Methods for Evaluating the Heterogeneity of Aging Processes in Human Populations Using Vital Statistics Data: Explaining the Black/White Mortality Crossover by a Model of Mortality Selection

By Kenneth G. Manton[1] and Eric Stallard[1]

ABSTRACT

The progress of physiological aging processes in human populations is often studied by analyzing the age trajectory of total and cause specific mortality rates under the assumption that the rate of increase in the risk of death indexes the average rate of change in the physiology of the individuals comprising the population. Unfortunately, such efforts have usually assumed that the age trajectory of the average mortality rate is the same as the age trajectory of any given individual's risk of death– equivalent to the assumption that all individuals in the population have the identical rate of aging. In this paper, we present a model that permits the evaluation of the aging process of individuals from the age trajectory of mortality rates in the population by a. positing a model of the distribution of individual constitutional differences in the age trajectory of mortality risks, and b. adjusting the population mortality rates under the model of heterogeneity to retrieve the individual risks. This model is applied to human mortality data from the U.S. Black and White populations for the period 1935 to 1975. This example was selected because of the observation of a mortality crossover (Blacks having relatively lower mortality rates) about age 75. The crossover could be explained under the proposed model of population heterogeneity and differential mortality selection. The implications of the model for estimating the heterogeneity of aging processes for individuals from human population data are discussed as well as implications for our existing perception of the human aging mechanism.

A number of investigations have assumed that the rate of aging of individual organisms in a population can be inferred from the age increase in the mortality rates for a population. Such a strategy has been employed for both animal models (Simms, 1942) and human models (Strehler, chapter 5, 1977; Brown and Forbes, 1974; Rosenberg et al. 1973; Sims, 1946). Unfortunately, the computational procedures usually employed in such efforts assume that the population under study is homogeneous, i.e., that all persons age identically and have the same rate of increase of mortality risks. If the study population is *not* homogeneous, then the rate of age increase in the risk of death for individuals will diverge from the rate of

[1]Center for Demographic Studies, Duke University, 2117 Campus Drive, Durham, North Carolina 27706

Human Biology, February 1981, Vol. 53, No. 1, pp. 47-67.

age increase in mortality rates for the population. This is because the least robust individuals (those who have the fastest rates of aging) will die first leaving, at successive ages, an increasingly robust population residual. Thus, because age increases in the mortality rates of a heterogeneous population represent the average risk in a systematically selected group of survivors, they provide a biased perception of the age rate of physiological changes for the individual (Vaupel et al. 1979a). As a consequence, calculation of certain biological rate constants of the aging process for individual organisms from human mortality data assuming either a Gompertz aging function (the risk of death is an exponential function of age; Strehler, chapter 5, 1977) or a Weibull aging function (the risk of death is a power function of age, Rosenberg et al. 1973) will be inappropriate if the human population is sufficiently heterogenous. Our knowledge of human genetic differences influencing aging and longevity and the considerable variability in susceptibility to major degenerative diseases (i.e., cardiovascular disease, cancer, stroke), suggests that the assumption of homogeneity for human populations is inappropriate.

Consequently, we present a model in this paper which allows us to examine the rate of aging of individuals from human population data by adjusting the age specific mortality rates for the degree of heterogeneity manifest in the human population and the magnitude of mortality selection to any given age. Operationally, we modified standard life table calculations (Chiang, 1968) to reflect the dependence of mortality rates at advanced ages upon the selection of earlier mortality levels on a heterogeneous population. An approach involving an adjustment of life table parameters was adopted since, for total mortality, the most common forms of population aging functions (e.g., Gompertz, Weibull) were derived without consideration of the effects of heterogeneity. By utilizing age specific life table parameters we are not constrained to a particular functional form. In other analyses, we have applied models which do assume a particular aging function for specific disease processes where biological models are available to rationalize the function selected, e.g., lung cancer mortality among females (Manton and Stallard, 1979) or breast cancer (Manton and Stallard, 1980). Total mortality, which represents a mixture of chronic disease mortality risks, will not have such a simple functional relation to age.

To illustrate the model, we have selected an example with important social and economic implications over which there has been much recent debate: the crossing to lower levels of White age specific mortality rates by Black age specific mortality rates at advanced ages (Thornton and Nam,

1972; Kitagawa and Hauser, 1973; Spiegelman, 1968). Since the socioeconomic factors that dictated the relative Black mortality disadvantage at younger ages are unlikely to be reversed at advanced ages, it seemed implausible that a convergence—let alone crossover—of the mortality rates should occur. The first attempts at explanation of the phenomenon suggested that the crossover was an artifact due to data problems (i.e., enumeration error and age misreporting). However, despite the fact that data quality is recognized as an important issue, the developing consensus is that this explanation is unsatisfactory because of the scope and variety of evidence supporting the existence of a crossover. For example, the crossover is manifest in the life tables produced by the National Center for Health Statistics for 1969 to 1971. Crossovers are found in cohort as well as period mortality rates (NCHS, 1972). Use of the Siegel (1974) census enumeration adjustments (the adjustments preferred by the Census Bureau) fails to eliminate the crossover (Rives, 1977). Even efforts to explain the crossover by more extreme assumptions about the effect of census enumeration error and age misreporting have failed to eliminate the crossover (Kitagawa and Hauser, 1973). Nam and Ockay (1977) found the existence of mortality crossovers to be internationally pervasive when contrasting populations with very different mortality patterns. Manton et al. (1979a) found that the Black/White mortality crossover existed in cause specific data, that the age at crossover varied by cause of death, and that the different age patterns of mortality rates specific to race and sex were consistent with independent epidemiologic evidence. Furthermore, even if the crossover could be explained by data problems at advanced ages, there would still be the difficulty of explaining a convergence in mortality rates—a convergence which begins at relatively early ages. Vaupel et al. (1979a) found a similar convergence to exist in male and female total mortality rates for Sweden, based on data which are much less likely to be subject to serious enumeration error. Consequently, the weight and consistency of the evidence indicate that careful consideration should be made of the population mechanisms by which the crossover (convergence) might occur (Lew and Seltzer, 1970; Spiegelman, 1968).

