American Journal of Epidemiology
Vol. 149, No. 12
Printed in U.S.A.

ORIGINAL CONTRIBUTIONS

Look before You Leap: Stratify before You Standardize

Bernard C. K. Choi,[1-3] Nicole A. de Guia,[2] and Peter Walsh[1]

This paper presents a mathematical model to show the conditions in which age standardization can be used to summarize age-specific rates for comparison purposes over calendar time. It shows that the conditions for valid comparison depend on the type of measure used for comparison, that is, difference, ratio, or percent change. If the measure for comparison is a difference of the standardized rates at two time points, then the age-specific rates need to maintain a constant rate difference over time for the comparison to be valid. If the measure for comparison is a ratio or percent change of the standardized rates at two time points, then the age-specific rates need to maintain a constant rate ratio over time for the comparison to be valid. Since in reality, as shown by our Canadian empirical data, age-specific rates do not always maintain a consistent pattern over time, it is recommended that one should always stratify the data to look at patterns of age-specific rates before applying age standardization. *Am J Epidemiol* 1999;149:1087–96.

confounding; data interpretation, statistical; epidemiologic methods; interaction; models, statistical

Age standardization is a method that is used extensively in epidemiology, because the individual age-specific rates "provide too much detailed information" (1, p. 2). Decision makers quickly lose interest when there is a large table of numbers. The method dates back to 1844, when Neison suggested comparing the mortality of two communities by having the population of one community "actually transferred" to the other community and subjecting it to "exactly the same rate of mortality as that prevailed" in the other community (2, p. 4). This could overcome the differences in age structure in the two communities. In modern day terminology, the first community is the standard population, and Neison's method is a direct standardization method.

While age standardization might remove the confounding effect of age, previous studies have shown problems with using standardized rates, even for comparison purposes. The choice of a standard can make a substantial difference in the standardized rates and therefore affect setting priorities. Spiegelman and Marks (3) compared mortality using the populations of 1940, 1950, and 1960 as standards and found a major change in the age-standardized rates for stroke and moderate change for diabetes. A study by the Metropolitan Life Insurance Company of mortality in the 50 American states in 1968 showed that 21 states were ranked differently, depending on whether the standard used was the 1940 or 1970 population (4). Feinleib (5) found that age-standardized cancer mortality rates show different degrees of increase from 1940 to 1968 when using five different standard populations, from 10 percent using the US 1940 population to 15 percent using the 2020 population. Kleinman (6) found that from 1980 to 1988, age-standardized cancer mortality rates changed by –0.1 percent (using the US 1940 population), +1.5 percent (US 1980 population), and +2.5 percent (US 2050 population). Rothenberg and Hahn (7) compared the effects of eight standard populations and found that the annual percent change in age-standardized rates for chronic obstructive pulmonary disease ranges from –0.1 percent (US 1940 population) to 0.4 percent (US 1980 population) and 0.7 percent (a uniform standard). Stoto (8) found that, using the US 1990 standard, the cancer death rate increased by 6.2 percent from 1970 to 1980 but that, using the 1940 standard, it increased by only 2.3 percent.

Because of the potential problems with the choice of standard populations, there has been renewed interest in

Received for publication May 12, 1998, and accepted for publication September 25, 1998.

[1] Bureau of Cardio-Respiratory Diseases and Diabetes, Laboratory Centre for Disease Control, Health Canada, Tunney's Pasture, Ottawa, ON, Canada.
[2] Department of Public Health Sciences, Faculty of Medicine, University of Toronto, Toronto, ON, Canada.
[3] Department of Epidemiology and Community Medicine, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada.
Reprint requests to Dr. Bernard C. K. Choi, Bureau of Cardio-Respiratory Diseases and Diabetes, Laboratory Centre for Disease Control, Health Canada, PL#1918C3, Tunney's Pasture, Ottawa, ON, Canada K1A 0K9.

studying the appropriateness of the age-standardized rate as a summary measure and the validity of comparisons between age-standardized rates (1). This paper examines the conditions when age-standardized rates can be used for secular trend analysis. A mathematical model is provided. Empirical data are given to illustrate the model.

MATHEMATICAL MODEL

The crude rate (r) in a group is the number of events (d), such as death or disease, in a year divided by the midyear population (n) (table 1),

$$r = d/n. \tag{1}$$

The crude rate is also the weighted average of the age-specific rates by the current midyear population,

$$r = (\Sigma r_i \cdot n_i)/(\Sigma n_i) = \Sigma r_i \cdot p_i, \tag{2}$$

where r_i is the rate at age i (or d_i/n_i), and p_i is the proportion of persons at age i in the total population in that year (or n_i/n) (table 1). Equation 2 is of interest because it emphasizes the composite nature of the crude rate. It also shows that the value of the crude rate depends on the age-specific rates r_i and on the age-specific distribution of the population p_i. Because of this, the crude rate is not recommended for comparison of rates between populations, such as populations in different geographic areas, or of rates within the same population at different times, because of differing population distributions.

The directly age-standardized rate (r') is a weighted average of the age-specific rates, where the weights (N_i) of the corresponding age-specific rates are derived from the age distribution of the standard population. Division by ΣN_i converts any set of weights N_i into a set of proportions P_i (or $N_i/\Sigma N_i$) that add to unity. Once the population proportion weights are selected, they are multiplied by the age-specific rates, and the results are summed to get a weighted average (9).

