State Center for Health Statistics
P.O. Box 29538 • Raleigh, NC 27626-0538
919/733-4728

STATISTICAL PRIMER

No. 13 August 1998

Age-Adjusted Death Rates

by

Paul A. Buescher

Introduction

Mortality or death rates are often used as measures of health status for a population. Population-based incidence or morbidity data are available in North Carolina in a few areas such as cancer and certain communicable diseases, but for most chronic diseases we know only how many people died from the disease and not how many are living with it. Given the importance of data from death certificates in measuring the health of populations, it is important that valid comparisons of death rates are made. Many factors affect the risk of death, including age, race, gender, occupation, education, and income. By far the strongest of these factors affecting the risk of death is age. Populations often differ in age composition. A "young" population has a higher proportion of persons in the younger age groups, while an "old" population has a higher proportion in the older age groups. Therefore, it is often important to control for differences among the age distributions of populations when making comparisons among death rates to assess the relative risk of death. This Statistical Primer describes how age-adjusted death rates are calculated and discusses some related issues.

The methods for adjusting death rates for age that are shown here could also be applied to other characteristics of a population, such as income or gender, if it were considered desirable to adjust for such characteristics before comparing death rates. Also, disease incidence rates, birth rates, or other types of rates could be adjusted for age, or other factors, using the general approach presented here.

Crude and Age-specific Death Rates

A crude or unadjusted death rate is simply the number of deaths divided by the population at risk, often multiplied by some constant so that the result is not a fraction. For example, for Hertford County, North Carolina during the period 1991 through 1995, there were 1,336 deaths to residents of the county. To get an annualized death rate for this five-year period, the estimated mid-year resident population is summed over the five years. For Hertford County, the sum of the population of those five years is 112,419. The crude death rate is 1,336 divided by 112,419 which equals .01188. This is the average annual proportion who died during the period 1991-95 (slightly more than one percent). When multiplied by 1,000, which is a common multiplier for a death rate for all causes of death, the rate is 11.9 deaths per 1,000 population per year (see last row of Table 1). For death rates for specific causes of death, a multiplier of 100,000 is often used so that the rate is not less than 1.0. For smaller geographic areas or when using cause-specific death rates, it is often desirable to calculate multi-year death rates to decrease random variation in the rates due to small numbers of deaths in a single year. These multi-year death rates are essentially average annual rates.

Table 1
Age-Specific Death Rates for All Causes of Death
Hertford County, North Carolina Residents
1991-1995 Combined

	1	2	3	4	5
	Number of	Sum of	Percentage	Age-Specific Death Rates	
Age Group	1991-95 Deaths	1991-95 Population	of Population by Age	Proportion Who Died	Per 1,000 Population
0-4	30	8,150	7.3	.00368	3.68
5-14	2	17,109	15.2	.00012	0.12
15-24	24	16,601	14.8	.00145	1.45
25-34	34	14,872	13.2	.00229	2.29
35-44	59	16,199	14.4	.00364	3.64
45-54	85	12,381	11.0	.00687	6.87
55-64	147	10,277	9.2	.01430	14.30
65-74	305	9,370	8.3	.03255	32.55
75-84	406	5,631	5.0	.07210	72.10
85+	244	1,829	1.6	.13341	133.41
Total	1,336	112,419	100.0	.01188	11.9

The crude death rate is a good measure of the overall magnitude of mortality in a population. If a population is old and has a high mortality rate as a result, then the crude rate is useful information for some purposes, such as planning for the delivery of health care services.

An age-specific death rate is simply a crude death rate for a specific age group. One can also calculate rates specific for race, gender, or other factors. Table 1 shows age-specific death rates for Hertford County residents for the period 1991-95, with ten commonly-used age groupings. Note that the death rate for ages 0-4 is substantially higher than the death rates for the other younger age groups (primarily due to a high death rate during infancy); only at ages 45-54 does the death rate exceed that for ages 0-4. It can be immediately seen that the death rates are many times higher in the oldest age groups. Therefore, a geographic area or demographic group with an older population will automatically have a higher overall death rate just because of the age distribution. The main purpose of age-adjusting death rates is to control for differences in the age distribution of various populations before making mortality comparisons.

