



STATISTICAL PRIMER

State Center for Health Statistics • 1908 Mail Service Center • Raleigh, NC 27699-1908 • 919/733-4728
www.schs.state.nc.us/SCHS/

No. 17

May 2007

Understanding Different Population Data Sources and Their Impact on Vital Statistics Rates

by

Kathleen A. Jones-Vessey

The State Center for Health Statistics (SCHS) routinely publishes reports that rely on population data for calculating rates. Some people falsely assume that annual population figures are precise, fixed counts that are not subject to change over time. In reality, population figures are essentially educated estimates of the population based on extrapolating figures from a variety of sources such as decennial Census data, births, deaths, school enrollment, and migration patterns.

Traditionally, the SCHS has relied on population figures generated by the State Demographer's Office. Most of our annual reports, including North Carolina Vital Statistics Volume I, Volume II, the Pregnancy report, and other SCHS publications continue to report rates based on population denominators derived from the annual July 1 Certified County Population Estimates produced by the State Demographer's Office. The State Demographer produces population figures for each county by single-year age groups and by race and sex. Using the State Demographer's estimates, racial population estimates are only available for whites and minorities and do not include figures for Hispanic populations or for specific races other than white. To produce their estimates and projections, the State Demographer's Office

develops and enhances complex mathematical computer models, and collects and reviews a variety of data from federal, state, and local government sources.¹ The certified population estimates produced by the State Demographer are periodically updated and revised after the initial release. Additionally, the State Demographer's files are typically "smoothed" after the next U.S. Census to adjust for population changes deduced from the decennial Census. For the latest North Carolina population data from the State Demographer's Office, go to demog.state.nc.us.

In recent years, alternative population data sources have become available. The Population Estimates Program of the U.S. Census Bureau in collaboration with the National Center for Health Statistics (NCHS) has begun to generate annual county-level resident post-census "bridged population estimates." NCHS bridged population files take 2000 Census population data which included 31 different race categories (and which allowed individuals to choose more than one race) and bridge it back to four single-race categories (White, Black/African American, American Indian/Alaska Native, and Asian/Pacific Islander). The file also includes population estimates for ethnicity (Hispanic/Latino and non-Hispanic/Latino) by race. The post-census



estimates are updated annually as additional data become available. The file was originally produced so that the new Census population data, which adhere to the 1997 federal Office of Management and Budget (OMB) standards for the collection of multiple race data, could be used by researchers analyzing vital records which still included only a single race.²

Given the changing population of North Carolina and the growing need for more information measuring the extent of health disparities, SCHS statisticians routinely receive requests for rates based on more detailed racial and ethnic breakdowns. In the past, these requests could not be answered. However, with the availability of NCHS bridged population data, race and ethnicity-specific rates are now produced. Several of our periodic reports, including the Racial and Ethnic Disparities Report Card and the Hispanic Health Fact Sheet, now use these files to generate rates for specific races and for Hispanics/Latinos.

Hispanic origin is not considered a racial category, but rather an ethnicity. As such, unless noted otherwise, rates presented for specific races include Hispanics/Latinos in one of the racial categories. Typically, according to NCHS coding specifications, Hispanic/Latinos are coded to “white” in vital statistics data.³ Using the NCHS bridged population file, the SCHS now often presents data which combine race and ethnicity categories: for example, white, non-Hispanic; African-American/Black, non-Hispanic; Other races/ non-Hispanic; and Hispanic/Latino.

Differences in Federal and State Generated Population Estimates

The SCHS continues to use the July 1 Certified Estimates of the State Demographer’s Office for our annual reports, while also using the NCHS bridged population files for special studies or ad hoc requests. One negative effect of this is that some of the rates reported in our annual publications using the State Demographer’s population data in the denominator do not match rates presented in special or periodic reports which use the NCHS bridged population data. Overall, there is not a large difference between the two files.