TABLE 1. Notations for age-specific distributions of events, population, and rates in a group and in the standard population

Age stratum	Group			Standard population		
	Event	Population	Rate	Event	Population	Rate
1	d_1	n_1	r_1	D_1	N_1	R_1
.						
.						
.						
i	d_i	n_i	r_i	D_i	N_i	R_i
Total	d	n	r	D	N	R

$$r' = (\Sigma r_i \cdot N_i)/(\Sigma N_i) = \Sigma r_i \cdot P_i. \tag{3}$$

Comparison of two age-standardized rates, r_1' and r_2', can be achieved using the following:

1. the rate difference, $r_2' - r_1'$ (10),
2. the rate ratio, r_2'/r_1' (11), or
3. the percent change, $(r_2' - r_1')/r_1'$ (11).

For comparison of rates over time, consider two time points, A and B. Consider further three scenarios with respect to the change of age-specific rates from time A to time B:

1. constant stratum-specific rates in time A and time B, that is, $r_{Ai} = r_{Bi}$,
2. constant stratum-specific rate difference, that is, $r_{Bi} - r_{Ai} = K$, where K is a constant, or
3. constant stratum-specific rate ratio, that is, $r_{Bi}/r_{Ai} = K$, where K is a constant.

The crude rate at time A is $r_A = (\Sigma r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai})$. As shown above, the crude rate at time A is also the directly standardized rate at time A using the population at time A as the standard, ${}_Ar_A'$. The directly standardized rate at time A using time B as the standard population is ${}_Br_A' = (\Sigma r_{Ai} \cdot n_{Bi})/(\Sigma n_{Bi})$.

The crude rate at time B is $r_B = (\Sigma r_{Bi} \cdot n_{Bi})/(\Sigma n_{Bi}) = {}_Br_B'$. The directly standardized rate at time B using time A as the standard population is ${}_Ar_B' = (\Sigma r_{Bi} \cdot n_{Ai})/(\Sigma n_{Ai})$.

The comparison of standardized rates at time A and time B (table 2) may yield different results using various measures for comparison under various scenarios as shown below.

Scenario 1: constant stratum-specific rates, that is, $r_{Ai} = r_{Bi}$

The rate difference.

$${}_Ar_B' - {}_Ar_A' = (\Sigma r_{Bi} \cdot n_{Ai})/(\Sigma n_{Ai}) - (\Sigma r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai})$$

$$= 0 = (\Sigma r_{Bi} \cdot n_{Bi})/(\Sigma n_{Bi}) - (\Sigma r_{Ai} \cdot n_{Bi})/(\Sigma n_{Bi})$$

$$= {}_Br_B' - {}_Br_A'.$$

The rate ratio.

$${}_Ar_B'/{}_Ar_A' = \frac{(\Sigma r_{Bi} \cdot n_{Ai})/(\Sigma n_{Ai})}{(\Sigma r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai})} = 1$$

$$= \frac{(\Sigma r_{Bi} \cdot n_{Bi})/(\Sigma n_{Bi})}{(\Sigma r_{Ai} \cdot n_{Bi})/(\Sigma n_{Bi})} = {}_Br_B'/{}_Br_A'.$$

TABLE 2. Comparison of standardized rates at time A and time B, using the population at time A or time B as the standard population

	Standardized rate	Measure for comparison		
		Rate difference	Rate ratio	% change
Using A as standard				
Time A	${}_Ar_A'$	${}_Ar_B' - {}_Ar_A'$	${}_Ar_B' / {}_Ar_A'$	$({}_Ar_B' - {}_Ar_A') / {}_Ar_A'$
Time B	${}_Ar_B'$			
Using B as standard				
Time A	${}_Br_A'$	${}_Br_B' - {}_Br_A'$	${}_Br_B' / {}_Br_A'$	$({}_Br_B' - {}_Br_A') / {}_Br_A'$
Time B	${}_Br_B'$			

The percent change.

$$({}_Ar_B' - {}_Ar_A')/{}_Ar_A'$$

$$= \frac{(\Sigma r_{Bi} \cdot n_{Ai})/(\Sigma n_{Ai}) - (\Sigma r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai})}{(\Sigma r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai})}$$

$$= 0 = \frac{(\Sigma r_{Bi} \cdot n_{Bi})/(\Sigma n_{Bi}) - (\Sigma r_{Ai} \cdot n_{Bi})/(\Sigma n_{Bi})}{(\Sigma r_{Ai} \cdot n_{Bi})/(\Sigma n_{Bi})}$$

$$= ({}_Br_B' - {}_Br_A')/{}_Br_A'$$

Therefore, when stratum-specific rates are constant from time A to time B, comparison of standardized rates at time A and time B using A or B as the standard population will yield identical results regardless of which of the three measures of comparison is used.

Scenario 2: constant stratum-specific rate difference, that is, $r_{Bi} - r_{Ai} = K$

The rate difference.

$${}_Ar_B' - {}_Ar_A' = (\Sigma r_{Bi} \cdot n_{Ai})/(\Sigma n_{Ai}) - (\Sigma r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai})$$

$$= (\Sigma(K + r_{Ai}) \cdot n_{Ai})/(\Sigma n_{Ai}) - (\Sigma r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai}) = K$$

$$= (\Sigma(K + r_{Ai}) \cdot n_{Bi})/(\Sigma n_{Bi}) - (\Sigma r_{Ai} \cdot n_{Bi})/(\Sigma n_{Bi})$$

$$= (\Sigma r_{Bi} \cdot n_{Bi})/(\Sigma n_{Bi}) - (\Sigma r_{Ai} \cdot n_{Bi})/(\Sigma n_{Bi}) = {}_Br_B' - {}_Br_A'.$$

The rate ratio.

