+1}	q_i	Q_{i1}	Q_{i2}	Q_{i3}	q_i	Q_{i1}	Q_{i2}	Q_{i3}
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
00–01	0.0096	0.0000	0.0003	0.0093	0.0075	0.0000	0.0002	0.0073
01–05	0.0021	0.0002	0.0001	0.0018	0.0016	0.0001	0.0001	0.0014
05–10	0.0013	0.0002	0.0000	0.0010	0.0009	0.0002	0.0000	0.0007
10–15	0.0017	0.0002	0.0001	0.0015	0.0010	0.0001	0.0001	0.0008
15–20	0.0062	0.0003	0.0002	0.0058	0.0025	0.0002	0.0001	0.0022
20–25	0.0082	0.0004	0.0002	0.0076	0.0026	0.0002	0.0001	0.0022
25–30	0.0078	0.0005	0.0004	0.0069	0.0027	0.0004	0.0002	0.0021
30–35	0.0090	0.0008	0.0008	0.0074	0.0033	0.0008	0.0004	0.0022
35–40	0.0106	0.0013	0.0019	0.0074	0.0047	0.0016	0.0006	0.0024
40–45	0.0147	0.0026	0.0043	0.0078	0.0078	0.0033	0.0014	0.0031
45–50	0.0224	0.0053	0.0081	0.0090	0.0124	0.0058	0.0027	0.0039
50–55	0.0367	0.0105	0.0150	0.0111	0.0206	0.0098	0.0051	0.0057
55–60	0.0594	0.0190	0.0254	0.0150	0.0320	0.0148	0.0092	0.0081
60–65	0.0928	0.0298	0.0413	0.0217	0.0515	0.0214	0.0174	0.0127
65–70	0.1359	0.0421	0.0625	0.0313	0.0775	0.0279	0.0304	0.0192
70–75	0.2072	0.0576	0.0991	0.0505	0.1197	0.0355	0.0545	0.0297
75–80	0.2983	0.0709	0.1488	0.0786	0.1823	0.0401	0.0959	0.0463
80–85	0.4242	0.0835	0.2243	0.1164	0.2921	0.0461	0.1721	0.0739

[a] Figures had been individually rounded off in computer. Expected relations may not hold exactly.

MULTIPLE DECREMENT TABLE In Table 5, typical multiple decrement tables show the relative importance of different risks of death and changes in the importance as age advances. In the present case, three risks are included: malignant neoplasms (R_1), major CV diseases (R_2), and other causes (R_3). The sum of the crude probabilities equals the probability of dying: $Q_{i1}+Q_{i2}+Q_{i3} = q_i$, for each age interval (x_i, x_{i+1}).

Under age 50, mortality level was low, with neither neoplasms nor CV diseases playing an important role in the probability of dying. From age 50 on, both CV diseases and neoplasms began to assert their influence. For white males, CV disease definitely was the greater risk of death. In age interval (50, 55), the probability of dying from CV diseases was about 40% the probability of dying from all causes ($Q_{i2}/q_i = .40$). This proportion increased steadily with age: from 45% for age interval (60, 65), to 50% for age interval (75, 80),

Annual Reviews
www.annualreviews.org/aronline

Table 6 Probability of dying in each age interval and the effect of eliminating major cardiovascular diseases as a risk of death, United States white male population and white female population, 1986[a]

