Number 6—Revised
March 1995

Revised to clarify computations in tables III and V

HEALTHY PEOPLE 2000 Statistical Notes

From the CENTERS FOR DISEASE CONTROL AND PREVENTION/National Center for Health Statistics

Direct Standardization (Age-Adjusted Death Rates)

Lester R. Curtin, Ph.D. and Richard J. Klein, M.P.H.

Introduction

Most population-based mortality objectives and subobjectives in *Healthy People 2000* are tracked using age-adjusted rates from the National Vital Statistics System (appendix table I). The exceptions are deaths from alcohol-related motor vehicle crashes, all motor vehicle crashes, and work-related injuries (objectives 4.1, 9.3, and 10.1), which are monitored with crude death rates from other data systems. In addition, objectives that refer to specific age groups are tracked with age-specific rather than age-adjusted rates.

Although the age-adjusted death rate (ADR) is one of the most frequently used indexes of mortality, there is often confusion concerning the basic concepts of its construction, use, and interpretation. Some of the persistent issues include the appropriateness of the ADR as a summary measure, the validity of comparisons between ADRs, the method of calculation, and the appropriateness of alternate summary measures.

Why use age-adjusted death rates?

The total number of health events (for example, the number of deaths) occurring in a population is useful for determining the magnitude of a public health problem. However, the absolute number of deaths is seldom useful for comparisons between population groups (for example, comparing males and females) or for comparing trends. Assuming equal risk, a larger population group will tend to generate more events (deaths) than a smaller group simply because of its size. Therefore, to compare relative differences in mortality among population groups, or for a given population group over time, the number of deaths must be related to the ‘‘population at risk’’ of dying to produce death rates. The population of interest may be the entire population of an area or a population subgroup (for example, people in a certain age group).

The simplest death rate is the crude death rate (CDR), defined as the total number of deaths divided by the midyear population. CDRs are usually expressed as a rate per 1,000 or 100,000 population. CDRs for individual age cohorts, called age-specific death rates (ASDRs), are the ratio of the number of deaths in a given age group to the population of that age group, again usually expressed per 1,000 or 100,000 population.

To compare the relative health of population groups or to assess change in mortality over time, two criteria must be considered. First, rates should relate the number of events to the population at risk. Second, because many health outcomes vary by age, the effect of the population’s age distribution must be taken into account.

Although it does relate the number of events to the population, the crude rate does not take into account the age distribution of the population. As such, it is not an appropriate measure for comparing differences between population groups or for assessing change in mortality over time. Because death rates for most diseases generally increase with age, a population group with a relatively

Table A. Crude death rate comparison

Age	Community A Deaths	Community A Population	Community A Rate[1]	Community B Deaths	Community B Population	Community B Rate[1]
0–34 years	20	1,000	20	180	6,000	30
35–64 years	120	3,000	40	150	3,000	50
65 years and over	360	6,000	60	70	1,000	70
Total	500	10,000	50	400	10,000	40

[1]Per 1,000 population.

Table B. Age-adjusted death rate calculation

Age	Standard population	Community A Rate[1]	Community A Rate × population	Community B Rate	Community B Rate × population
0–34 years	3,000	20	60	30	90
35–64 years	3,000	40	120	50	150
65 years and over	4,000	60	240	70	280
Total	10,000	42	420	52	520

[1]Per 1,000 population.

younger age distribution will tend to have fewer total deaths from a given disease than a comparably sized population group with an older age distribution. Similarly, even if the age-specific risks of dying for a group remain unchanged between two time points, the number of deaths will increase as the population ages.

As an alternative to crude rates, ASDRs can be used. The most comprehensive and reliable method of comparing death rates over time or between different population groups is to compare individual ASDRs for all age groups of interest. However, this method often requires an extremely large number of comparisons and tends to overwhelm both the investigator and the intended audience.

Because the crude death rate is not appropriate and ASDRs provide too much detailed information, a summary measure that controls for a population's age distribution is needed. A commonly used measure is the ADR (1,2).

Age-adjusted rates were developed in 1841 for the analysis of mortality data (3). In the 19th century, mortality data provided the most useful, and often the only, measure of the health of a population. About that time, it was observed that a community could have ASDRs that were lower than the national average at each age interval, but, because the community's population was older, the overall CDR was higher for the community than for the Nation.