$${}_Ar_B'/{}_Ar_A' = \frac{(\Sigma r_{Bi} \cdot n_{Ai})/(\Sigma n_{Ai})}{(\Sigma r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai})}$$

$$= \frac{(\Sigma(K + r_{Ai}) \cdot n_{Ai})/(\Sigma n_{Ai})}{(\Sigma r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai})}$$

$$= \frac{(\Sigma(K + r_{Ai}) \cdot n_{Ai})}{(\Sigma r_{Ai} \cdot n_{Ai})} \neq {}_Br_B'/{}_Br_A'.$$

The percent change.

$$({}_Ar_B' - {}_Ar_A')/{}_Ar_A'$$

$$= \frac{(\Sigma r_{Bi} \cdot n_{Ai})/(\Sigma n_{Ai}) - (\Sigma r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai})}{(\Sigma r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai})}$$

$$= \frac{(\Sigma(K + r_{Ai}) \cdot n_{Ai})/(\Sigma n_{Ai}) - (\Sigma r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai})}{(\Sigma r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai})}$$

$$= \frac{K}{(\Sigma r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai})}$$

$$\neq ({}_Br_B' - {}_Br_A')/{}_Br_A'.$$

Therefore, when differences of stratum-specific rates are constant from time A to time B, comparison of standardized rates at time A and time B using A or B as the standard population will yield identical results only when using the rate difference as a measure of comparison.

Scenario 3: constant stratum-specific rate ratio, that is, $r_{Bi} / r_{Ai} = K$

The rate difference.

$${}_Ar_B' - {}_Ar_A' = (\Sigma r_{Bi} \cdot n_{Ai})/(\Sigma n_{Ai}) - (\Sigma r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai})$$

$$= (\Sigma K \cdot r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai}) - (\Sigma r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai})$$

$$= (\Sigma (K - 1) \cdot r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai}) \neq {}_{B}r_{B}' - {}_{B}r_{A}'.$$

The rate ratio.

$${}_{A}r_{B}'/{}_{A}r_{A}' = \frac{(\Sigma r_{Bi} \cdot n_{Ai})/(\Sigma n_{Ai})}{(\Sigma r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai})}$$

$$= \frac{(\Sigma K \cdot r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai})}{(\Sigma r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai})} = K.$$

$$= {}_{B}r_{B}'/{}_{B}r_{A}'.$$

The percent change.

$$({}_{A}r_{B}' - {}_{A}r_{A}')/{}_{A}r_{A}'$$

$$= \frac{(\Sigma r_{Bi} \cdot n_{Ai})/(\Sigma n_{Ai}) - (\Sigma r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai})}{(\Sigma r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai})}$$

$$= \frac{(\Sigma K \cdot r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai}) - (\Sigma r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai})}{(\Sigma r_{Ai} \cdot n_{Ai})/(\Sigma n_{Ai})} = K - 1$$

$$= ({}_{B}r_{B}' - {}_{B}r_{A}')/{}_{B}r_{A}'.$$

Therefore, when ratios of stratum-specific rates are constant from time *A* to time *B*, comparison of standardized rates at time *A* and time *B* using *A* or *B* as the standard population will yield identical results only when using the rate ratio or percent change as measures of comparison.

The above mathematical model demonstrates that, since comparison of the summary measures would be a comparison of the entire set of age-specific rates for the two population groups, the difference, ratio, or percent change in rates for the two groups should reflect exactly the same difference, ratio, or percent change for every age group. When the difference, ratio, or percent change for age-specific rates is not consistent over time, age standardization is not valid. It also demonstrates that, even if the age-specific rates have a constant rate difference, they should not be summarized by age standardization if a ratio of the two standardized rates would be used for comparison. Moreover, it demonstrates that, if the age-specific rates have a constant rate ratio, age standardization would be wrong if the difference of the two standardized rates is used for comparison.

EMPIRICAL EXAMPLES

We reviewed Canadian hospital separation data from 1971 to 1991 and found several empirical examples demonstrating the validity of age-standardized rates for secular trend analysis.

First, when age-specific rates are consistent over time, age-standardized rates can be used to show the secular trend.

Figure 1 shows that the age-specific hospital separation rates from 1971 to 1991 for cerebrovascular disease (*International Classification of Diseases*, Ninth Revision, codes 430–438) in Canada are basically constant over time. In other words, there is no age and calendar year interaction. In this case, the age-standardized rates are valid. As shown in figure 2, age-standardized rates using the 1971 or 1991 Canadian population as the standard show a very similar time trend. The crude rate, on the other hand, presents a very different time trend pattern from the standardized rates and is deceptive. Note that the crude rate curve crosses over the 1971 standardized rate curve in 1971 and crosses over the 1991 standardized rate curve in 1991 (figure 2).

Second, when age-specific rates are not consistent over time, age-standardized rates cannot be used to show the secular trend.

Figure 3 shows the age-specific hospital separation rates from 1971 to 1991 for asthma (*International Classification of Diseases*, Ninth Revision, code 493) in Canada. The hospital separation rates for the 0- to 1- and for the 1- to 4-year age groups increased by about fourfold over the 20-year period. In this case, since the hospital separation rates increased over the years for some age groups but stayed the same for others, a summary statistic, such as an age-standardized rate, will conceal rather than reveal the true underlying trends and is not valid. This is readily shown in figure 4; the age-standardized rates for asthma using the 1971 Canadian standard population and the 1991 Canadian standard population show distinctly different patterns. The choice of standard population may therefore affect priority ranking and other public health decisions.

Third, when age-standardized rates using different standard populations show similar secular trends over time, it does not always mean that age standardization is valid.