For some causes of death, such as injuries and AIDS, older persons do not have the highest death rates. But even in these cases it is important to standardize for age when comparing death rates across different populations, since some populations may have a higher proportion of persons in the age groups with the highest death rates.

Another Statistical Primer by the State Center for Health Statistics discusses the issue of random error in vital rates and presents formulas for quantifying this error and calculating confidence intervals around the measured rates.[1] Those formulas are applicable to the crude and age-specific rates presented here, and to any simple or unadjusted rate. Random error may be substantial when a rate or percentage has a small number of events in the numerator (e.g. less than 20).

Age-adjusted Death Rates

Direct Method

The direct method of age adjustment is frequently used to compare the death rates of different populations, by controlling for differences in age distribution. The age-specific death rates of the population of interest (sometimes called the "study" population) are applied to the age distribution of a "standard" population in order to calculate "expected deaths." These are the deaths that would occur in the standard population IF the age-specific rates of the study population were in operation. These expected deaths for each age group are then summed and divided by the total standard population to arrive at the age-adjusted death rate. Stated another way, this is the death rate that the study population would have IF it had the same age distribution as the standard population.

Table 2 provides an example. The age-specific death rates for all causes of death for Hertford County are applied to the 1980 North Carolina population by age, which is used as the standard. (Any population could be used as the standard; the 1980 North Carolina population was chosen somewhat arbitrarily for purposes of illustration.) To generate the expected deaths in column 4, the rates shown in column 1 are converted to a proportion by moving the decimal point three places to the left and then multiplied by the standard population groups by age in column 2. The total expected deaths are then divided by the total standard population and the result multiplied by 1,000 to yield an age-adjusted death rate for Hertford County of 8.7. Usually it would not be necessary to show the age-specific death rates to two decimal places (false precision), but in this case the extra digits are needed to get a more accurate estimate of the number of expected deaths.

Table 2
Age-Adjustment by the Direct Method
Hertford County, North Carolina Residents
1991-1995 Combined; All Causes of Death

Age Group	1 Hertford County 1991-95 Age-Specific Death Rate (per 1,000)	2 1980 N.C. Population (Standard)	3 Percentage of Standard Population By Age	4 Expected Deaths
0-4	3.68	404,560	6.9	1,489
5-14	0.12	927,836	15.7	111
15-24	1.45	1,144,204	19.4	1,659
25-34	2.29	968,215	16.4	2,217
35-44	3.64	689,838	11.7	2,511
45-54	6.87	601,866	10.2	4,135
55-64	14.30	552,494	9.4	7,901
65-74	32.55	389,244	6.6	12,670
75-84	72.10	172,408	2.9	12,431
85+	133.41	45,956	0.8	6,131
Total	11.9 (Crude Rate)	5,896,621	100.0	51,255

$$\text{Age-Adjusted Death Rate} = (51{,}255 \div 5{,}896{,}621) \times 1{,}000$$
$$= 8.7$$

This adjusted death rate is considerably lower than the crude death rate of 11.9. This is mainly because the percentages in the age groups 65 and older are substantially lower in the 1980 North Carolina standard population (column 3 of Table 2) than the same percentages in the 1991-95 Hertford County population (column 3 of Table 1). When the Hertford County age-specific death rates are adjusted to a younger standard population, the overall adjusted rate is lower.

The crude death rate for North Carolina for the 1991-95 period for all causes of death was 8.9, compared to the crude rate of 11.9 for Hertford County. The 1991-95 North Carolina death rate adjusted to the 1980 North Carolina age distribution is 7.4, compared to the age-adjusted rate of 8.7 for Hertford County. The difference in the crude rates between North Carolina and Hertford County is larger partly because Hertford County had an older population. The fact that the age-adjusted rate for Hertford County is still higher than that for North Carolina suggests that the 1991-95 age-specific death rates for Hertford County were generally higher than those for the state in 1991-95.