To examine the Black/White mortality crossover, life tables are separately calculated for the Black and White populations in the U.S. over the period 1935 to 1975 based upon the assumptions: a. that each population is heterogeneous; b. that the initial distribution of individuals in each population (Black versus White) is identical (within sex) with respect to variables relevant to longevity, and c. that an individual's susceptibility to environmental conditions is fixed at birth. The life tables adjusted for

heterogeneity and mortality selection will thus represent the hypothetical mortality experience of "standard" individuals (i.e., individuals who, at birth, have the average endowment for longevity) and not the mortality rates for the population alive at each age—rates which are based on the mixture of persons still alive at that age. The ratio of the Black and White individual age specific mortality risks (within sex) can then be calculated to determine if the adjustment for heterogeneity and mortality selection is sufficient to remove the crossover. Examination of the age patterns of these ratios (or "relative risks") will also help to suggest additional factors affecting the mortality experience of these two groups.

Materials and Methods

A Model of Selection

Calculation of the hypothetical life tables for standard individuals requires that certain assumptions be made both about the initial distribution of individuals with respect to variables relevant to survival and about the change of that distribution due to selection. We assume that the partial differential equation governing the change of the distribution as a cohort ages is:

$$\frac{\partial}{\partial x} f_x(z) = f_x(z) \ (\bar{\mu}_x - \mu_x(z)). \tag{1}$$

In (1), $f_x(z)$ is the probability density function (pdf) of longevity characteristics, z's, for survivors to age x; $\mu_x(z)$ is the force of mortality at age x for a person at a given value of z; and $\bar{\mu}_x$ is the average force of mortality in the cohort at age x. In writing (1), we assume that each person is "tagged" with a value of z at birth and retains this value throughout his life. Thus, $f_0(z)$ represents the initial (at birth) pdf of the z's in the cohort while $f_x(z)$, $x>0$, represents the pdf at some later time. Note that because $\mu_x(z)$ is dependent on z, $f_0(z)$ and $f_x(z)$ will, in general, be different—implying that the proportion of the cohort at any given level of z will change over time as mortality selectively removes individuals at different z values.

Substantively, for a birth cohort, a longevity characteristic z will be taken to mean the sum of effects of a series of genetic traits determining both the individual's initial ability to resist environmental stress and the age trajectory of that ability. Strehler (1977) both identifies a number of genetic traits that are suspected to underlie human longevity and attempts to relate them to the specific functional age losses (e.g., neurological and sensory performance, age declines in immunological performance)

that decrease the viability of the organism with age. Note that one of the assumptions of polygenic determination of factors affecting longevity is that genetic traits will tend to covary, i.e., that a person with a given endowment for longevity will have a strong probability of possessing a pattern of traits which functions collectively to determine his ability to survive the environment.

Note that the model is sufficiently general to permit the origin of the age dimension to be any age past birth, so that a cohort's survival need not be monitored from birth. Indeed, in developed societies mortality selection may be quite moderate before age 50 and rapidly increasing only thereafter. In examining the partial survival experience of a cohort the longevity traits do not have to be assumed fixed prior to the age of interest (say age 50). This would permit us to examine a much broader range of longevity characteristics since environmental effects that had affected the viability of the organism before that age could be studied. For example, the family history of breast cancer is known to be an important risk factor likely due to hereditary characteristics. The genetic traits determining the age specific risks of the disease would determine the longevity characteristics associated with this disease. In addition to familial risk, however, age at first pregnancy has been identified as a factor that permanently alters the women's risk of breast cancer. In particular, early age at first pregnancy (before age 25) has been identified as a behavioral factor which permanently lowers a woman's risk of breast cancer. By applying a model of heterogeneity and mortality selection to the age trajectory of breast cancer risks for females aged 25 and over (25 being an age after which both early age at first pregnancy and familial risk were fixed) we were able to estimate the likely degree of heterogeneity in the female population (Manton and Stallard, 1980). Indeed it was found that familial susceptibility, which induces a risk of about 2 or 3 to 1 in the general population, could be identified with extremely high risk levels (on the order of 50 to 1) in a small group of persons—a finding consistent with the results of studies of breast cancer risk in women with early breast cancer onset and specific family pedigrees of risk. It should be emphasized, however, that one of the prime functions of positing a model of heterogeneity is to search for the manifestation of fixed determinants of longevity in the same sense that analyses are made of the transmissions of traits to discover the likely existence of underlying genetic effects. Hence, one of the functions of such analyses is to discover evidence from natural human populations for the existence of such longevity characteristics and to examine certain of its properties, e.g., its dimensionality.

In the following development we will, however, assume that each

cohort is monitored from birth. In this case, we can solve (1) to express $f_x(z)$ as a function of $f_0(z)$:

$$f_x(z) = f_0(z)\, s_x(z)/\bar{s}_x, \tag{1a}$$

where $s_x(z)$ is the probability that an individual at a given value of z survives to at least age x, i.e.,

$$s_x(z) = \exp[-\int_0^x \mu_t(z)dt];$$

and where $\bar{s}_x$ is the average probability of survival to at least age x, i.e.,

$$\bar{s}_x = \exp[-\int_0^x \bar{\mu}_t\, dt].$$

Equation (1a) shows that the proportion of the cohort at age x with a given value of z, $f_x(z)\Delta z$, is simply equal to the proportion who initially had that value of z, $f_0(z)\Delta z$, times the probability of surviving to age x at that value of z, $s_x(z)$, and renormalized via $\bar{s}_x$ to adjust for those who died in the interval 0 to x.