Table A presents a hypothetical comparison to illustrate this situation.

In each of the two comparably-sized communities in table A, the ASDRs increase with age; at each age the age-specific rates are higher for community B than for community A. Yet, the total (or crude) death rate is *lower* for community B. This occurs because community A is an older population; 60 percent of its citizens are 65 years and over. In contrast, only 10 percent of community B's population is in the oldest age group. Because the death rate is highest in the oldest age group, there are fewer total deaths in community B.

Direct standardization

There are two basic methods of standardization, or age-adjustment; both were introduced in the 19th century. These two methods have become known as the direct and indirect methods. (Indirect standardization is discussed in a later section.) When the direct standardization method is applied to ASDRs, the resultant summary index is called the ADR. Two assumptions are made when this index is computed for a population: The population's observed age-specific rates are assumed to be valid, and the age distribution of the population is assumed to be that of a standard, or reference, or population.

Table B illustrates the calculation of the ADR using the hypothetical data from table A. Specific computational formulae for the ADR are given in the technical appendix.

To calculate the ADR, the standard population and the age-specific death rate for each age interval are multiplied and these products are summed. In this example, the total for community A is 420. This sum is divided by the total standard population (10,000 in this case) to obtain the ADR. As with crude rates, the ADR is usually expressed in terms of a rate per 1,000 or per 100,000 population. Thus, the ADR for community A is 42 deaths per 1,000 population and the ADR for community B is 52 per 1,000. Note that, although the crude rate for community A was larger than that for community B, the ADR for community A is *smaller* than the ADR for community B. This is consistent with each of the age-specific rates for community A being smaller than those of community B.

Because of the method of computation, the age-adjusted rate is often interpreted as the hypothetical death rate that would have occurred if the observed age-specific rates were present in a population whose age distribution is that of the standard population. It is very important to realize that the ADR is an artificial measure whose absolute value has no intrinsic meaning. The ADR is useful for comparison purposes only, not to measure absolute magnitude. (To compare absolute magnitude, crude rates are used.) It is also important to note that in order to compare two age-adjusted rates, the *same* standard population must have been used.

Selection of a standard population

After the decision to use an ADR is made, the standard population must be selected. There are two basic types of standard populations, internal and external. Internal standards are created from the data to be used in the analysis; for example, the average age distribution of all populations to be compared. The use of an internal standard has certain statistical advantages for the ADR (4). However, if an internal standard is used, the results cannot be directly compared to other studies that use adjusted rates computed using a different standard population.

External standards are standard populations drawn from sources outside the analysis. For example, the National Center for Health Statistics (NCHS) typically uses a standard based on the 1940 United States population (5). This U.S. standard population is usually given in terms of a ‘‘standard million’’ in 10-year age groups. The U.S. standard million population is presented in appendix table I. A specific example of age-adjustment for one of the *Healthy People 2000* mortality objectives using the standard million and 10-year ASDR is given in appendix table II. The calculation of the variances of the age-adjusted rates in table II are shown in appendix table III.

NCHS publishes a large number of ADRs based on the U.S. standard population. This standard is also used to track those mortality objectives in *Healthy People 2000* and those Health Status Indicators which are monitored with age-adjusted rates. Several States use the same standard population in their publications. As long as the identical standard is used, ADRs from various national and State publications can be compared. But, if different standard populations are used to compute the ADR, then these ADRs are *not* comparable. Thus, there are considerable advantages to using the U.S. standard when computing State and local ADRs.

In recent years there have been discussions about whether the 1940 standard should be supplanted by a more contemporary ‘‘standard’’ that more closely reflects the U.S. population’s current (or future) age distribution (6,7). In considering this issue, it is important to remember that the actual magnitude of the ADR is beside the point. The ADR is an index number used for relative comparisons that should not be affected by the choice of a standard. (See section on ‘‘When not to adjust.’’) Although the *magnitude* of the ADRs may be greatly affected by the choice of a standard population, *relative mortality*, as measured by trends, race ratios, and sex ratios, is generally unaffected (8).