Ten age groups are often used for age adjustment of death rates. This provides enough detail to capture differences in the age distributions of the populations that are being compared, but not so many age categories that the data are "spread too thin." For many years, the State Center for Health Statistics used 18 five-year age groups for age adjustment, but during the 1980's changed to ten age groups because the 18 categories often resulted in the numerators of the age-specific rates being very small, leading to unstable rates.

An alternate formula for computing the age-adjusted death rate by the direct method is simply to sum the products of the age-specific death rate and the proportion of the standard population in that age group across all ten age groups. This weighted sum is represented by the following formula:

$$\text{Age-adjusted death rate} = \sum_{i=1}^{10} w_i \, p_i$$

where p_i is the age-specific mortality rate for age group i and w_i (or the weight) is the proportion of the standard population in age group i (move the decimal point of the percentages in column 3 of Table 2 two places to the left). The crude death rate can also be expressed as a weighted sum of the age-specific death rates and the proportions of the population by age, where the proportions in this case are simply the proportions of the study population itself in each age group (rather than the standard population). Try to reproduce the crude and age-adjusted death rates in Tables 1 and 2 using this weighted sum method! Any minor differences are due to rounding.

An age-adjusted death rate is a summary measure that condenses a lot of information into one figure. Where feasible, it is always desirable to inspect the age-specific death rates of the populations being compared. This additional detail often provides further insights into the nature of the mortality differences between the populations.

Indirect Method

The indirect method of age-adjustment applies the age-specific death rates of the standard population to the age distribution of the study population in order to generate expected deaths in the study population. These are the deaths that would occur in the study population IF the age-specific death rates in the standard population were in operation. This method may be used in situations where the numbers of deaths in each age group in the study population are too small to calculate stable age-specific rates. Also, this method is often used in developing countries or other areas where there is no information available on age-specific deaths for the study population, but there is such information available for a national or standard population. The expected deaths are then summed across the age groups and compared to the actual or observed number of deaths for the study population. This ratio of observed/expected deaths is often referred to as the **standardized mortality ratio,** or SMR. A ratio greater than 1.0 indicates higher mortality in the study population compared to the standard population (controlling for age distribution), while a ratio less than 1.0 indicates lower mortality in the study population. The SMR controls for age distribution since both the observed and expected deaths are based on the age distribution of the study population. Multiplying the SMR times the crude death rate in the standard population produces the **indirectly standardized death rate** for the study population.

Table 3 presents an example. The age-specific death rates in the 1993 North Carolina standard population (column 2), after moving the decimals three places to the left, are multiplied by the 1991-95 Hertford County population in column 1 to produce the expected deaths in column 3. These expected deaths by age may be compared to the actual 1991-95 deaths by age in Hertford county, shown in column 1 of Table 1. Dividing the 1,336 total deaths observed in Hertford County during 1991-95 by the 1,187 total expected deaths results in an SMR of 1.13. This indicates that the 1991-95 death rate in Hertford County was on the whole higher than the rate in the 1993 North Carolina standard population, controlling for age. Multiplying the crude death rate in the standard population of 9.0 by 1.13 gives an indirectly standardized death rate for Hertford County of 10.2. It is usually desirable to use a standard population that is close to the same year(s) as the data for the study population, to avoid differences between the observed and expected deaths due to changing (often declining) age-specific death rates over time. This is why the 1993 (midpoint) North Carolina standard was used in this example.

Table 3
Age-Adjustment by the Indirect Method
Hertford County, North Carolina Residents
1991-1995 Combined; All Causes of Death

	1	**2**	**3**
Age Group	**Hertford County 1991-95 Population**	**Age-Specific Death Rates in 1993 North Carolina Standard Population (per 1,000)**	**Expected Deaths in Hertford County**
0-4	8,150	2.44	20
5-14	17,109	0.25	4
15-24	16,601	0.98	16
25-34	14,872	1.53	23
35-44	16,199	2.55	41
45-54	12,381	5.03	62
55-64	10,277	12.41	128
65-74	9,370	28.48	267
75-84	5,631	63.19	356
85+	1,829	147.85	270
Total	112,419	9.0 (Crude Rate in Standard)	1,187