To describe the model, it is necessary to make assumptions about the functional forms of both $\mu_x(z)$ and $f_x(z)$. We consider $\mu_x(z)$ first. The simplest assumption we can make concerning $\mu_x(z)$ is that it is proportional to z:

$$\mu_x(z) = z\, \mu_x(1) \tag{2}$$

where $\mu_x(1)$, or more compactly, μ_x, is the force of mortality for the standard individual. Thus, the force of mortality for any given level of $z > 0$ is assumed to be proportional to the force of mortality for the standard individual. Note that because z is assumed to be fixed for the life of each individual, those individuals with $z > 1$ will face a higher force of mortality at all ages than that faced by the standard individual. Conversely, those individuals with $z < 1$ will face a force of mortality less than that of the standard individual. Thus z may be taken to be a measure of relative (to the standard individual) frailty or "susceptibility to death." Alternately, 1/z may be considered as a measure of vitality or "robustness."

Of particular interest is the cohort force of mortality, $\bar{\mu}_x$, which is defined as:

$$\bar{\mu}_x = \bar{z}_x\, \mu_x \tag{3}$$

where

$$\bar{z}_x = \int_0^\infty z\, f_x(z)\, dz$$

is the mean of z among survivors to age x. Clearly, the mean decreases with age since (1) and (2) imply that the derivative of the mean is proportional to the variance, or

$$\frac{\partial \bar{z}_x}{\partial x} = -\mu_x \, \mathrm{var}_x(z)$$

which is always negative. This means the "frailer" population members (with high z's) are being selected earlier than their more "robust" contemporaries (with low z's). The important point to note is that, except in the case of a homogeneous population, i.e., $\mathrm{var}_x(z) = 0$, $\bar{z}_x$ must decrease as x increases. In combination with equation (3), this implies that the cohort force of mortality will, in general, be different from the standard individual force of mortality and will tend to increase at a less rapid rate at older ages where the impact of mortality selection is stronger (see appendix in Vaupel et al. 1979).

The proportionality assumption has additional implications for the selection of the initial form of the distribution of z. First of all, since the force of mortality cannot be negative, it follows from (2) that z must be non-negative. Second, since the standard individual is defined to have the average endowment at birth, $\bar{z}_0$, for longevity, it follows that

$$\bar{\mu}_0 = \mu_0,$$

so that (3) implies:

$$\bar{z}_0 = 1.$$

Third, for parsimony it would be desirable that the parameters of f_x be unchanged for any x. A family of distributions which satisfies all three criteria is the one-parameter gamma distribution with pdf given by

$$f_x(z) = z^{k-1} \lambda_x^k \exp(-z \lambda_x)/\Gamma(k),$$

where $k>0$ is the shape parameter of the distribution; where

$$\lambda_x = k + \int_0^x \mu_t \, dt;$$

and where $\Gamma(k)$ is the gamma function defined as

$$\Gamma(k) = \int_0^\infty z^{k-1} \exp(-z) \, dz.$$

We chose the one parameter form of the gamma distribution because of our proportionality assumption which meant that the effects of selection were always represented as the ratio of quantities so that the absolute

level of mortality cancelled out in the ratio. This allowed us to fix the initial scale parameter, λ_0, of the gamma distribution without loss of generality. For the gamma distribution, the mean and variance of z among survivors to age x are given by

$$\bar{z}_x = k/\lambda_x$$

and

$$\text{var}_x(z) = \bar{z}_x^2/k. \tag{4}$$

By solving (4) for k, i.e.,

$$k = \bar{z}_x^2/\text{var}_x(z),$$

we see that k may be alternately interpreted as the inverse of the squared coefficient of variation in z at any given age, x. This means that the assumption of a gamma pdf for $f_x(z)$ is equivalent to the assumption of a constant coefficient of variation in the force of mortality faced by the survivors to each given age.

Though the gamma pdf is, mathematically, a satisfactory choice for $f_x(z)$ under assumptions (1) and (2), its primary advantage as a representation of population heterogeneity is based on its great "flexibility". This flexibility is due to the property that the shape parameter k may be adjusted to represent many of the more familiar distributional forms, including the exponential and chi-squared distributions. This topic will be more fully discussed in the next section where the theoretical bases for the choices of k used in our analysis are presented.

Use of (3) and (4) and the definition

$$\bar{s}_x = \exp[-\int_0^x \bar{\mu}_t \, dt],$$

where $\bar{s}_x$ is the proportion of the cohort surviving to age x, yield the important relationship between the mean of z among survivors to age x and the proportion surviving:

$$\bar{z}_x = \bar{s}_x^{1/k}. \tag{5}$$

An alternate derivation of (5) based on the distribution of z among those who die at age x has been presented by Vaupel et al. (1979a). With this result, the equations for the hypothetical life tables can be expressed as simple functions of the cohort survivorship variables $\bar{s}_x$ and $\bar{s}_{x+n}$.

The age specific probability of death, ${}_nq_x$, for the standard individual is defined to be

$$_nq_x = 1 - \exp[-\int_x^{x+n} \mu_t \, dt]. \tag{6}$$

Using (5) and (3) in (6) and simplifying yields

$$_nq_x = 1 - \exp[k/\bar{s}_x^{1/k} - k/\bar{s}_{x+n}^{1/k}]$$

which shows that the hypothetical life table for the standard individual can be constructed given knowledge of the cohort survivorship at each age interval. This information can be obtained from the "l_x column" of published cohort life tables. Indeed, since (2) and (6) imply that the age specific probability of death, $_n q_x(z)$, for any level of z is

$$_nq_x(z) = 1 - (1 - {}_nq_x)^z,$$

it is clear that the hypothetical life table for individuals at any level z can be derived via

$$_nq_x(z) = 1 - \exp[kz/\bar{s}_x^{1/k} - kz/\bar{s}_{x+n}^{1/k}].$$

We now consider how this model may be used to compare the relative risks of two populations with a crossover in observed (cohort) mortality rates. The relative risk at age x can be defined as the ratio, $\bar{r}_x$, of the age specific forces of mortality in the two populations. Thus, using (3) and (5), we have

$$\bar{r}_x = r_x (\bar{s}_{x1}/\bar{s}_{x2})^{1/k} \tag{7}$$

where

$$r_x = \mu_{x1}/\mu_{x2}. \tag{8}$$

Thus, even if $r_x > 1$, a crossover ($\bar{r}_x \leq 1$) of the cohort forces of mortality would occur if, for any x,

$$[r_x]^k \leq \bar{s}_{x2}/\bar{s}_{x1}.$$

Interestingly, if r_x were constant for all x, (7) implies that $\bar{r}_x$ would approach 1, indicating a convergence but not a crossover of the mortality rates for the two populations.