Despite this, controversy continues over which standard population to use when age adjusting death rates to measure temporal changes in cause-specific mortality. Examination of the issue shows that standard populations generally yield only a small effect on trend comparisons by cause of death (9). Thus, any standard population is adequate so long as comparison populations are not very ‘‘unusual’’ or ‘‘abnormal’’ with respect to the population under study (10). This means that the age distribution of the standard population should be somewhat similar to the population of interest.

Small number issues

One problem with ADR is that rates based on small numbers of deaths will exhibit a large amount of random variation. (See the technical appendix for more detail.) Therefore, if the number of deaths is small, mortality data should be aggregated over a number of years, or several small geographic areas must be combined into larger areas before computing the ADR (11). A very rough guideline is that there should be at least 25 total deaths over all age groups.

When not to adjust

The general consensus of the scientific literature is that, if it is appropriate to standardize, then the selection of the standard population should not affect relative comparisons. However, standardization is not appropriate when age-specific death rates in the populations being compared do not have a consistent relationship (12).

For example, evaluating trends in age-adjusted cancer death rates over time can be difficult because the ASDRs for younger ages have been decreasing while death rates at older ages are increasing. If a relatively young standard population is used, the trend in ADR may show a small increase or even a decrease; if a relatively older standard population is used, cancer mortality shows a much larger increase. Thus, using a more current (i.e., older) population than the 1940 standard, such as the 1990 U.S. population, as a standard population yields a much greater increase in the cancer mortality trend curve than does an analysis using the 1940 standard. Under these circumstances, a single summary measure is likely to be inappropriate for describing trends over time. Here, one should not use ADR, but should look at trends among ASDRs.

This does not mean it is inappropriate to publish ADR for cancer mortality. Within a defined time interval, e.g., 1990, geographic or race-sex comparisons may still be appropriate. It can be noted that when age-adjusted rates are computed using two distinct standards and the comparisons are different, then it is not appropriate to standardize in the first place. Again, only age-specific comparisons may be valid. Kitagawa illustrates this situation with an example of the mortality of white males living in metropolitan counties compared with those residing in nonmetropolitan counties in 1960 (13). In this case, ASDRs for white males under age 40 were lower in metropolitan counties than in nonmetropolitan counties. After age 40, the reverse was true. A summary index, such as the ADR, does not adequately describe the mortality differentials in the two groups. In cases such as these, the ADR is an imprecise indicator of mortality; the age-specific comparisons would be a better choice.

Indirect standardization

Because of concerns with the use of ADR, some mortality analysts prefer indirect standardized rates. Indirect standardization is generally thought of as an approximation to direct standardization. That is, when data needed to compute a direct measure (e.g., ASDRs) are not available, there may still be enough information to compute an indirectly standardized measure. However, indirect standardization has intrinsic value and should be considered on its own merits, not solely as an approximation to direct standardization (14,15).

For indirect standardization, a standard set of age-specific death rates are assumed to apply to the observed population. For example, the age-specific U.S. death rates could be applied to the age-specific local area population. This technique yields an ‘‘expected’’ number of deaths in a population, assuming the standard set of ASDRs was operating in the population.

An indirect adjusted death rate (IADR) can be computed from the expected number of deaths, but the index most often used is the ratio of the expected to the actual observed number of deaths. This ratio is called the standardized mortality ratio (SMR). The mathematics of indirect standardization and an example of the calculation of an SMR are given in the appendix.

Summary

This paper describes some of the issues related to the computation and use of age-adjusted rates. The following points were made:

- The age-adjusted rate is an index measure, the magnitude of which has no intrinsic value. It should be used for comparison purposes only.
- If it is appropriate to use age-adjustment, then the comparison should not be affected by the selection of a standard population. Conversely, if the comparison can be affected by the choice of a standard population, then it is not appropriate to age-adjust for that comparison.
- The standard population should not be ‘‘abnormal’’ or ‘‘unnatural’’ when compared to populations under study. Considering the amount of published material, there are advantages to using the U.S. standard population.
- Standardization is not a substitute for the examination of age-specific rates.