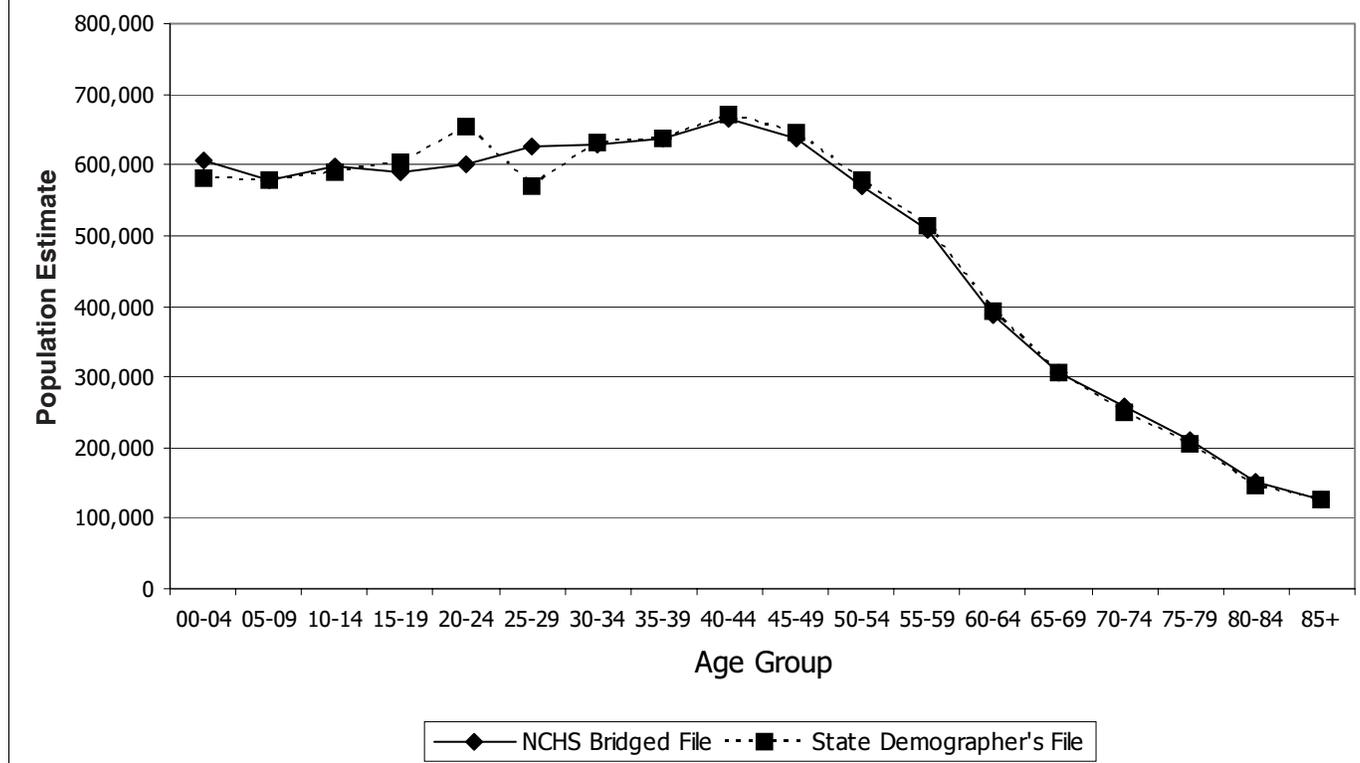
Table 1 presents population estimates from 2005 using the NCHS bridged population files and the North Carolina State Demographer’s population files. In total, there was only a 1,176-person difference in the population estimates from the two files – with the State Demographer’s file being lower than the NCHS file overall for 2005.

The two files also show some variability with regard to population estimates by age group. **Figure 1** presents 2005 North Carolina population estimates from the two files by age group. While the overall total percentage difference was 0.0%, the percentage difference varies for specific age groups. The young adult age groups show the greatest differences between the two files with the 20-24 and 25-29 age groups reflecting differences of more than eight percent.

Table 1: 2005 North Carolina Resident Population Using the NCHS Bridged Population File and the State Demographer’s Certified July 1 Population Estimates Maintained by SCHS

	State Demographer Certified Estimate	NCHS Bridged Estimate	Numeric Difference	Percentage Difference
Gender:				
Male	4,277,092	4,271,569	5,523	0.1%
Female	4,404,974	4,411,673	6,699	-0.2%
Race:				
White	6,479,702	6,478,615	1,087	0.0%
Minority	2,202,364	2,204,627	2,263	-0.1%
TOTAL	8,682,066	8,683,242	1,176	0.0%

Figure 1: 2005 North Carolina Resident Population Estimates



Differences between the two population files by race are harder to measure since the State Demographer's file does not contain specific race estimates other than white and minority. Overall, the white population estimates for 2005 generated by the two files differ by 1,087 people, with the State Demographer's file estimating higher figures than the bridged file. With regard to the minority population, again there is virtually no percentage difference between the files, though the estimated minority populations differed by more than 2,200 people.

Population estimates by gender are also available in both of the population files. Comparing the bridged population file estimates for 2005 and the 2005 State Demographer's population file, we do see some differences in the estimates by gender. The 2005 State Demographer's file has 5,523 more male residents of North Carolina than the NCHS bridged estimates. In

contrast, the NCHS bridged population file shows 6,699 more females for 2005 than the State Demographer's July 1 certified estimate. However, given North Carolina's large population, the percentage differences with regard to gender in the two files is extremely small.