The standardized hospital separation rates for chronic respiratory disease (*International Classification of Diseases*, Ninth Revision, codes 490–508 and 515–517)

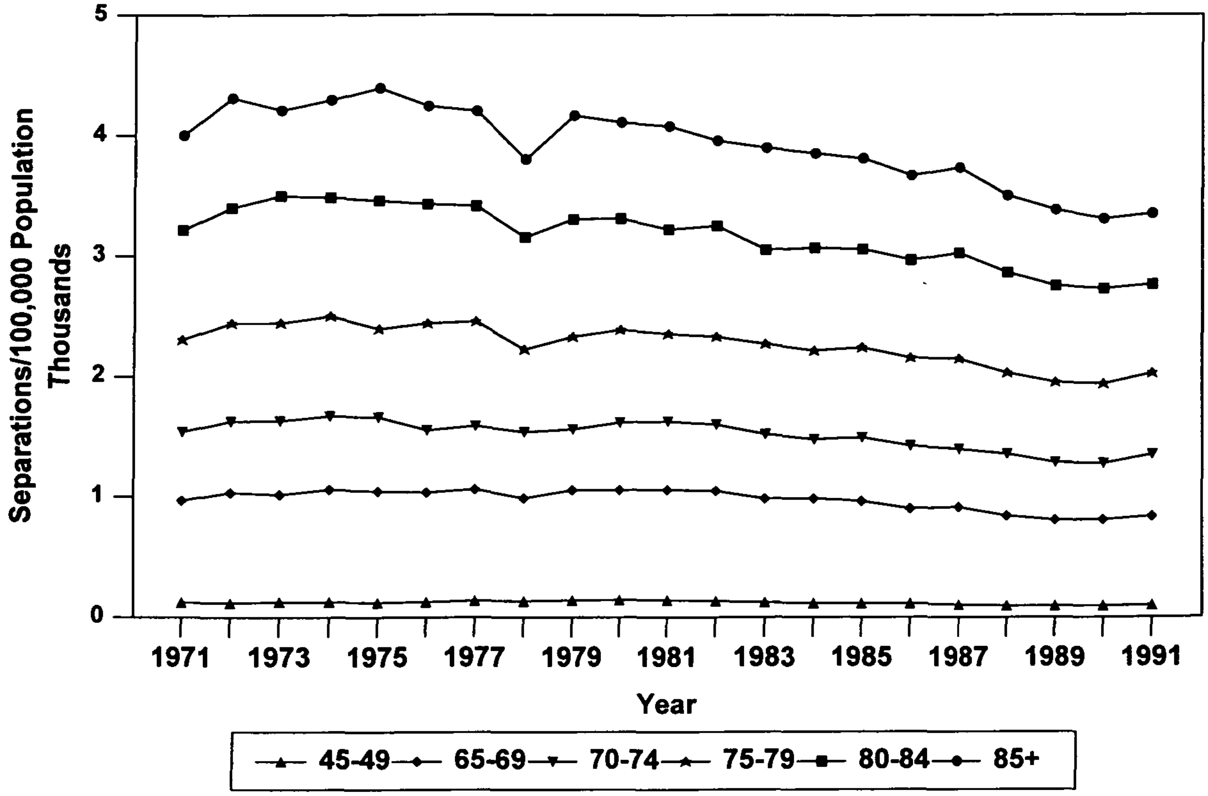

FIGURE 1. Age-specific hospital separation rates for cerebrovascular disease (*International Classification of Diseases*, Ninth Revision, codes 430–438), Canada, 1971–1991. Only selected age-specific curves are shown.

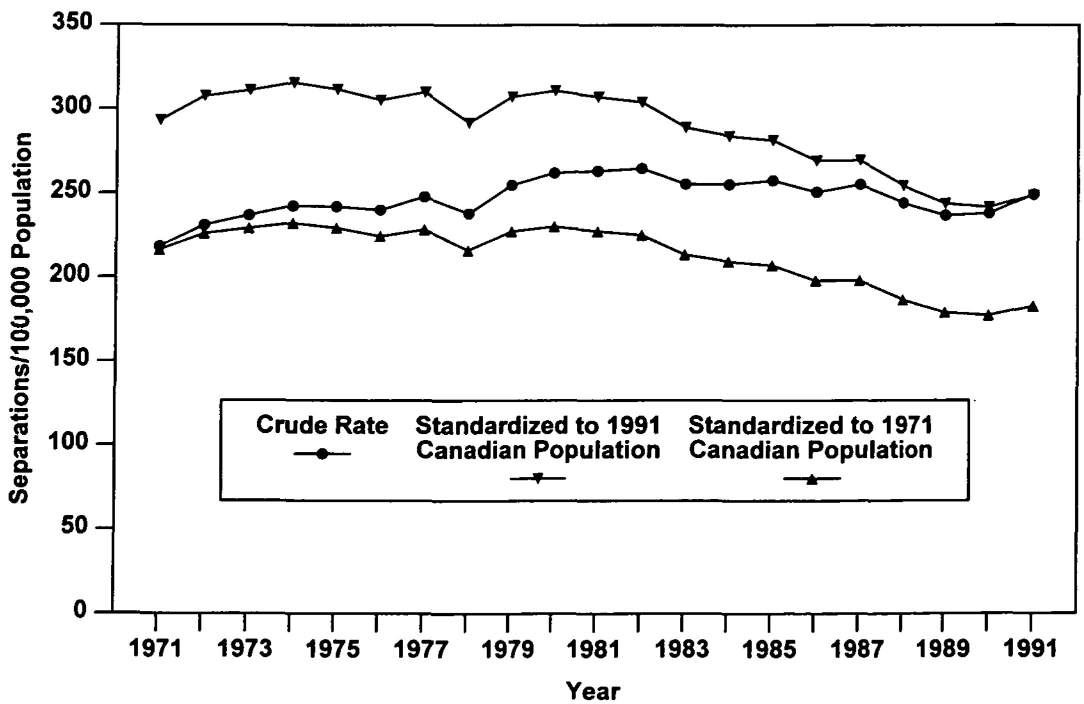

FIGURE 2. Crude and age-standardized hospital separation rates for cerebrovascular disease (*International Classification of Diseases*, Ninth Revision, codes 430–438), Canada, 1971–1991.