While standardization is most often applied to a series of age-specific death rates, direct or indirect standardization can also be applied to variables other than age. For example, infant mortality rates can be adjusted for birthweight distribution (16). Age-adjustment can also be used to monitor other measures of health at the local level, such as incidence or prevalence of disease.

Throughout the history of the ADR, the utility of the measure has often come into question. Any summary index, including direct or indirect standardization, will mask age-specific differences. Therefore, some authors have stressed the importance of comparing individual age-specific rates rather than attempting to summarize differences among the age-specific rates (17,18). A summary index, however, is more easily compared than an entire table of age-specific rates. Thus, the age-adjusted rate continues to be an integral part of the analysis of mortality trends and differentials. Accepting this, the need for a summary index must be balanced with recognition of the limitations of summary measures.

References

1. Shyrock HS, Siegel JS. The methods and materials of demography, vol 2. U.S. Bureau of the Census. Washington: U.S. Government Printing Office. 1971.
2. Spiegelman M. Introduction to demography. Rev. ed. Cambridge, MA. Harvard University Press. 1968.
3. Neison FGP. On a method recently proposed for conducting inquiries into the comparative sanatory condition of various districts. Journal of the Royal Statistical Society of London (now the Royal Statistical Society), vol 7, pp 40–68. 1844.
4. Kalton G. Standardization: A technique to control for extraneous variables. Applied Statistics, 17:118–36. 1968.
5. National Center for Health Statistics. Vital Statistics of the United States, 1989, vol II, mortality, part A. Washington: Public Health Service. 1992.
6. Johnson R. Proposed new standard population. Proceedings of the social statistics section, American Statistical Association, pp 176–81. 1990.
7. Feinleib MF, Zarate AO, eds. Reconsidering age adjustment procedures: Workshop proceedings. National Center for Health Statistics. Vital Health Stat 4(29). 1992.
8. Spiegelman M, Marks HH. Empirical testing of standards for the age adjustment of death rates by the direct method. Human Biology, 38:280–92. 1966.
9. Curtin LR, Maurer J, Rosenberg HM. On the selection of alternative standards for the age-adjusted death rate: Proceedings of the social statistics section, American Statistical Association, pp 218–23. 1980.
10. Wolfenden HH. On the theoretical and practical considerations underlying the direct and indirect standardization of death rates. Population Studies, 16:188–90. 1962.
11. Kleinman JC. Age-adjusted mortality indices for small areas: Applications to health planning. American Journal of Public Health, 67:834–40. 1977.
12. Fleiss JL. Statistical methods for rates and proportions. John Wiley and Sons, New York. 1973.
13. Kitagawa EM. Theoretical considerations in the selection of a mortality index and some empirical comparison. Human Biology, 38:293–308. 1966.
14. Inskip H, Beral V, Fraser P. Methods for age-adjustment of rates. Statistics in Medicine 2:455–66. 1983.
15. Tukey JW. Statistical mapping: What should and should not be plotted. Proceedings of the 1976 workshop on automated cartography and epidemiology. National Center for Health Statistics. DHEW (PHS) 79–1254. 1979.
16. Foster JE, Kleinman JC. Adjusting neonatal mortality rates for birthweight. National Center for Health Statistics. Vital Health Stat 2(94). 1982.
17. Woosley TD. Adjusted death rates and other indices of mortality. Chapter 4 in Vital Statistics Rates in the United States, 1900–1940.Washington: U.S. Government Printing Office. 1959.
18. Elveback LR. Discussion of indexes of mortality and tests of their statistical significance. Human Biology 38:322–24. 1966.
19. Chiang CL. Standard error of the age-adjusted death rate. U.S. Department of Health, Education, and Welfare: Vital Statistics Special Reports 47:271–85. 1961.
20. Keyfitz N. Sampling variance of standardized mortality rates. Human Biology 38:309–17. 1966.