Models of Longevity Distribution

To apply the selection model just discussed we must be able to specify the distribution of individual z's (i.e., longevity characteristics). This amounts to selecting the gamma shape parameter k. In other papers we have explored estimation procedures to derive values of k from the relative mortality risks of large population groups (Manton and Stallard,

1980). In this paper, however, we will select values of k on theoretical grounds in order to focus upon the biological (rather than statistical) properties of the selection mechanisms. In order to do so we will specify four theoretical assumptions about the distribution of longevity characteristics:

1. That the biological dimensions underlying longevity, say y_i, are normally distributed at birth. The y_i would be normally distributed if each is the sum of the effects of a large number of independent biological factors;
2. That for each y_i an "optimal" biological point can be defined such that deviations in either direction from this point will be associated with decreased survival. A first approximation to this point might be the mean of each y_i, denoted by $\bar{y}_i$;
3. That, conditionally on age, the force of mortality, $\mu_x(z)$, is a quadratic function of the y_i, i.e., for person p, $z_p = (y_{p1}-\bar{y})^2 + \ldots + (y_{pn}-\bar{y}_n)^2$ where n is the number of dimensions underlying longevity;
4. That each individual's endowment for longevity (z) is fixed at birth, i.e., z_p is independent of age.

Under these four assumptions the problem of determining the value of k for our selection model is reduced to a determination of n, the number of dimensions relevant to longevity. The relation of n to the gamma shape parameter is simply $n=2k$. Note that the lower the value of n the greater is the heterogeneity of individuals within a population (see equation 4).

The values of n that we will investigate are 1 and 2, suggesting that longevity is unidimensional and bidimensional, respectively. These values imply distributions of z values with k's of 0.5 and 1.0. The gamma distribution with $k=0.5$ is a chi-squared distribution with one degree of freedom and, for $k=1.0$ either a chi-squared distribution with 2 degrees of freedom or an exponential distribution. We restrict our attention to these values for a variety of reasons. First, Strehler (1977) has argued that the genetic determinants of longevity may be very few in number. Second, an analysis of the chronic disease process of cardiovascular morbidity suggests that, conditional on age, the risk of this type of disease was a quadratic function of a single linear component of the risk variables under study (Manton et al. 1979b). Third, attempts to estimate k, from data on the relative risks of males and females in the Swedish population (data which are not subject to the same degree of error as the U.S. Black population data, and for which an identifying assumption of constant relative risks beyond age 50 seemed defensible) suggest values of k between 0.5 and 1.0 (Manton et al. 1980). Fourth, models of individual disease

processes suggest very marked heterogeneity in disease susceptibility, implying low values of k (Manton and Stallard, 1979). Finally, it appears that values of $k > 1.0$ do not eliminate the crossover. Hence, the analysis is directed to the study of the implications of the two selected values of k, though it is clear that the model is sufficiently general to permit any number of dimensions to underlie longevity.

Data

A requirement of an analysis of population heterogeneity and selection is that a historical series of mortality and population data is available to estimate the cohort survivorship variable, $\bar{s}_x$. For several European nations such data are available, but for the U.S. there is a problem in that the death registration system was not completed until 1933. Furthermore, there are questions about the reliability of such data, especially for the Black population. The data used in this report came from many sources and were adjusted using census preferred estimates of underenumeration (Siegel, 1974; Keyfitz, 1979). Note that in general, adjustments for census underenumeration will tend to be greater for Blacks than Whites, especially for young adult males. As a consequence, these adjustments will tend to reduce the magnitude of the crossover.

Mortality data on U.S. race specific mortality before 1900 are sparse. However, life table estimates are available for Blacks in 1850 (Evans, 1962) and for Whites in 1850 (Jacobson, 1957). For years after 1900, the "official" vital statistics life tables for the Black and White populations are available though they are unadjusted for enumeration error. However, a number of alternates to the "official" life tables adjusted for enumeration error were available for this analysis. For the White population for 1901 and the decennial years 1910 to 1950 the life table estimates were taken from Coale and Zelnik (1963). For the Black population for 1901 and the decennial years 1910 to 1950 the "official" life tables were adjusted for enumeration error using the estimates of net census undercounts provided by Coale and Rives (1973). For 1960 to 1970 Black and White populations the census preferred estimates of Siegel (1974) were employed in recalculating the "official" unadjusted life tables. The adjustments employed by Siegel involved the Coale and Zelnik and Coale and Rives methodology for estimated underenumeration. Consequently, it is unlikely that bias would be introduced into the data at the point of crossing over from the Coale and Zelnik or Coale and Rives data to the Siegel adjusted data. For 1975, life tables were constructed from mortality counts provided by the National Center for Health Statistics and

the 1975 resident population projections provided by the U.S. Bureau of the Census. Single year of age estimates for all years from 1850 to 1975 were constructed from the base data by linear interpolation methods (described in Vaupel et al. 1979b). Using the interpolated mortality rates it was possible to compute the cohort survivorship variable for persons up to age 85 for both races and sexes for the years 1935 (1850 + 85) to 1975. The fact that the earliest available mortality data for both races were for 1850 restricted our attention to ages up to 85 in order to examine some temporal variation (i.e., the 40 years 1935 to 1975).

Results

In this section we will present a. examples of the effects of various levels of population heterogeneity on our perception of the individual aging trajectory, and b. the relative risk for standard individuals in the U.S. Black and White populations after adjustment for heterogeneity at the theoretically specified levels implied by the k values of 0.5 and 1.0. By using a pre-specified value of k we can examine age specific change in the relative risks of individuals in the two populations. In contrast, estimation of k would require some assumption about the behavior of the relative risks over age, thereby precluding such comparisons.