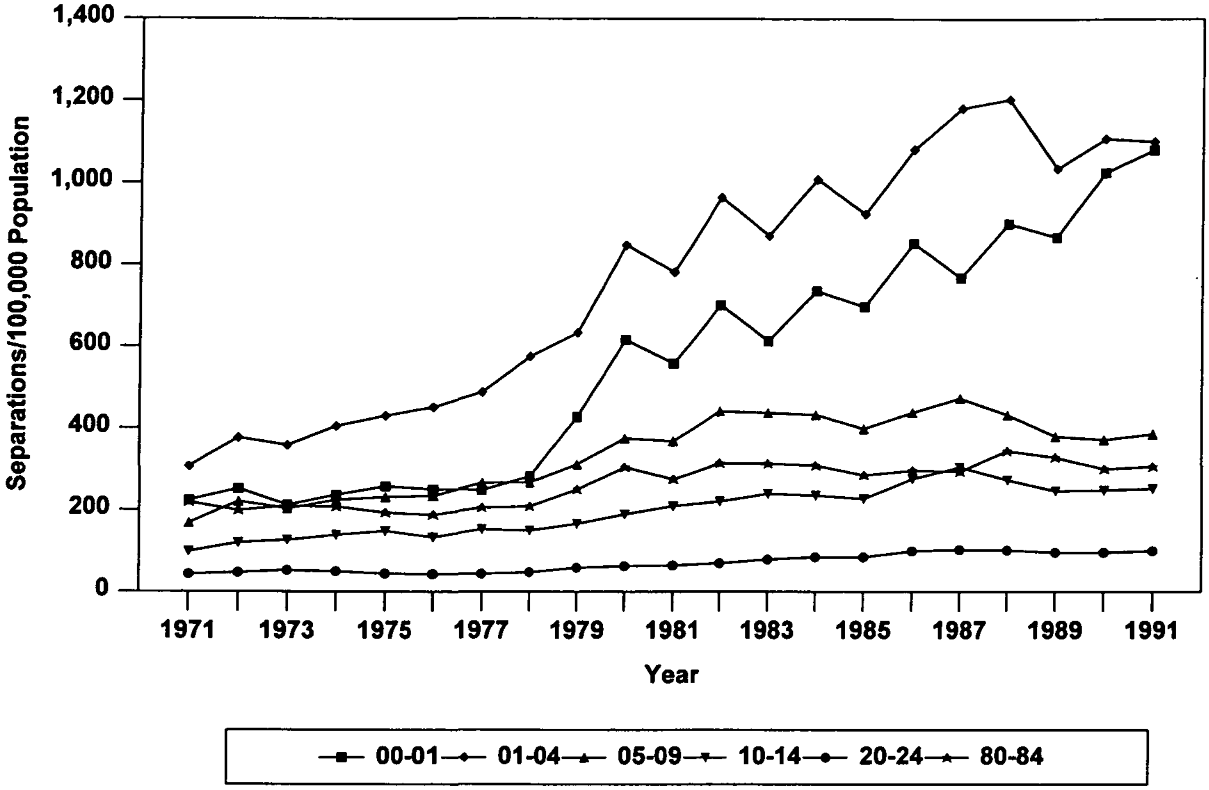

FIGURE 3. Age-specific hospital separation rates for asthma (*International Classification of Diseases*, Ninth Revision, code 493), Canada, 1971–1991. Only selected age-specific curves are shown.

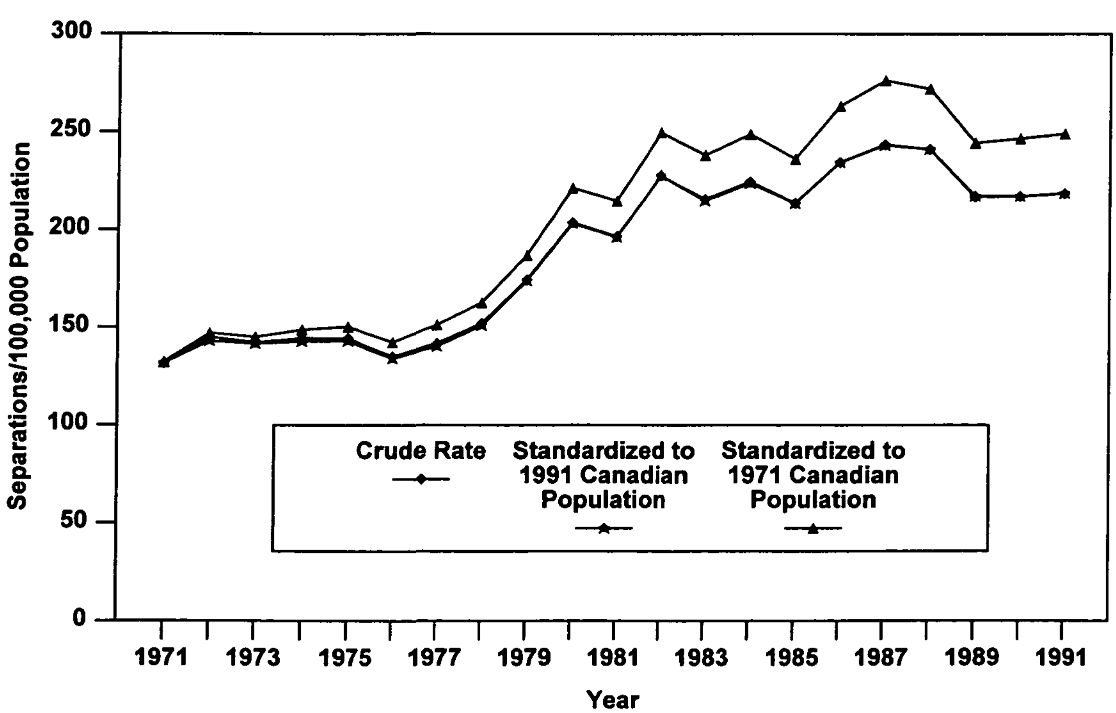

FIGURE 4. Crude and age-standardized hospital separation rates for asthma (*International Classification of Diseases*, Ninth Revision, code 493), Canada, 1971–1991.