	White Males				White Females			
Age Interval (in years) x_i to x_{i+1}	CV diseases present q_i	CV diseases eliminated $q_{i.2}$	Difference $q_i - q_{i.2}$	Percent change $\frac{q_i - q_{i.2}}{q_i} \times 100$	CV diseases present q_i	CV diseases eliminated $q_{i.2}$	Difference $q_i - q_{i.2}$	Percent change $\frac{q_i - q_{i.2}}{q_i} \times 100$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
00–01	0.0097	0.0094	0.0003	2.96	0.0075	0.0073	0.0002	2.82
01–05	0.0021	0.0020	0.0001	4.59	0.0016	0.0015	0.0001	5.65
05–10	0.0013	0.0012	0.0000	3.95	0.0009	0.0008	0.0000	5.17
10–15	0.0017	0.0017	0.0001	3.74	0.0010	0.0009	0.0001	5.53
15–20	0.0062	0.0060	0.0002	2.49	0.0025	0.0024	0.0001	3.46
20–25	0.0082	0.0080	0.0002	2.77	0.0026	0.0024	0.0001	5.61
25–30	0.0078	0.0075	0.0004	4.74	0.0027	0.0025	0.0002	8.86
30–35	0.0090	0.0082	0.0008	8.99	0.0033	0.0030	0.0004	10.88
35–40	0.0106	0.0087	0.0019	17.82	0.0047	0.0041	0.0006	13.56
40–45	0.0147	0.0104	0.0042	28.87	0.0078	0.0064	0.0014	18.08
45–50	0.0224	0.0144	0.0080	35.78	0.0124	0.0097	0.0027	21.76
50–55	0.0367	0.0218	0.0149	40.55	0.0206	0.0156	0.0051	24.63
55–60	0.0594	0.0345	0.0249	41.96	0.0320	0.0229	0.0091	28.42
60–65	0.0928	0.0525	0.0402	43.36	0.0515	0.0344	0.0171	33.25
65–70	0.1359	0.0759	0.0601	44.20	0.0775	0.0479	0.0296	38.23
70–75	0.2072	0.1141	0.0931	44.93	0.1197	0.0670	0.0527	43.99
75–80	0.2983	0.1627	0.1356	45.46	0.1823	0.0910	0.0913	50.08
80–85	0.4242	0.2291	0.1951	46.00	0.2921	0.1323	0.1598	54.70
85+	1.0000	1.0000	0.0000	0.00	1.0000	1.0000	0.0000	0.00

[a] Figures had been individually rounded off in computer. Expected relations may not hold exactly.

Annual Reviews
www.annualreviews.org/aronline

Table 7 Expectation of life at each age x_i and the effect of eliminating major cardiovascular diseases as a risk of death, United States white male population and white female population, 1986

	White Males				White Females			
Age (in years) x_i	CV diseases present e_i	CV diseases eliminated $e_{i.2}$	Difference $e_{i.2} - e_i$	Percent change $\frac{e_{i.2} - e_i}{ei} \times 100$	CV diseases present e_i	CV diseases eliminated $e_{i.2}$	Difference $e_{i.2} - e_1$	Percent change $\frac{e_{i.2} - e_i}{ei} \times 100$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
00	72.08	80.39	8.31	11.52	79.08	91.24	12.16	15.37
01	71.79	80.15	8.36	11.65	78.68	90.91	12.23	15.54
05	67.93	76.30	8.37	12.32	74.81	87.05	12.24	16.36
10	63.02	71.39	8.37	13.29	69.87	82.12	12.25	17.53
15	58.12	66.51	8.39	14.43	64.94	77.19	12.25	18.87
20	53.47	61.90	8.43	15.77	60.09	72.37	12.28	20.44
25	48.89	57.38	8.49	17.36	55.24	67.54	12.30	22.27
30	44.26	52.79	8.53	19.28	50.38	62.70	12.32	24.45
35	39.63	48.20	8.57	21.62	45.54	57.88	12.34	27.09
40	35.03	43.60	8.57	24.47	40.74	53.11	12.37	30.34
45	30.51	39.03	8.52	27.93	36.04	48.43	12.39	34.37
50	26.15	34.56	8.41	32.18	31.46	43.88	12.42	39.46
55	22.04	30.27	8.23	37.35	27.07	39.53	12.46	46.03
60	18.27	26.26	7.99	43.77	22.88	35.39	12.52	54.73
65	14.86	22.57	7.71	51.86	18.97	31.56	12.59	66.34
70	11.79	19.21	7.42	62.90	15.34	28.01	12.67	82.56
75	9.19	16.35	7.16	77.81	12.07	24.83	12.76	105.76
80	7.02	14.03	7.01	99.89	9.18	22.06	12.88	140.30
85	5.38	12.47	7.09	131.61	6.90	20.03	13.13	190.45

Annual Reviews
www.annualreviews.org/aronline

and 53% for age interval (80, 85). Thus, from age 75 on, about one in every two white male deaths was attributable to CV diseases.

The absolute value of the probability of dying from CV diseases also increased with age. For white males, the probability was $Q_{i2} = 150$ per 10,000 for age interval (50, 55), to $Q_{i2} = 2243$ per 10,000 for age interval (80, 85)—a 1400% increase over 35 years of life, or 40% per year!