Appendix

Table I. *Healthy People 2000* mortality objectives

Objective number	*Cause of death*	*ICD–9 identifying codes*
1.1	Coronary heart disease	410–414, 402, 429.2
1.1a	[Blacks]	
2.1	See 1.1	
2.1a	See 1.1a	
2.2	Cancer (all sites)	140–208
3.1	See 1.1	
3.1a	See 1.1a	
3.2	Lung cancer	162.2–162.9
3.3	Chronic obstructive pulmonary disease	490–496
4.1	Alcohol-related motor vehicle crashes	E810–E819
4.1a	[American Indians/Alaska Natives]	
4.1b	[Ages 15–24]	
4.2	Cirrhosis	571
4.2a	[Black males]	
4.2b	[American Indians/Alaska Natives]	
4.3	Drug-related deaths	292, 304, 305.2–305.9, E850–E858, E950.0–E950.5, E962.0, E980.0–E980.5
6.1	Suicides	E950–E959
6.1a	[Ages 15–19]	
6.1b	[Males 20–34]	
6.1c	[White males 65 and older]	
6.1d	[American Indian/Alaska Native males]	
7.1	Homicides	E960–E969
7.1a	[Children 0–3]	
7.1b	[Spouses 15–34]	
7.1c	[Black males 15–34]	
7.1d	[Hispanic males 15–34]	
7.1e	[Black females 15–34]	
7.1f	[American Indians/Alaska Natives]	
7.2	See 6.1	
7.2a	See 6.1a	
7.2b	See 6.1b	
7.2c	See 6.1c	
7.2d	See 6.1d	
7.3	Firearm injuries	E922.0–E922.3, E922.8–E922.9, E955.0–E955.4, E965.0–E965.4, E970, E985.0–E985.4
	Knife injuries	E920.3, E956, E966, E986, E974
9.1	Unintentional injuries	E800–E949
9.1a	[American Indians/Alaska Natives]	
9.1b	[Black males]	
9.1c	[White males]	
9.3	Motor vehicle crashes	E810–E825
9.3a	[Ages 14 and younger]	
9.3b	[Ages 15–24]	
9.3c	[Ages 70 and older]	
9.3d	[American Indians/Alaska Natives]	
9.3e	[Motorcyclists]	
9.3f	[Pedestrians]	
9.4	Falls and fall-related injuries	E880–E888
9.4a	[Ages 65–84]	
9.4b	[Ages 85+]	
9.4c	[Black males 30–69]	
9.5	Drowning	E830, E832, E910
9.5a	[Ages 0–4]	
9.5b	[Males 15–34]	
9.5c	[Black males]	
9.6	Residential fires	E890–E899
9.6a	[Ages 0–4]	
9.6b	[Ages 65 and older]	
9.6c	[Black males]	
9.6d	[Black females]	
10.1	Work-related injuries[1]	E800–E999
10.1a	[Mine workers]	
10.1b	[Construction workers]	
10.1c	[Transportation workers]	
10.1d	[Farm workers]	
13.7	Cancer of the oral cavity and pharynx	140–149
14.3	Maternal mortality	630–676
14.3a	[Blacks]	
15.1	See 1.1	
15.1a	See 1.1a	
15.2	Stroke	430–438
15.2a	[Blacks]	
16.1	See 2.2	
16.2	See 3.2	
16.3	Breast cancer in women	174
16.4	Cancer of the uterine cervix	180
16.5	Colorectal cancer	153.0–154.3, 154.8,159.0
17.9	Diabetes-related deaths[1]	250
17.9a	[Blacks]	
17.9b	[American Indians/Alaska Natives]	
20.2	Epidemic-related pneumonia and influenza deaths for ages 65+	480–487

[1] *Healthy People 2000* uses multiple cause-of-death data.

This appendix presents examples of the computation of the age-adjusted death rate (ADR), indirect adjusted death rate (IADR), and the standard mortality ratio (SMR). These examples demonstrate that each standardized index is a weighted average of the age-specific rates. For the ADR, the weights are determined by the standard population. A discussion of the variability of the ADR is also included.

Suppose the data are aggregated into i = 1, 2, ..., I age groups. Let:

d_i = the number of deaths in the i-th age interval, and
p_i = the population size in the i-th age interval.