The age trajectory of mortality probabilities estimated from the population and estimated for the standard individual (i.e., a person at a pre-specified level of susceptibility) are presented in Figure 1.

The data presented in Figure 1 are for the White female birth cohort of 1875 so that the mortality experience of persons aged 90 occurred in 1965. The line representing the cohort mortality probabilities indicates the age trajectory of the average mortality risks of the population surviving to any given age. In contrast to this curve, we have the age trajectory for the person at the average level of frailty (i.e., $z = 1.0$) under three different assumptions about the degree of heterogeneity manifest in the population at birth. One can see that the age trajectories for both $k = 0.5$ and 1.0 are very different from the cohort age trajectory of risk. One essential feature of the age trajectories based on the heterogeneity model is that the annual percentage increase in any given individual's risk is larger than the annual percentage increase in the average risk of survivors, suggesting that an individual's age at death is more narrowly determined than the range implied by cohort mortality rates. It should be noted that the systematic bias of cohort mortality rates away from individual mortality rates will vary in degree (depending upon the level of

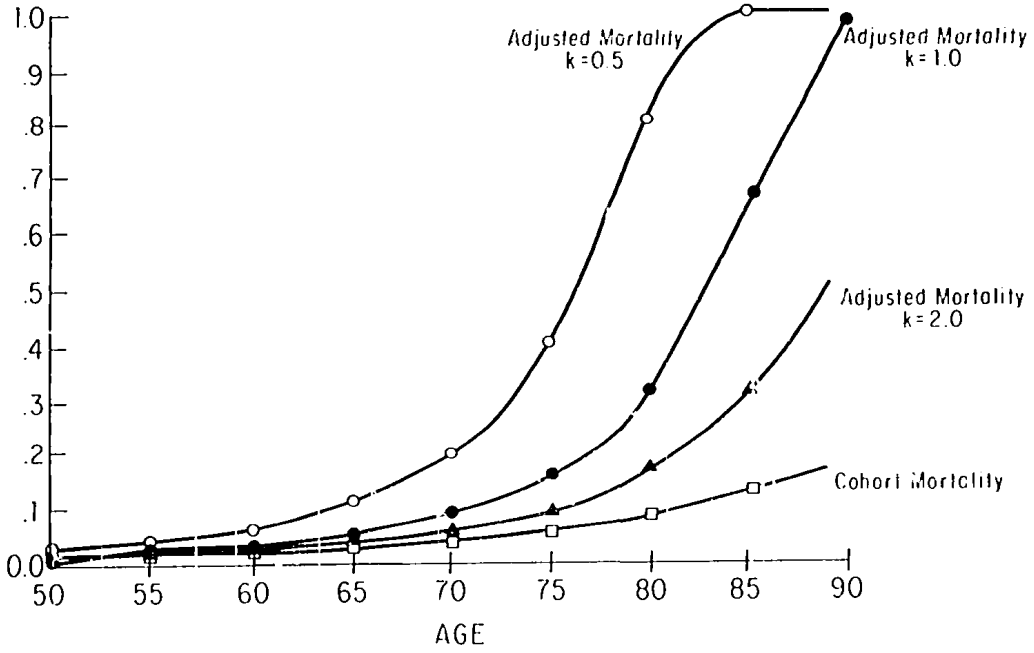

FIG. 1. Cohort and "Individual" Age Specific Mortality Probabilities for the 1875 White Female Birth Cohort

heterogeneity existing in the birth cohort) but not in kind. That is, examination of the third individual age trajectory (k = 2.0) shows the same pattern of deviation from the cohort experience so that with increasing k, the individual age trajectory will more closely approximate the cohort age trajectory but always with the same pattern of deviations. This can be understood by realizing the cohort mortality experience is an appropriate model for the individual only if k were infinite—corresponding to the case of a homogeneous population. In fact, of the four populations considered in this report, the presentation of the White female data represents the case where heterogeneity effects are the smallest because of the larger proportions of white females surviving to advanced ages. To understand the cumulative effects of the mortality age trajectories presented in Figure 1, consider the mean residual life expectancies for different ages and levels of heterogeneity presented in Table 1.

An examination of the residual life expectancies shows that the different age trajectories in Figure 1 imply quite different life chances for individuals than are perceived from the cohort life tables. The magnitude of the differences in life expectancy across different assumptions about

Table 1

Life Expectancy at Selected Ages and Different Assumed Levels of Heterogeneity for the 1875 White Female Birth Cohort

	Heterogeneity Level (k)			
AGE	0.5	1.0	2.0	(Cohort)
0	36.68	43.20	46.94	51.53
50	13.97	18.66	21.70	26.03
65	5.25	8.70	11.25	15.30

heterogeneity may be contrasted with the estimate that total elimination of cancer (about 17% of all U.S. female deaths) as a cause of death would increase life expectancy at birth by only 2.3 years (Keyfitz, 1977; naturally this calculation was based on the assumption of a homogeneous population). Even when compared with the effects of total elimination of cardiovascular mortality, about 56% of all U.S. female deaths, yielding a 17.1 year gain in life expectancy at birth (Preston et al. 1972), the 14.9, 8.3, and 4.6 year differences in cohort and individual life expectancy at birth in Table 1 are still seen to have a substantial impact on our perception of a given individual's life chances. Also, it should be remembered that, even for a heterogeneous population, the cohort life expectancy at birth represents the average number of years lived by the cohort members. Thus, whereas the standard individual has a life expectancy substantially below the cohort average, there will be numerous others with a life expectancy much larger than the cohort average.

Given the illustration of the magnitude of the effect of heterogeneity on individual age trajectories of mortality risk we can now present (Table 2) the mortality risks for Black males relative to White males (see equation (8)) for the years 1935, 1955, and 1975 for the unidimensional model (k = 0.5), the bidimensional model (k = 1.0) and the observed values. The corresponding results for females are presented in Table 3.