Standardized Mortality Ratio (SMR) $= 1{,}336 \div 1{,}187$

$= 1.13$

Indirectly Standardized Mortality Rate $= 1.13 \times 9.0$

$= 10.2$

Comparison of the Direct and Indirect Methods

The direct method of adjustment is generally preferred where the numbers of deaths in the study population are large enough to produce stable age-specific death rates. A big advantage of the direct method is that the adjusted rates of a number of different study populations (e.g. all counties in North Carolina) can be directly compared to each other if they are all adjusted to the same standard population. This allows mortality comparisons assuming a constant age distribution across all of the study populations. The indirect method is often used if mortality rates by age cannot be calculated for the study population, or if the numbers of deaths in the study population are too small to produce stable age-specific death rates. A problem with the indirect method is that the adjusted rate for the study population can be compared only with the rate of the standard population. Different study populations cannot be compared to each other since the adjusted rates are not based on a common age distribution. In other words, differences in the rates may still be due to differences in age distribution, since the rates are adjusted to the age distribution of each particular study population rather than to a common standard.

Issues in the Choice of the Standard Population

An age-adjusted death rate is a hypothetical index, designed to facilitate comparisons among populations, rather than a true measure of risk. An age-adjusted death rate (by the direct method) answers the question: What would the death rate in a study population be IF that population had the same age distribution as the standard population? So in theory any population distribution can be used as the standard; it is only a set of weights applied to the age-specific death rates. The choice of the standard population will not usually have a great effect on the **relative** levels of the age-adjusted rates that are being compared. But it is important to remember that age-adjusted death rates can be compared to each other **only** if they are adjusted to the same standard.

For many years the National Center for Health Statistics has used the 1940 United States population as the standard for age-adjusting death rates. Converted to a population of one million with the same proportions at each age as in the 1940 population, this standard is sometimes referred to as the "standard million." An advantage of consistently using this same standard population is that it promotes comparisons of age-adjusted death rates, especially in looking at trends over time from 1940 to the present. A disadvantage of using this standard is that the size of the adjusted rate is often much different from the size of the crude rate in the study population.

Take the example of heart disease mortality in North Carolina. In 1993 the crude heart disease death rate was 277.0 per 100,000 population. Age-adjusted to the 1940 United States population standard, the 1993 heart disease death rate for North Carolina was 151.4. The 1993 United States heart disease death rate, age-adjusted to the 1940 United States standard, was 145.3. This shows that the 1993 heart disease death rate in North Carolina was slightly higher than that in the United States, after adjustment for differences in age distribution. However, the North Carolina adjusted rate of 151.4 is much lower than the crude rate in 1993 of 277.0 (i.e. it is not an accurate measure of the risk of death from heart disease in 1993). This is primarily due to the following: a) the 1940 United States population is much younger than the 1993 North Carolina population, and b) heart disease death rates are much higher in the older age groups. So standardizing to a much younger population results in a much lower age-adjusted death rate. In recognition of this problem, the

National Center for Health Statistics has proposed to begin using the year 2000 United States population as the recommended standard population. This will mean that the age-adjusted death rates will generally be much more similar in size to contemporary crude death rates. However, it will also mean that time series comparisons of age-adjusted death rates will have to be re-computed using the new standard, and that rates adjusted to the 1940 standard cannot be compared to rates adjusted to the new standard.

For a number of years the State Center for Health Statistics used the current-year North Carolina population as the standard for computing adjusted rates in the annual publication *Leading Causes of Death*. This was not a problem as long as comparisons of adjusted rates were made within the current year. It did, however, preclude comparisons of adjusted rates over time since the standard population was changed every year. To address this problem, beginning in the late 1980s, the State Center for Health Statistics started using the 1980 North Carolina population as the standard for adjustment in each annual publication. This made it possible to compare adjusted death rates for different years. But as the North Carolina population has become older over time, the current-year crude death rates have generally become increasingly different in size from the adjusted death rates. In the 1996 edition of *Leading Causes of Death* we changed from computing age-race-sex-adjusted death rates to computing age-adjusted rates. (The reasons for this are discussed below.) Since the adjusted rates for 1996 forward would not be comparable to previously published rates anyway, we took this as an opportunity to update the standard population to the 1990 North Carolina population by age, which has made the crude and adjusted death rates less different in size than when the 1980 standard population was used.