Malignant neoplasms are the second most important risk of death. During 1986, for white males, the probability of dying from malignant neoplasms also increased with age: from $Q_{i1} = 105$ per 10,000 for age interval (50, 55) to $Q_{i1} = 835$ per 10,000 for age interval (80, 85), which is nearly a 700% increase over the 35 years. For white males, the risk of dying from neoplasms was about 70% as high as major CV diseases in the age interval 50–65 years. Beyond age 65, the relative importance of neoplasms decreases with age, because CV diseases became the dominant risk of death.

The mortality pattern among white females differs from that among white males. Table 5 confirms the general impression that females live longer. Between ages 45 and 75, the sex ratio of the probability of dying, $q_i(f):q_i(m)$, was consistently lower than 60%. Also, CV diseases were not as overwhelming a risk of death among white females as among white males. Below age 65, the probability of dying from CV diseases was lower than that from neoplasms. Beyond age 65, CV diseases overtook neoplasms and assumed the role of the major risk of death among white females.

IMPACT OF CARDIOVASCULAR DISEASES ON HUMAN MORTALITY AND HUMAN LONGEVITY Major cardiovascular diseases, as illustrated in Tables 5–7, have caused more deaths in the human population than has any other diseases. To evaluate the impact of the CV diseases on human longevity, we can compare the mortality and survival experience of the current population with the hypothetical experience of the same population under the conditions that would exist if major CV diseases were removed as a risk of death. The basic quantities needed for this purpose are the probability q_i and the net probability $q_{i.2}$ that a person alive at age x_i will die in age interval (x_i, x_{i+1}) if CV diseases (R_2) were eliminated as a risk of death.

The required data are age-specific death rate M_i and age-cause-specific death rate M_{i2} for each age interval (x_i, x_{i+1}), given in Table 3 for white males and in Table 4 for white females. Using a procedure described in Chiang (12), two life tables had been constructed for each group: one based on the probability q_i using formula 28, and other based on the net probability $q_{i.2}$ computed from formula 32. The probabilities and the corresponding expectations of life shown in Tables 6 and 7 reflect in different ways the effect of the CV diseases on human mortality.

Table 6 gives a comparison between the probabilities q_i and $q_{i.2}$. The

Annual Reviews
www.annualreviews.org/aronline

difference $q_i - q_{i.2}$ is the reduction in probability of dying if CV diseases were eliminated, or alternatively, the excess probability of dying because of the presence of CV diseases. The difference was not pronounced before age 40 because the disease then is quite rare, but advances with age at an accelerated rate. If CV diseases were removed, the reduction in the probability of dying in age interval (40, 45) would be 28.8% for white males and 18% for white females. For age interval (50, 55), the reduction in the probability of dying would be over 40% for white males and nearly 25% for white females. At age 70 or older, the reduction would be about 45% for both white males and females.

The impact of the major CV diseases on the expectation of life are shown in Table 7, where e_i is the "real" expectation of life in the current population under the normal condition with the presence of all causes of death, whereas $e_{i.2}$ is the (hypothetical) expectation of life if CV diseases were eliminated as a cause of death. The difference $e_{i.2} - e_i$ is the increase in the life expectancy if CV diseases were eliminated, or the number of years lost because the presence of CV diseases as a risk of death. The nearly constant difference $e_{i.2} - e_i$ under age 50 was because CV diseases cause death mainly among older persons and the reduction in the expectation of life because of CV diseases occurred almost entirely in persons older than age 50.

As the CV diseases became an increasingly dominant cause of death in older ages, the expectations of life $e_{i.2}$ became much greater than the expectation e_i. At age 60, the expectation of life for white males was $e_{i.2} = 26.2$ years and $e_i = 18.2$ years, with a difference of $e_{i.2} - e_i = 8$ years, a 43% reduction because of the presence of CV diseases. The corresponding reduction for white females was 54.7%. At age 70, the expectations were $e_{i.2} = 19.2$ and $e_i = 11.7$ for white males with a relative reduction of 62%, and $e_{i.2} = 28$ and $e_i = 15.3$ for white females with a relative reduction of 82.6%. Thus, if the major CV diseases were eliminated as a cause of death, a white male could expect an additional 8 years of life at age 60, and nearly 7½ additional years of life at age 70. For white females, the increase in the life expectancy would be even more impressive. If CV diseases were eliminated, a white female could enjoy an additional 12½ years of life at age 60 and at age 70.