An examination of Table 2 shows that, as expected, a mortality crossover (i.e., the relative risk drops below 1.0) is manifested in the observed ratios for all three years. The observed ratios increase over the age range 15 to 35 years of age. For example, 35-year-old Black males in 1975 were at 2.9 times higher risk than their White male contemporaries whereas 15-year-old Black males were only at 1.3 times the White risk. From age 40 to 70 the ratios exhibit a substantial decline with the crossover occur-

ring by age 75 in 1935 and 1955 and by age 80 in 1975. With such high levels in the observed relative risk ratios, it would be expected that the effects of differential mortality selection would be increasingly manifest over age. In particular, if the crossover at advanced ages is an artifact of the early differential mortality selection, then the observed ratios at advanced ages understate to a considerable degree the greater hazards to which Black males are subjected in the U.S. population. In fact, Table 2 shows that, for both the bidimensional and unidimensional models, the crossover observed after age 75 is eliminated for the comparison of standard individuals in all three years. Interestingly, in both models the individual risk ratios begin to diverge from the observed risk ratios at about age 20, implying that throughout most of the adult life, not just at advanced ages, the observed risk ratios present a biased picture of the relative mortality risks of Blacks in the U.S. male population (Spiegelman,

Table 2

Age-Specific Mortality Risk Ratios—Black Males vs. White Males—for the Years 1935, 1955, and 1975

	Observed Ratios			Bidimensional Model (k = 1.0)			Unidimensional Model (k = 0.5)		
Age	1935	1955	1975	1935	1955	1975	1935	1955	1975
0	1.26	1.62	2.04	1.26	1.63	2.05	1.27	1.65	2.07
5	1.18	1.60	1.44	1.19	1.63	1.47	1.21	1.66	1.50
10	1.31	1.27	1.47	1.32	1.30	1.50	1.33	1.34	1.54
15	1.91	1.40	1.30	1.91	1.45	1.33	1.91	1.49	1.36
20	2.48	1.52	1.44	2.65	1.56	1.48	2.83	1.60	1.52
25	2.66	1.88	2.51	3.12	1.92	2.60	3.66	1.96	2.69
30	2.65	2.19	3.01	3.31	2.26	3.16	4.15	2.34	3.33
35	2.66	2.27	2.93	3.53	2.38	3.13	4.67	2.50	3.35
40	2.43	2.04	2.84	3.34	2.33	3.06	4.59	2.67	3.30
45	2.29	1.78	2.14	3.25	2.28	2.34	4.62	2.92	2.56
50	2.05	1.61	2.00	3.02	2.25	2.26	4.45	3.15	2.55
55	1.67	1.56	1.66	2.57	2.38	1.95	3.96	3.65	2.28
60	1.29	1.23	1.57	2.05	2.03	2.05	3.25	3.34	2.68
65	1.25	1.28	1.31	1.98	2.17	1.93	3.12	3.67	2.85
70	1.14	1.16	1.24	1.78	2.05	2.04	2.79	3.64	3.35
75	1.00	.97	1.27	1.60	1.72	2.43	2.55	3.04	4.63
80	.92	.84	.97	1.52	1.38	1.85	2.52	2.27	3.56
84	.81	.75	.64	1.39	1.03	1.09	2.36	1.41	1.84

1968). We can also see that the unidimensional model suggests the greatest differential in mortality risks between Blacks and Whites with the risk ratio for standard individuals rising as high as 4.7 for age 35 in 1935 compared with the observed ratio of 2.7.

An examination of the secular trend in the mortality ratios shows that, for the bidimensional model, standard individuals in the Black male population, ages 25 to 50, improved from 1935 to 1955 and that from 1955 to 1975 their situation deteriorates relative to Whites, though still improved over 1935. However, in the age range 70 to 80, Blacks were relatively worse off in 1975 compared to 1935. The most plausible explanation for this secular increase in the racial mortality differentials at older ages is the relatively more rapid reduction in individual White male mortality in recent years. An examination of the unidimensional model shows the same general pattern with a. much higher relative risks for Black males at a specified level of frailty; b. much of this increased relative risk maintained into advanced ages; and c. a secular decrease in mortality differentials again implied up to age 65 with an increase after age 65. The last point suggests that selection possibly might obfuscate a secular trend in the age specific relative risks, i.e., since the observed risk ratios in the over-65 population remain virtually constant over time.

An alternate explanation of the secular trends is that the age variability in the unidimensional model is the result of cohort effects. In particular, the standard individuals in the Black male cohort born between 1885 and 1900 (ages 35 to 50 in 1935) appear to be more disadvantaged than those in the adjacent cohorts and generally exhibit the highest risk ratios, regardless of which period is considered. For example, standard individuals in the Black male population aged 35 in 1935 had a risk ratio of 4.7. Twenty years later, they (aged 55 in 1955) had a ratio of 3.7, while in 1975 they had a ratio of 4.6. A similar pattern for standard individuals in this Black male cohort is also seen under the bidimensional model.