Taking the example of Hertford County, the 1991-95 crude death rate was 11.9 and the age-adjusted death rate using the 1980 North Carolina population as the standard was 8.7. Using the 1990 North Carolina population as the standard, the 1991-95 age-adjusted death rate for Hertford County for all causes of death is 10.0. The main reason that this latter adjusted rate is higher is because the 1990 North Carolina population used as the standard is older than the 1980 North Carolina population.

One should be especially careful when assessing trends over time using age-adjusted death rates. It is essential that rates for different years be adjusted to the same standard population before making comparisons. Also, if the standard population is very different from the populations of the years being compared (as is often the case when using the 1940 U.S. standard), changes in the adjusted rates over time may not be an accurate reflection of the actual changes in the risk of death.

Errors of Adjusted Rates

A detailed discussion of random errors in age-adjusted death rates is beyond the scope of this paper. The reader should refer to the Statistical Primer cited in reference 1 for information on the general concepts of random errors in rates, confidence intervals, and determining if the difference between two rates is statistically significant. Using the terminology in that paper, a 95% confidence interval around a proportion can be computed as

$$p \pm 1.96 \sqrt{pq/n}$$

The $\sqrt{pq/n}$ is commonly known as the **standard error** of the proportion. In this case a death rate is treated as the proportion (p) who died during the time period of interest. If the proportion who died is small, then q (which is 1 – p) will be very close to 1.0 and the formula becomes $\sqrt{p/n}$, where n is the population or the denominator of the proportion.

We saw from the discussion above that an age-adjusted death rate (by the direct method) is a weighted sum of the age-specific death rates. Using ten age groups, the formula for the standard error of an age-adjusted death rate is as follows:

$$\sqrt{\sum_{i=1}^{10} w_i^2 (p_i/n_i)}$$

This is the square root of the sum across the ten age groups of the square of the weight times the standard error of the age-specific death rate squared. Remember that the weight is simply the proportion of the standard population in age group i. This standard error of the age-adjusted death rate times 1.96 is the half-width of the 95% confidence interval around the age-adjusted rate.

This is a very brief discussion of a lengthy topic. For questions or assistance, contact the author.

Issues in Adjusting for Race and Gender

For many years, the death rates in the *Leading Causes of Death* publication of the State Center for Health Statistics were adjusted simultaneously for age, race, and gender. This was done for five-year death rates for specific causes of death, by county of residence. With 40 age-race-gender-specific rates being computed for the adjustment process (10 age groups x 2 race groups: white/minority x 2 gender groups), it became apparent that the data were being spread too thin. A particular problem was in the western North Carolina counties, which generally have very small minority populations, and there were also problems in other counties with a small population. Just one or two deaths in a small population group could result in a very high age-race-gender-specific rate, which would severely inflate the adjusted death rate. If this rate were applied to the appropriate age-race-gender group of the standard population, a very large number of expected deaths could result and the adjusted rate would be extremely high. On the other hand, zero deaths in several population groups may result in a very low age-race-gender-adjusted rate.

For example, in the 1995 *Leading Causes of Death* publication of the State Center for Health Statistics, Macon County was shown to have a 1991-95 age-race-adjusted death rate for female breast cancer of 64.6 (per 100,000 population) compared to the age-race-adjusted rate for North Carolina of 28.3. (Breast cancer death rates are already gender-specific and do not need to be adjusted for gender; in this case there were 20 age-race-specific death rates.) The 1980 North Carolina population was used as the standard for adjustment. Further investigation showed that this very high adjusted rate of 64.6 was due to two breast cancer deaths in a very small minority female population group (ages 55-64). When the 1991-95 breast cancer death rate was adjusted only for age, the Macon County rate was 30.4 compared to a North Carolina age-adjusted rate of 25.6. The crude 1991-95 breast cancer death rates for Macon County and North Carolina were 56.5 and 32.1, respectively. Adjusting only for age avoids the bias due to small numbers, and the similar size of the county and state age-adjusted rates appropriately shows that the elevated crude death rate for breast cancer in Macon County was due mainly to an older population distribution in Macon County compared to North Carolina as a whole.