The total number of deaths is

$$d = \sum_i d_i$$

the total population is

$$p = \sum_i p_i$$

Age specific death rates (ASDRs) are defined as

$$\text{ASDR} = \frac{\textit{number of deaths for age interval i}}{\textit{midyear population for age interval i}}$$

thus,

$$ASDR = m_i = \textit{the death rate in the i-th age interval.}$$

The age-specific death rate is given by

$$m_i = d_i / p_i$$

In this form, the death rate (*m*) on a unit basis (i.e., per person) will be between 0 and 1. ASDRs are usually expressed as a rate per 1,000 or per 100,000 population. For example, if there are 10 deaths in an age group that has a total population of 1,000 persons, the ASDR on a unit basis is 0.01; per 1,000 it is 10, per 100,000 it is 1,000.

The annual crude death rate is defined as the total number of deaths over all ages divided by the midyear population. The crude death rate is then

$$m = \textit{total deaths} / \textit{total population}$$

Again, it is usually expressed per 1,000 or per 100,000 population.

Algebraically, the direct standardized (or age-adjusted) rate is a weighted average of the age-specific death rates. To compute the ADR, the standard population is used to determine a set of weights. For convenience, let

p_{si} = population in age group i in the standard population and let the standard weights be given by

$$w_{si} = \frac{p_{si}}{\sum_i p_{si}}$$

[NOTE: That in this form $0<w_{si}<1$ and the w_{si} sum to 1. The weights are often expressed as a standard million so that w_{si} sum to 1,000,000.]

Then the ADR is given by

$$\text{ADR} = \sum_i w_{si} * m_i$$

The ASDRs used by NCHS to compute the ADR are rounded to one decimal place. The weights used by NCHS are based on the 1940 U.S. population and are called the ‘‘standard million.’’ As the name implies, the standard million weights sum to one million. The standard million is shown in table II.

The age-adjusted rates shown in most NCHS publications and those used to track the *Healthy People 2000* objectives are computed using the standard million and ASDRs in 10-year age groups. A specific calculation for stroke mortality (*Healthy People 2000* objective 15.2) for males and females is shown in table III. For illustrative purposes, the deaths and populations are those of a hypothetical medium-sized State.

In this example, d_i = deaths in 10-year age-groups, m_i = 10-year ASDR per 100,000, and $w_{si} = p_{si} / \Sigma_i p_{si}$ (weights on a unit basis).

Indirect standardization

For direct standardization, the observed ASDRs and a standard population are used. For indirect standardization, the observed population and a standard set of ASDRs are used. Indirect standardized rates are sometimes calculated and presented, but more often an SMR is presented. The indirect standardized rate and SMR are defined as follows:

$$\text{SMR} = \frac{\textit{number of observed deaths}}{\textit{number of expected deaths}}$$

or

$$\text{SMR} = \frac{\sum_i d_i}{\sum_i m_{si} * p_i}$$

where m_{si} are the standard ASDRs on a unit basis. The indirect adjusted death rate is then

$$\text{IADR} = \text{SMR} * \textit{(crude rate for the standard population)}$$

or

$$\text{IADR} = \frac{M_s * \sum_i d_i}{\sum_i m_{si} * p_i}$$

For the data in table A, the age-specific rates for community A can be used as the standard rates. Then the crude rate and the indirect standardized rates are the same for community A (50 per 1,000); the SMR for community A is 1. The calculation for the IADR and SMR for community B is shown in table IV.

Then

$$\text{SMR} = \frac{400}{300} = 1.33 \qquad \text{IADR} = 1.33 * 50 = 67$$

Each index has advantages and disadvantages. Indirect rates can be used when age-specific numbers of deaths are not available or when the number of deaths is small. Also, the indirect standardized rates have smaller variability. However, indirect rates may not be comparable across areas; they can be used for comparisons of areas only if age and area effects are independent.

Variability

The numbers of deaths reported for a community represent complete counts. As such, numbers of deaths and death rates are not subject to sampling error, although they are subject to errors in the registration process. However, when used for analytic purposes, such as comparison of rates over time or for different areas, the number of events that actually occurred may be considered as one of a large series of possible results that could have arisen under the same circumstances. The probable range of values may be estimated from the actual figures according to certain statistical assumptions. From these assumptions the standard errors of ASDRs and ADRs can be calculated (19,20).