Table 3 shows that the secular trends in the observed female ratios are somewhat different from the corresponding secular trends for males. For example, in 1935, Black females were relatively more disadvantaged, to about age 70, than Black males; however, by 1975, the relative progress for Black females had been greater between ages 25 and 45 (the child bearing years). This suggests the importance of the reduction of maternal mortality among Black females from 1935 to 1975. For older ages, the observed relative disadvantage for Black females in 1975 is greater than for Black males in 1975. A comparison of the unidimensional and bidimensional selection models shows that the general pattern of age

Table 3

Age-Specific Mortality Risk Ratios—Black Females vs. White Females—for the Years 1935, 1955, and 1975

	Observed Ratios			Bidimensional Model (k = 1.0)			Unidimensional Model (k = 0.5)		
Age	1935	1955	1975	1935	1955	1975	1935	1955	1975
0	1.31	1.72	2.24	1.31	1.74	2 26	1.32	1.75	2.27
5	1.29	1.76	1.26	1.30	1.79	1.28	1.31	1.82	1.31
10	1.46	1.52	1.27	1.47	1.55	1.30	1.49	1.59	1.32
15	2.79	1.89	1.19	2.81	1.94	1.22	2.84	2.00	1.24
20	3.20	2.41	1.59	3.44	2.47	1.63	3.70	2.54	1.66
25	3.19	2.78	2.27	3.74	2.86	2.33	4.39	2.94	2.38
30	3.00	3.07	2.33	3.76	3.22	2.41	4.72	3.38	2.49
35	3.08	3.17	2.80	4.13	3.39	2.94	5.53	3.64	3.08
40	2.93	2.95	2.50	4.11	3.44	2.65	5.78	4.01	2.80
45	2.73	2.64	2.14	3.96	3.43	2.31	5.75	4.46	2.49
50	2.41	2.45	2.14	3.58	3.48	2.40	5.32	4.96	2.69
55	1.91	2.21	1.97	2.90	3.48	2.30	4.40	5.48	2.69
60	1.43	1.75	1.88	2.19	2.97	2.47	3.36	5.02	3.23
65	1.30	1.70	1.56	1.98	3.00	2.36	3.02	5.27	3.58
70	1.10	1.27	1.67	1.66	2.35	2.94	2.50	4.35	5.19
75	.84	.93	1.74	1.22	1.66	3.68	1.78	2.96	7.80
80	.68	.73	1.01	.98	1.14	2.06	1.14	1.78	4.22
84	.55	.60	.89	.60	.75	1.24	.65	.94	2.23

specific mortality differentials for females is similar to the pattern for males. The prominent differences seemed to be that, in 1935, individual Black females were even more disadvantaged than individual Black males, with a high relative risk ratio of 5.8 (unidimensional model) at age 40 and that the mortality ratios for ages 25 to 45 declined due to the relatively (to White females) greater reduction of maternal mortality for individual Black females. In addition, a larger degree of crossover was observed for Black females in 1935 and 1955 so that a k value of 0.5 was insufficient to remove the crossover in these years. By 1975, this situation had changed dramatically where,with k = 0.5, the relative risk at age 84 was 2.2.

Additionally, for the Black female cohort born between 1885 and 1900 standard individuals under both the unidimensional and bidimensional models appear to be at a greater disadvantage than the adjacent cohorts. The fact that this cohort pattern appears for both male and female stan-

dard individuals and for both models, but does not appear in the observed mortality rates, underscores the importance of adjusting for population heterogeneity *prior* to making population comparisons.

These results are based on the assumption that longevity characteristics in the population are gamma distributed with k = 0.5 or k = 1.0. However, if k is greater than 1.0, the same type of change in the relative risks would be noted, though to a lesser degree. If, for example, there are additional measurement factors to the ones represented in the Siegel (1974), Coale and Zelnik (1963), and Coale and Rives (1973) estimates for the Black and White populations, they could serve to decrease the degree of crossover (or reduce it to convergence), thus possibly permitting the crossover to be eliminated with a value of $k > 1.0$.

One strong assumption that has been employed in the above discussion is that Blacks and Whites have the same distribution of longevity characteristics at birth. The assumption could, of course, be relaxed so that k might be equal to 0.5 for Blacks (greater heterogeneity due to greater admixture of genetic elements) and 1.0 for Whites. By generalizing in this way it is possible to remove the convergence of the individual risk ratios at older ages observed in Tables 2 and 3 and replace it with a divergence. Conversely, if k were set to 0.5 for Whites and 1.0 for Blacks, this would accentuate the rate of convergence and crossover of the individual risk ratios in the two groups. Thus, in this case, with theoretical specifications of k consistent with the notion that a discrete number of dimensions underlie the heterogeneity in a population, the assumption of identical heterogeneity distributions seems justifiable. It is obvious that by using empirical procedures "optimal" (in the sense of preserving constant relative risks for the two groups) values of k could be generated. However, it will not generally be possible to relate these values of k directly to a discrete number of independent dimensions. Note that if any convergence exists in the observed mortality rates then the hypothesis of differential selection, and a model such as presented here, would have to be entertained as an explanation. The alternative to such an assumption would be that individuals are homogeneous and that the cohort life tables accurately reflect the age trajectory of risks for all individuals, i.e., that differentials in life chances are functions solely of environmental "shocks." One implication of the assumption of homogeneity would be that the biological mechanisms thought to determine longevity would have to behave in a more complicated fashion over age. For example, in the case of a specific disease, cancer, which exhibits a slowing, at advanced ages, of the rate of increase of mortality risks, the rejection of a selection

mechanism as an explanation would require a biological model indicating that individuals became increasingly resistant to cancer past some advanced "threshold" age.

Discussion

The major declines in mortality over most of this century have resulted in an approximate doubling of life expectancy in the U.S. over that of the last century. The concomitant shift of the majority of deaths to the advanced ages has led to research on all aspects of aging and age related changes in the biology of the human organism. In this research age specific population mortality rates often are used as a proxy measure for individual aging changes (Taylor, 1962; Jones, 1962; Brown and Forbes, 1974; Strehler, 1977). Alternately, mortality rates for different populations are used to index the relative health status and longevity potential of individuals. Thus, the utility of age-specific mortality rates in the analysis of a number of important facets of aging research has been well established. In this paper procedures were presented which correct an important misspecification of the methods used to analyze mortality data, i.e., as listed above, the use of cohort (i.e., population level) mortality rates when the true focus of analysis is the probability of death for individuals within the population. We have shown that individual differences in longevity endowment will result in a divergence between the increase with age of the cohort mortality rates and the age increase in the probabilities of death for individuals within the cohort. This divergence, due to the earlier selection of the less "robust" population members, implies that individuals age "faster" than their cohorts. This phenomenon has important theoretical implications for the study of aging and human survival, and practical implications for epidemiologic studies of the elderly population where one is studying disease in a highly selected group.