in Canada using 1971 or 1991 standard populations show very similar time trends from 1971 to 1991 (figure 5). This observation would have led many researchers to believe that age standardization is valid. However, this is not true. The age-specific rates in fact show very different patterns across age groups over time, with even cross-overs (figure 6). Hospital separation rates for the 0- to 1-year age group decreased over the 20-year period, while those rates for the 80- to 84- and 85+-year age groups increased. Cross-overs of the rates occurred in the early 1980s. In this case, the age-specific rates should have been reported. Any reports based on the age-standardized rates or the crude rates would be invalid.

DISCUSSION

This paper has provided both theoretical and empirical evidence relating to concepts from the literature regarding the validity of age standardization and of age-standardized rates. First, it has been pointed out that, when all the age-specific rates have "a consistent relationship" (1, p. 3) or move in the same direction at the same "relative magnitude" (6, p. 23), the age-standardized rate is a valid summary measure for comparison. This paper, through a mathematical model, demonstrates that the definition of a consistent relation changes with the measure chosen for the comparison (i.e., rate difference, rate ratio, or percent change). For example, the age-standardized rate is a valid summary measure for comparison of the absolute difference of direct standardized rates only when all the age-specific rates move in the same direction at the same absolute magnitude. The algebraic consequences of age adjustment under three scenarios provide theoretical evidence that summary comparisons appropriate at the rate ratio level may not be appropriate at the rate difference level. Although it is not in our data set, we believe that a real-world scenario can be found that satisfies the constant rate ratio criteria between age groups and time but provides a misleading inference at the rate difference level.

Second, it has been suggested that, "in order to compare two age-adjusted rates, the same standard population must have been used" (1, p. 3). This paper shows that this is true only when the age-specific rates for the two population groups being compared have a consistent relation. When age is an effect modifier, comparison of two age-standardized rates would be invalid even if the same standard population is used. In fact, when age is an effect modifier, any age standardization would be invalid. Again, by using a mathematical model, this paper further shows the meaning of a consistent relation, which depends on whether a difference, ratio, or percent change is used for comparison.

Third, it has been suggested by Curtin that, "if it is appropriate to use age adjustment, then the results should not be affected by the selection of a standard population; but if the results can be affected by the choice of a standard population, then it is not appropri-

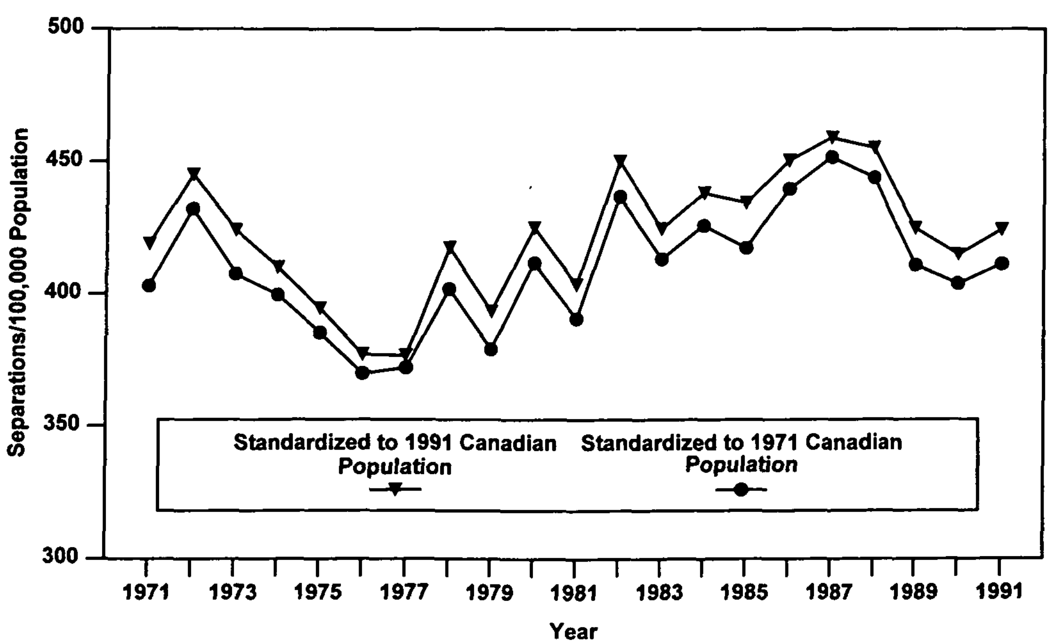

FIGURE 5. Age-standardized hospital separation rates for chronic respiratory disease (*International Classification of Diseases*, Ninth Revision, codes 490–508 and 515–517), Canada, 1971–1991.