The variance of an ASDR is assumed to be determined by a binomial distribution. This assumes that the chance of dying in an age interval is constant within the age interval and that everyone has the same chance of dying; this is an assumption of homogeneity. Under the homogeneity assumption, the variance of an age-specific rate on a unit basis is given by

$$\text{Variance}\ (m_i) = \frac{m_i * (1 - m_i)}{p_i}$$

NOTE: If the rates are per 100,000, then the (1–m_i) term becomes (100,000–m_i).

The variance of an ADR can be defined as weighted average of the variances of the ASDRs. Under the assumption that ASDR are independent, or that the covariance (m_i, m_j) = 0 for i not equal to j, then the standardized rates are simply weighted averages of the age-specific rates, and the variance is given by

$$\text{Variance (ADR)} = \sum_{\text{i}} w_i^2 * \text{Variance}\ (m_i)$$

An example of the calculation of the variances of the age-adjusted rates for males and females computed in table III is shown in table V. In order to obtain meaningful variances, the number of deaths and the populations for males and females used in table V are those of the same hypothetical medium-sized State used in table III. The variance of an age-adjusted death rate for the entire U.S. population is extremely small.

Confidence intervals can be formed using the variances. If the number of deaths is large enough (again, a rough principle is 25 or more) then a 95-percent confidence interval for the age-adjusted rate is formed as:

$$(\text{ADR} - 1.96 * \sqrt{var\ (m_i)}\ ,\ \text{ADR} + 1.96 * \sqrt{var\ (m_i)}\ )$$

For the example in table III, the 95-percent confidence interval for the ADR for strokes for males for the hypothetical State is

$$[33.0 - (1.96 * 1.05)\ ,\ 33.0 + (1.96 * 1.05)]$$

or

$$(30.9, 35.1)$$

For females, the 95-percent confidence interval is

$$[27.9 - (1.96 * 0.80)\ ,\ 27.9 + (1.96 * 0.80)]$$

or

$$(27.1, 28.7)$$

Because the 95-percent confidence intervals do not overlap, the difference between the ADRs for males and females is statistically significant at the 0.05 level.

Some care has to be exercised when both the rates are low and the number of deaths is small. In this case, the above formula can result in the lower bound being less than zero, and death rates cannot be negative. One way to avoid this is to use log transformations of the rates; another way is to use a discrete distribution function. These methods are beyond the scope of this report. The NCHS Office of Research and Methodology can provide assistance on computing variances for ADRs based on small frequencies.

When ASDRs are based on sufficiently small numbers, a simple Poisson approximation may be used to compute the variance of the ASDRs, as follows:

$$\frac{m_i^2}{d_i}$$

where m_i is the ASDR on a unit basis for the i-th age group, and d_i is the corresponding number of deaths

In these cases, the resultant Poisson age-specific variances can be used in the formulae described in this section to compute the variance of the ADR. More information on random variation can be found in the annual vital statistics volumes (5).

Table II. Standard million age distribution used to adjust death rates to the U.S. population in 1940

Age	*Standard million (p_{si})*	*Unit basis (p_{si} / 1,000,000)*
All ages	1,000,000	1.00000
Under 1 year	15,343	0.015343
1–4 years	64,718	0.064718
5–14 years	170,355	0.170355
15–24 years	181,677	0.181677
25–34 years	162,066	0.162066
35–44 years	139,237	0.139237
45–54 years	117,811	0.117811
55–64 years	80,294	0.080294
65–74 years	48,426	0.048426
75–84 years	17,303	0.017303
85 years and over	2,770	0.002770

Revised Table III

Table III. Age-adjusted death rate calculation for stroke (ICD-9 430–438) for males and females: Hypothetical medium-sized State