ACKNOWLEDGMENTS

A computer program and Tables 5–7 have been prepared by C. Y. Liu. Copies of the program, which are not included because of the length of the paper, are available for distribution to interested readers. I wish to thank L. Le Cam for a helpful comment, B. J. van den Berg and M. White for reading the manuscript, and Mr. Liu for his help in the computations.

Annual Reviews
www.annualreviews.org/aronline

Literature Cited

1. Armitage, P. 1959. The comparison of survival curves. *J. R. Stat. Soc. A.* 122:279–300
2. Bailey, W. G., Haycocks, H. W. 1946. *Some Theoretical Aspects of Multiple Decrement Life Table*. Edinburgh: Inst. Actuar., London Fac. Actuar. 40 pp.
3. Barlow, R. E., Prochan, F. 1975. *Statistical Theory of Reliability and Life Testing*. New York: Holt, Rinehart & Winston. 290 pp.
4. Berkson, J., Gage, R. P. 1952. Survival curve for cancer patients following treatment. *J. Am. Stat. Assoc.* 47:501–15
5. Bernoulli, D. 1760. Essai d'une nouvelle analyse de la mortalité causée par la petite verole et des avantages de l'inoculation pour la prévenir. *Historie avec les Memoire,* pp. 1–45. Paris: Acad. Sci.
6. Birnbaum, Z. W. 1979. On the mathematics of competing risks. *Natl. Cent. Health Stat. Vital Health Stat.* Ser. 2, No. 77:1–58
7. Boardman, T. J., Kendall, P. J. 1970. Estimation in compound exponential failure models. *Technometrics* 12:891–900
8. Bur. Census. 1987. *United States Population Estimates, by Age, Sex, and Race:* 1980 to 1987
9. Chiang, C. L. 1961. On the probability of death from specific causes in the presence of competing risks. *Proc. 4th Berkley Symp. Math. Stat. Prob.* J. Neyman ed., IV:169–180
10. Chiang, C. L. 1961a. A stochastic study of the life table and its applications: III. The follow-up study with the consideration of competing risks. *Biometrics* 17:57–78
11. Chiang, C. L. 1961b. Standard error of the age-adjusted death rate. *Natl. Cent. Health Stat. Vital Stat. Spec. Rep.* 47:273–85
12. Chiang, C. L. 1968. *Introduction to Stochastic Processes in Biostatistics*. New York: Wiley. 313 pp.
13. Chiang, C. L. 1970. Competing risks and conditional probabilities. *Biometrics* 26:767–76
14. Chiang, C. L. 1972. On constructing current life tables. *J. Am. Stat. Assoc.* 67:538–41
15. Cornfield, J. 1957. The estimation of the probabilities of developing a disease in the presence of competing risks. *Am. J. Public Health* 47:601–7
16. Cutler, S. J., Ederer, F. 1958. Maximum utilization of the life table method in analysing survival. *J. Chron. Dis.* 8:699–712
17. David, H. A. 1970. On Chiang's proportionality assumption in the theory of competing risks. *Biometrics* 26:336–39
18. David, H. A. 1974. Parametric approaches to the theory of competing risks. *Reliab. Bio.* 275–90
19. David, H. A., Moeschberger, M. L. 1978. *The Theory of Competing Risks*. London: Griffin. 103 pp.
20. Dorn, H. 1950. Methods of analysis for follow-up studies. *Hum. Biol.* 22:238–48
21. Drolette, M. E. 1975. The effect of incomplete follow-up. *Biometrics* 31:135–44
22. Elandt-Johnson, R. C., Johnson, N. L. 1980. *Survival Models and Data Analysis*. New York: Wiley. 457 pp.
23. Elveback, L. 1958. Estimation of survivalship in chronic disease. *J. Am. Stat. Assoc.* 53:420–40
24. Fix, E., Neyman, J. 1951. A simple stochastic model of recovery, relapse, death and loss of patients. *Hum. Biol.* 23:205–41
25. Frost, W. H. 1933. Risk of persons in familiar contact with pulmonary tuberculosis. *Am. J. Public Health* 23:426–32

25a. Gail, M. 1975. A review and critique of some models in competing risk analysis. *Biometrics* 31:209–22

25b. Greville, T. N. E. 1943. Short methods of constructing abridged life table. *Rec. Am. Inst. Actuar.* 32:29–43