Though the effects of heterogeneity have been presented in terms of human populations, it is clear that the concepts presented in this paper may be generalized to include other types of populations. For example, a cell culture may be considered as a population of cells in a laboratory experiment where the effects of application of specific chemical substances are being studied. The timing of the occurrences of certain cell transformations is recorded for each transformed cell to generate, say, a cellular "dose/response" function. However, if the cells are in fact heterogeneous with respect to risk of transformation (analogous to force of mortality), then the average dose/response level will be a biased repre-

sentation of the dose/response level of cells at the "standard cell" (analogous to standard individual) *susceptibility* level. Furthermore, as in the study of human mortality at advanced ages, the bias will be particularly severe in those experiments where a large proportion of the cells undergoes such a transformation.

Acknowledgements

The research reported in this paper was supported by N.I.A. grant No. AG01159-03. The authors wish to thank Drs. J. W. Vaupel and M. A. Woodbury for comments which greatly strengthened this paper.

Received: 31 January 1980.
Revision received 30 May 1980.

Literature Cited

Brown, K. S. and W. F. Forbes 1974 A mathematical model of aging processes. J. Gerontology, 29: 46-51.

Chiang, C. L. 1968 Introduction to stochastic processes in biostatistics. Wiley, New York.

Coale, A. J. and N. Rives 1973 A statistical reconstruction of the Black population of the United States 1880-1970. Population Index, 39: 3-36.

Coale, A. J. and M. Zelnik 1963 New estimates of fertility and population in the United States. Princeton University Press, Princeton.

Evans, E. 1962 The economics of American Negro slavery, 1830-1860. *In:* Aspects of labor economics. Princeton University Press, Princeton.

Jacobson, P. H. 1957 An estimate of the expectation of life in the United States in 1850. Milbank Memorial Fund Quarterly, 35: 197-201.

Jones, H. B. 1962 Mechanisms of aging suggested from study of altered death risks. *In:* Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. IV, Biology and problems of health (J. Neyman, ed.). University of California Press, Berkeley.

Keyfitz, N. 1977 What difference would it make if cancer were eradicated? An examination of the Taeuber paradox. Demography, 14: 411-418.

_____ 1979 Information and allocation: Two uses of the 1980 census. Am. Statistician, 33: 45-50.

Kitagawa, I. M. and P. M. Hauser 1973 Differential mortality in the United States: A study in socioeconomic epidemiology. Harvard University Press, Cambridge, Mass.

Lew, E. A. and F. Seltzer 1970 Uses of the life table in public health. Milbank Memorial Fund Quarterly, 48: 15-37.

Manton, K. G., S. S. Poss and S. Wing 1979a The Black/White mortality crossover: Investigation from the perspective of the components of aging. Gerontologist, 19: 291-300.

MANTON, K. G. AND E. STALLARD 1979 Maximum likelihood estimation of a stochastic compartment model of lung cancer latency: Lung cancer mortality among White females in the U.S. Computers and Biomedical Research, 12: 313-325.

MANTON, K. G. AND E. STALLARD 1980 A two-disease model of female breast cancer: Mortality in 1969 among White females in the United States. J. Natl. Cancer Inst. 64: 9-16.

MANTON, K. G., E. STALLARD AND J. W. VAUPEL 1980 Methods for comparing the mortality experience of heterogeneous populations. Demography, in review.

MANTON, K. G., M. A. WOODBURY AND E. STALLARD 1979b Analysis of the components of CHD risk in the Framingham study: New multivariate procedures for the analysis of chronic disease development. Computers and Biomedical Research, 12: 109-123.

NAM, C. B. AND K. A. OCKAY 1977 Factors contributing to the mortality crossover pattern. XVIII General Conference of the International Union for Scientific Study of Population, Mexico City.

NATIONAL CENTER FOR HEALTH STATISTICS 1972 Cohort mortality and survivorship: United States death registration states, 1900-1968. DHEW Pub. No. (HSM) 73-1400.

PRESTON, S., N. KEYFITZ AND R. SCHOEN 1972 Causes of death: Life tables for national populations. Seminar Press, New York.

RIVES, N. W. 1977 The effect of census errors on life table estimates of Black mortality. Am. J. Public Health, 67: 867-868.

ROSENBERG, B., G. KEMENY, L. G. SMITH, I. D. SKURNICK AND M. J. BANDURSKI 1973 The kinetics and thermodynamics of death in multicellular organisms. Mechanisms of Aging and Development, 2: 275-288.

SIEGEL, J. S. 1974 Estimates of coverage of the population by sex, race, and age in the 1970 census. Demography, 11: 1-23.

SIMMS, H. 1942 The use of measurable cause of death (hemorrhage) for the evaluation of aging. J. General Physiology, 26: 169-178.

_____ 1946 Logarithmic increase in mortality as a manifestation of aging. J. Gerontology, 1: 13-26.

SPIEGELMAN, M. 1968 The curve of human mortality. Proceedings of Seminars of Duke University Council on Aging and Human Development, pp. 286-298.

STREHLER, B. L. 1977 Time, cells, and aging. Academic Press, New York.

TAYLOR, W. F. 1962 On the methodology of studying aging in humans. *In:* Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. IV, Biology and problems of health J. Neyman, (ed.). University of California Press, Berkeley.

THORNTON, R. G. AND C. B. NAM 1972 The lower mortality rates of non-Whites at the older ages: An enigma in demographic analysis. Research Reports in Social Science, 11: 1-8.

VAUPEL, J. W., K. G. MANTON AND E. STALLARD 1979a The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography, 16: 439-454.

VAPUEL, J. W., K. G. MANTON AND E. STALLARD 1979b What demographic difference does it make that the frail die first?: A model of mortality and some suggestive results based on Swedish mortality since 1778. Center for the Study of Health Policy, Institute of Policy Sciences and Public Affairs, Working Paper No. 1079, Duke University, Durham, N.C.