Age	d_i	p_i (thousands)	m_i (per 100,000)	w_{si} (unit basis)	$m_i * w_{si}$
Males					
Under 1 year	1	38	2.6	0.015343	0.0398918
1–4 years	–	150	–	0.064718	0.0000000
5–14 years	1	322	0.3	0.170355	0.0511065
15–24 years	2	344	0.6	0.181677	0.1090062
25–34 years	8	443	1.8	0.162066	0.2917188
35–44 years	21	379	5.5	0.139237	0.7658035
45–54 years	46	256	18.0	0.117811	2.1205980
55–64 years	103	189	54.5	0.080294	4.3760230
65–74 years	254	136	186.8	0.048426	9.0459768
75–84 years	371	57	650.9	0.017303	11.2625227
85 years and over	212	12	1,766.7	0.002770	4.8937590
					Age-adjusted rate = Sum ($m_i * w_{si}$) = 33.0 (per 100,000)
Females					
Under 1 year	1	36	2.8	0.015343	0.0429604
1–4 years	–	143	–	0.064718	0.0000000
5–14 years	1	309	0.3	0.170355	0.0511065
15–24 years	2	337	0.6	0.181677	0.1090062
25–34 years	7	458	1.5	0.162066	0.2430990
35–44 years	21	401	5.2	0.139237	0.7240324
45–54 years	41	267	15.4	0.117811	1.8142894
55–64 years	83	208	39.9	0.080294	3.2037306
65–74 years	245	178	137.6	0.048426	6.6634176
75–84 years	553	100	553.0	0.017303	9.5685590
85 years and over	661	34	1,944.1	0.002770	5.3851570
					Age-adjusted rate = Sum ($m_i * w_{si}$) = 27.8 (per 100,000)

Table IV. Calculation of SMR and indirect adjusted death rate

Age	m_{si} (per 1,000)	m_{si} (unit basis)	d_i	p_i	$m_{si} * p_i$
0–34 years	20	0.02	180	6,000	120
35–64 years	40	0.04	150	3,000	120
65 years and over	60	0.06	70	1,000	60
Total	120	0.05	400	10,000	300

Revised Table V

Table V. Variance calculation for the age-adjusted death rate calculation for stroke (ICD-9 430–438) for males and females: Hypothetical medium-sized State

Age	d_i	p_i (thousands)	m_i (per 100,000)	w_{si} (unit basis)	Var (m_i)	w_i^2* Var(m_i)
Males						
Under 1 year	1	38	2.6	0.015343	6.84193	0.00161
1–4 years	–	150	–	0.064718	0.00000	0.00000
5–14 years	1	322	0.3	0.170355	0.09317	0.00270
15–24 years	2	344	0.6	0.181677	0.17442	0.00576
25–34 years	8	443	1.8	0.162066	0.40631	0.01067
35–44 years	21	379	5.5	0.139237	1.45111	0.02813
45–54 years	46	256	18.0	0.117811	7.02998	0.09757
55–64 years	103	189	54.5	0.080294	28.82026	0.18581
65–74 years	254	136	186.8	0.048426	137.09637	0.32150
75–84 years	371	57	650.9	0.017303	1134.49700	0.33966
85 years and over	212	12	1,766.7	0.002770	14462.39759	0.11097
Variance of age-adjusted death rate = Sum w_1^2 * Var (m_i) = 1.10						
Standard error of ADR = Square root of variance = 1.05						
Females						
Under 1 year	1	36	2.8	0.015343	7.77756	0.00183
1–4 years	–	143	–	0.064718	0.00000	0.00000
5–14 years	1	309	0.3	0.170355	0.09709	0.00282
15–24 years	2	337	0.6	0.181677	0.17804	0.00588
25–34 years	7	458	1.5	0.162066	0.32751	0.00860
35–44 years	21	401	5.2	0.139237	1.29669	0.02514
45–54 years	41	267	15.4	0.117811	5.76690	0.08004
55–64 years	83	208	39.9	0.080294	19.17504	0.12362
65–74 years	245	178	137.6	0.048426	77.19700	0.18103
75–84 years	553	100	553.0	0.017303	549.94191	0.16465
85 years and over	661	34	1,944.1	0.002770	5606.77868	0.04302
Variance of age-adjusted death rate = Sum w_1^2 * Var (m_i) = 0.64						
Standard error of ADR = Square root of variance = 0.80						

DEPARTMENT OF
HEALTH & HUMAN SERVICES

Public Health Service
Centers for Disease Control and Prevention
National Center for Health Statistics
6525 Belcrest Road
Hyattsville, Maryland 20782

BULK RATE
POSTAGE & FEES PAID
PHS/NCHS
PERMIT NO. G-281

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

DHHS Publication No. (PHS) 95–1237
4-1476 (3/95)