26. Greville, T. N. E. 1948. Mortality tables analysis by cause of death. *Rec. Am. Inst. Actuar.* 37:283–94
27. Halley, E. 1693. An estimate of the degrees of the mortality of mankind, drawn from curious tables of the births and funerals at the city of Breslau. *Philos. Trans. R. Soc. London* 17:569–610
28. Hooker, P. F., Longley-Cook, L. H. 1957. *Life and Other Contingencies,* Vol. II. Cambridge: Cambridge Univ. Press. 256 pp.
29. Kalbfleisch, J. D., Prentice, R. L. 1980. *The Statistical Analysis of Failure Time Data*. New York: Wiley. 321 pp.
30. Kaplan, E. L., Meier, P. 1958. Nonparametric estimation from incomplete observations. *J. Am. Stat. Assoc.* 53:457–81
31. Karn, N. M. 1931. An inquiry into various death rates and the comparative influence of certain diseases on the duration of life. *Ann. Eugen.* 4:279–326

31a. Keyfits, N. 1966. A life table that

Annual Reviews
www.annualreviews.org/aronline

agrees with the data. *J. Am. Stat. Assoc.* 61:305–12
32. Kimball, A. W. 1969. Models for the estimation of competing risks from grouped data. *Biometrics* 25:329–37
32a. King, G. 1914. On a short method of constructing an abridged mortality table. *J. Inst. Actuar.* 48:294–303
33. Kuzma, J. W. 1967. A comparison of two life table methods. *Biometrics* 23:51–64
34. Lee, L., Thompson, W. A. 1974. Results on failure time and pattern for the series system. *Reliab. Bio.* Phila. Soc. Ind. Appl. Math, pp. 291–302
35. Littel, A. S. 1952. Estimation of T-year survival rate from follow-up studies over a limited period of time. *Hum. Biol.* 24:87–116
36. Makeham, W. M. 1874. On an application of the theory of the composition of decremental forces. *J. Inst. Actuar.* 18: 317–22
37. Mantel, N., Bailer, J. C. 1970. Letter to the editor. *Biometrics* 26:861–63
38. Marshall, A. W., Olkin, I. 1967. A multivariate exponential distribution. *J. Am. Stat. Assoc.* 62:30–44
39. Miller, D. R. 1977. A note on independence of multivariate lifetimes in competing risks models. *Ann. Stat.* 5:576–79
40. Miller, R. G. 1981. *Survival Analysis.* New York: Wiley. 238 pp.
41. Moeschberger, M. L., David, H. A. 1971. Life tests under competing causes of failure and the theory of competing risks. *Biometrics* 27:909–33
42. Moeschberger, M. L., Klein, J. P. 1984. Consequence of departure from independence in exponential series system. *Technometrics* 262:277–84
43. Natl. Cancer Inst., Div. Cancer Prev. Control Surveill. Program. 1989. *Cancer Statistics Review,* 1973–1986. US Dep. Health Hum. Serv. PHS, NIH Publ. No. 89–2789
44. Natl. Cent. Health Stat. 1986. *Vital Stat. of the US,* Vol. II, Part A
45. Natl. Cent. for Health Stat. 1987. *Life Table. Vital Stat. of the US,* 1987. Vol. II, Sec. 6:1–19
46. Pike, M. C. 1970. A note on Kimball's paper "Models for the estimation of competing risks from grouped data." *Biometrics* 26:579–81
47. Prentice, R. L., Kalbfleisch, J. D., Peterson, A. V., Jr., Flournoy, N., Farewell, V. T., Breslow, N. E. 1978. The analysis of failure times in the presence of competing risks. *Biometrics* 34:541–54
47a. Reed, L. H., Merrell, M. 1939. A short method of constructing an abridged life table. *Am. J. Hyg.* 30:33–62
47b. Sirken, M. G. 1964. Comparisons of two methods of constructing abridged life tables by reference to a "standard" table. *Vital Health Stat.*, Ser. 2, No. 4 1–11. Natl. Cent. Health Stat., US Dep. Health Educ. Welf.
48. Spurgeon, E. F. 1922. *Life Contingencies.* Cambridge: Cambridge Univ. Press. 479 pp.
49. Todhunter, I. 1865. *A History of the Mathematical Theory of Probability.* New York: Chelsea. 624 pp.
50. Tsiatis, A. 1975. A nonidentifiability aspect of the problem of competing risks. *Proc. Natl. Acad. Sci. USA* 72: 20–22
51. Wong, O. 1977. A competing risks model based on the life table procedure in epidemiological studies. *Int. J. Epidemiol.* 6:153–59