Age generally has a much stronger impact on mortality than race or gender, and therefore is the most important factor to adjust for. Also, there are other questions about adjusting for race. Age differences in mortality are not easily modified. Racial differences in mortality, on the other hand, are often due to factors that can be changed through public health, medical care, or socioeconomic interventions. Adjusting for race may cover up the fact that certain geographic areas, for example, have higher mortality because they have a larger percentage of minority populations (who often have higher death rates). In many cases we would want to target these areas for public health interventions and not produce statistics that adjust downward a higher level of mortality that is potentially modifiable.

For example, it was shown above that the 1991-95 age-adjusted death rate for Hertford County for all causes of death (using the 1980 North Carolina population as the standard) was 8.7 per 1,000, compared to the 1991-95 age-adjusted death rate for North Carolina (using the same standard) of 7.4. If the 1991-95 Hertford County death rate is age-race-gender-adjusted to the 1980 North Carolina population, the resulting adjusted rate is 7.7 and the comparable age-race-gender-adjusted rate for North Carolina is 7.4. This shows that the age-race-gender-specific death rates in Hertford County were similar to those in the state as a whole, but does not reveal that Hertford County had a higher overall mortality rate than the state due to a higher percentage of minorities. Minorities (primarily African Americans) are approximately 61 percent of the total population in Hertford county, compared to 24 percent for the state as a whole.

Rather than adjusting for race, a better approach would be to examine racial differences in mortality by calculating race-specific death rates, perhaps adjusted for age. Minority populations often have a younger age distribution than whites. Adjusting for age usually results in relatively higher death rates for minorities, and larger differences between whites and minorities than when comparing crude death rates. In adjusting the death rates of different race (or race-gender) groups for age, it is important to use the same standard population (or set of age-specific weights) in all cases so that the adjusted rates will be directly comparable. At the county level in North Carolina, small numbers of deaths generally preclude calculating statistically reliable death rates for minority populations other than African Americans. For this reason, we usually calculate death rates for two broad racial groups: white and all minorities combined. In North Carolina as a whole, African Americans comprise more than 90% of the minority population. (Hispanics are considered an ethnic rather than a racial group and most Hispanics are counted within the white racial group). Another problem with calculating death rates for specific minority sub-groups is the lack of accurate population estimates to use in the denominators of the death rates.

For the reasons discussed above, beginning with the 1996 *Leading Causes of Mortality* publication, we have adjusted all death rates only for age, using the 1990 North Carolina population in ten age groups as the standard.

Readers with questions or comments about this Statistical Primer may contact Paul Buescher at (919) 715-4478 or through e-mail at paul_buescher@mail.ehnr.state.nc.us. Further reading on the topic of adjusted rates may by found in references 2, 3, and 4.

Reference

1. Buescher PA. Problems with rates based on small numbers. *Statistical Primer*, No. 12, State Center for Health Statistics, April 1997.

2. Rothman KJ. *Modern Epidemiology* (Chapter 5). Little, Brown and Co., Boston/Toronto, 1986.

3. Curtin LR, Klein RJ. Direct standardization (age-adjusted death rates). *Healthy People 2000 Statistical Notes*, No. 6 (revised). National Center for Health Statistics, 1995.

4. Anderson RN, Rosenberg HR. Age standardization of death rates: implementation of the year 2000 standard. *Monthly Vital Statistics Report* (forthcoming). National Center for Health Statistics, 1998.

Department of Health and Human Services
State Center for Health Statistics
P. O. Box 29538
Raleigh, N.C. 27626-0538

BULK RATE
U.S. Postage
PAID
Raleigh, N.C. 27626-0538
Permit No. 1862

1,900 copies of this public document were printed at a cost of $272.94 or 14 ¢ per copy.