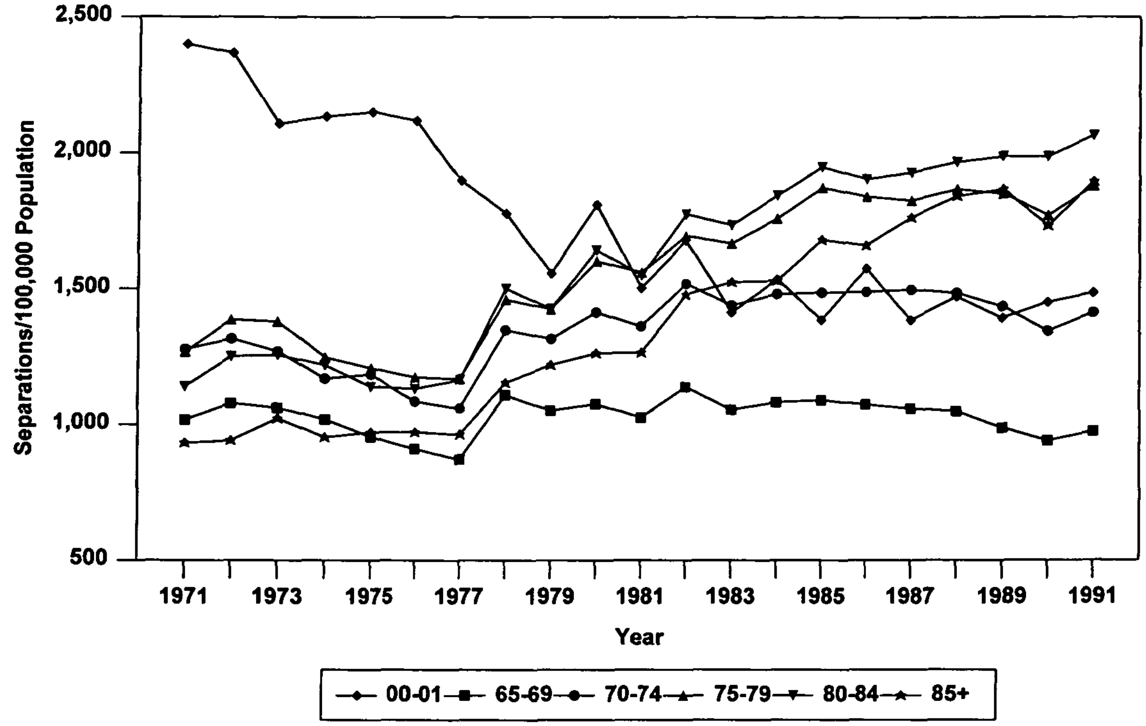

FIGURE 6. Age-specific hospital separation rates for chronic respiratory disease (*International Classification of Diseases*, Ninth Revision, codes 490–508 and 515–517), Canada, 1971–1991. Only selected age-specific curves are shown.

ate to use standardization at all" (12, p. 15). This paper shows that the Curtin test to decide whether age standardization is appropriate or not may not always be true. In some cases, such as our example of chronic respiratory disease, the comparison is not affected by the choice of the 1971 or 1991 Canadian population, but the age-specific rates show distinctly different trends over the years, indicating that age standardization is not valid. Curtin and Klein specified correctly a condition in which the test becomes invalid: "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" (1, p. 3).

It is evident, then, that the use of different standard populations may lead to different perceived trends. Therefore, the question arises as to how to choose the correct standard population. Unfortunately, there is no "standard" way to do this. When direct standardized death rates were first published in reports of the Registrar General of England and Wales in 1883, the standard population was based on the most current census population available, namely, the 1881 census population (12). The initial practice was to change the standard population every 10 years using each new census population. However, this was problematic because of the need to recompute historical rates to assess current trends. At the beginning of the 20th century, England and Wales decided to adopt the 1901 population, again the most current census population, as "the" standard population and to use this standard even after a new decennial census. In the early 20th century, methods for vital statistics analysis in the United States often followed the lead of England and Wales. Age-standardized rates in the first US mortality reports were based on the standard populations of England and Wales in 1901 (13) to provide comparability between the mortality measures produced by the two countries. In the 1940s, it was decided that the US population was different enough from those populations of England and Wales in 1901 that a new standard should be used. At the time, the 1940 US population was the latest census population available and was thus used as the standard. Since that time, all age-standardized rates published in the US *Vital and Health Statistics* series have used the 1940 population as the standard. From time to time, there have been challenges to the use of the 1940 population as the standard, especially when a new census becomes available (5). In 1992, the US National Center for Health Statistics reviewed the considerations and practical implications for implementing new standards and adopted the recommendation that the 1940 population would continue to be used as the basis for calculating age-standardized rates (14). In Canada, no official standard population has been adopted. Most standardized rates have been using either the 1971 or 1991 Canadian populations as standards.

Previous researchers have suggested several considerations, both statistical and nonstatistical, when choosing the standard population. There are two statistical considerations. First, in the comparison among several different populations, a "pooled" standard is recommended for statistical reasons since it minimizes the variance of the direct-standardized rates (15). Second, in the comparison of trends over time, the base-year population is recommended as a standard (5), because statistically the difference between the crude rates for the base-year and a subsequent year can be decomposed into a difference of the age-specific rates weighted by the base-year population (thus, a difference in two standardized rates with the base-year population as a standard), a difference between the populations in the 2 years, and a rate-population interaction term (10).

Three nonstatistical considerations have been proposed. First, the standard population chosen should not be considered abnormal or unnatural relative to the populations under study (12). For example, a public health program whose focus is youth should use a standard that gives greater emphasis to the younger population than a program whose target is the elderly population (11). Second, when a decision is made to change standards, the most current census population should be used (5). Third, standards should not be changed frequently, because of the possibility of enormous resources needed to recompute historical figures and the confusion caused by the publication of a set of data that may no longer be comparable to previously published data (16). A change in standard may make a difference in the perception of disease burden, the ranking of public health problems, and economic efforts and costs (7). Because a change in the standard can lead to a different perceived trend, this reiterates the need for stratification before standardization. After all, age-specific analysis is the only way to establish whether age standardization was appropriate in the first place.

The basic characteristics of the age standardization rate remind us of its many limitations. The age-standardized rate is an index number, and its magnitude changes with the standard population. Therefore, the numerical value of the rate itself has no meaning. In other words, a standardized rate is "artificial" (10) and should be used only for comparison purposes, such as between geographic areas, or over time (1, 17). Furthermore, the age-standardized rate is a summary measure. Thus, situations can arise where such summary measures are inappropriate. For example, according to Inskip et al. (18), if there are large variations in the ratios of the age-specific rates between the populations, then any summary measure can give misleading results. Therefore, researchers must be particularly wary in time trend analysis for, if the age-specific rate trends vary across age groups, an age and calendar time interaction may exist, and thus summary statistics such as the age-standardized rate may actually conceal more than they reveal (16).

REFERENCES

1. Curtin LR, Klein RJ. Direct standardization (age-adjusted death rates). In: Healthy People 2000 statistical notes. No. 6. Atlanta: Centers for Disease Control and Prevention, 1995.
2. Neison FGP. On a method recently proposed for conducting inquiries into the comparative sanitary condition of various districts. Lond J R Stat Soc 1844;7:40–68. Cited in: Feinleib M. Chap 1. A reconsideration of age adjustment. In: National Center for Health Statistics. Reconsidering age-adjustment procedures: workshop proceedings. Atlanta: Centers for Disease Control and Prevention, 1992:4. (Vital and health statistics, Series 4: Documents and committee reports, no. 29) (DHSS publication no. (PHS) 0276-4733).
3. Spiegelman M, Marks HH. Empirical testing of standards for age adjustment of death rates by the direct method. Hum Biol 1966;38:280–92.
4. Metropolitan Life Insurance Company. Standardized mortality rates. Stat Bull Metrop Insur Co 1981;62:5–14.
5. Feinleib M. Chap 1. A reconsideration of age adjustment. In: National Center for Health Statistics. Reconsidering age-adjustment procedures: workshop proceedings. Atlanta: Centers for Disease Control and Prevention, 1992:3–7. (Vital and health statistics, Series 4: Documents and committee reports, no. 29) (DHSS publication no. (PHS) 0276-4733).
6. Kleinman JC. Chap 4. The use of multiple standards. In: National Center for Health Statistics. Reconsidering age-adjustment procedures: workshop proceedings. Atlanta: Centers for Disease Control and Prevention, 1992:23–8. (Vital and health statistics, Series 4: Documents and committee reports, no. 29) (DHSS publication no. (PHS) 0276-4733).
7. Rothenberg R, Hahn R. Chap 6. The choice of the standard for age adjustment. In: National Center for Health Statistics. Reconsidering age-adjustment procedures: workshop proceedings. Atlanta: Centers for Disease Control and Prevention, 1992;49–59. (Vital and health statistics, Series 4: Documents and committee reports, no. 29) (DHSS publication no. (PHS) 0276-4733).
8. Stoto MA. Chap 8. Age adjustment for the year 2000 health objectives. In: National Center for Health Statistics. Reconsidering age-adjustment procedures: workshop proceedings. Atlanta: Centers for Disease Control and Prevention, 1992:65–7. (Vital and health statistics, Series 4: Documents and committee reports, no. 29) (DHSS publication no. (PHS) 0276-4733).
9. Rothman KJ. Modern epidemiology. Boston: Little, Brown and Co, 1986.
10. Kitagawa EM. Components of a difference between two rates. J Am Stat Assoc 1955;50:1168–94.
11. Rosenberg HM, Curtin LR, Maurer J, et al. Chap 5. Choosing a standard population: some statistical considerations. In: National Center for Health Statistics. Reconsidering age-adjustment procedures: workshop proceedings. Atlanta: Centers for Disease Control and Prevention, 1992:29–45.

(Vital and health statistics, Series 4: Documents and committee reports, no. 29) (DHSS publication no. (PHS) 0276-4733).
12. Curtin LR. Chap 2. A short history of standardization for vital events. In: National Center for Health Statistics. Reconsidering age-adjustment procedures: workshop proceedings. Atlanta: Centers for Disease Control and Prevention, 1992:11–16. (Vital and health statistics, Series 4: Documents and committee reports, no. 29) (DHSS publication no. (PHS) 0276-4733).
13. Linder FE, Grove RD. Vital statistics rates in the United States, 1900–40. Washington, DC: US GPO, 1943.
14. Feinleib M. Chap 10. Summary and recommendations. In: National Center for Health Statistics. Reconsidering age-adjustment procedures: workshop proceedings. Atlanta: Centers for Disease Control and Prevention, 1992b:73. (Vital and health statistics, Series 4: Documents and committee reports, no. 29) (DHSS publication no. (PHS) 0276-4733).
15. Kalton G. Standardization: a technique to control for extraneous variables. Appl Stat 1968;17:118–36.
16. Gray-Donald K. Chap 3. Inadequacies of age adjustment as illustrated in cancer mortality. In: National Center for Health Statistics. Reconsidering age-adjustment procedures: workshop proceedings. Atlanta: Centers for Disease Control and Prevention, 1992:17–20. (Vital and health statistics, Series 4: Documents and committee reports, no. 29) (DHSS publication no. (PHS) 0276-4733).
17. Wolfenden HH. On the methods of comparing the mortalities of two or more communities, and the standardization of death rates. J R Stat Soc 1923;86:399–411.
18. Inskip H, Beral V, Fraser P. Methods for age-adjustment of rates. Stat Med 1983;2:455–66.