Both the NCHS bridged population file and the State Demographer's certified July 1 file provide resident population estimates for North Carolina counties. Differences in county population estimates between the two files range from a low of only five (Gates County) to a high of 6,219 residents (Wake County – a difference of less than one percent). More than half of the counties (53) had higher 2005 population totals in the State Demographer's population file, while the NCHS bridged population file revealed higher population estimates for 47 counties.

Differences in Population Estimates and the Impact on Vital Statistics Rates

The State Center for Health Statistics frequently presents vital statistics rates based on population figures. Differences in population denominators generated from the two different population files might have an impact on some of our vital statistics rates. As shown in **Table 2**, overall, the difference between rates based on the State Demographer's population file and rates based on the NCHS bridged population file is negligible with the exception of the 2005 teen pregnancy rate and breast and prostate cancer death rates. The teen pregnancy rate difference of three percent between the two files is likely a result of the differences we see in the population estimates for females as well as for the younger age groups as outlined in **Figure 1**. Caution should be taken when generating age-specific rates for these younger age groups using the two different population files. With regard to age-adjusted breast and prostate cancer mortality, differences in the rates are likely a reflection of the differences in the population estimates by gender in the two files, as well as the differences in the younger age groups.

Differences in the two files with regard to county population estimates have a small impact on county-specific vital statistics rates as well. Unadjusted mortality rates vary in some cases depending on which population denominators are used. Six counties, specifically Lee, Chatham, Orange, Swain, Hyde, and Onslow counties, have rates that differ by more than three percent depending on which 2005 population estimates are used to calculate their resident mortality rates. However, for the vast majority of counties, the difference is negligible.

An additional issue to keep in mind with the use of the NCHS bridged population file is that the file is updated every year based on the latest available information. Therefore, estimated North Carolina population figures for 2003, for example, will vary as they are regularly updated in later years. Variations in North Carolina resident population estimates from the bridged population files from 2002 to 2005 are presented in **Table 3**. North Carolina's 2001 population estimate declined by approximately 7,800 from the estimates produced in 2002 to the latest estimates produced in 2005. In contrast, the 2003 estimate increased by more than 15,000 from 2003 to 2005.

Table 2: North Carolina Resident Rates Using State Demographer's and NCHS Bridged Population Files for 2005

	Rate Based on State Demographer's Population Data	Rate Based on NCHS Bridged Population Data	Percentage Difference in the Rates
2005 Teen (ages 15-19) pregnancy rate ¹	61.7	63.6	-3.1
2005 Birth rate ²	14.2	14.2	0.0
2005 Age-adjusted mortality rate, all causes ³	884.2	873.1	1.3
2005 Age-adjusted mortality rate, heart disease ³	211.5	208.6	1.4
2005 Age-adjusted mortality rate, cancer ³	194.2	192.2	1.0
2005 Age-adjusted mortality rate, prostate cancer ⁴	25.9	25.1	3.1
2005 Age-adjusted mortality rate, breast cancer ⁵	25.8	25.4	1.6
2005 Age-adjusted mortality rate, unintentional motor vehicle injuries ³	18.7	18.8	-0.5
2005 Age-adjusted mortality rate, suicide ³	11.4	11.4	0.0

¹ Rate per 1,000 women ages 15-19

² Rate per 1,000 population

³ Rate per 100,000 population

⁴ Rate per 100,000 male population

⁵ Rate per 100,000 female population

Table 3: 2002-2005 Estimates of the 2001-2005 N.C. Resident Population from the NCHS Bridged Population Files

	2001 Population Estimate	2002 Population Estimate	2003 Population Estimate	2004 Population Estimate	2005 Population Estimate
2002 Bridged Population File	8,206,105	8,320,146	n/a	n/a	n/a
2003 Bridged Population File	8,195,249	8,305,820	8,407,248	n/a	n/a
2004 Bridged Population File	8,198,256	8,311,899	8,421,190	8,541,221	n/a
2005 Bridged Population File	8,198,279	8,312,755	8,422,375	8,540,468	8,683,242

Conclusion

It is not known how long NCHS will continue to produce the bridged population data. The NCHS bridged files were originally created to bridge the gap between the old vital statistics race/ethnicity standards, where a single race was reported on vital records, and the 1997 federal OMB standards used by the 2000 Census. Once national vital statistics data are converted to the new OMB race/ethnicity standards, which involve reporting multiple races, the creation of these files could be deemed unnecessary and NCHS would no longer generate these files. Even if the files are created in the future, the single-race population estimates would not be consistent with the multiple-race data collected on vital records.

While the State Demographer's Office does update and revise their annual population estimates periodically, the State Center for Health Statistics maintains static or "closed" population files based on the July 1 certified estimates. These files are updated by the State Center for Health Statistics only when smoothed files are generated after every Census. We do not update our files with multiple revisions so that our calculated rates are as consistent as possible with data published in our annual Vital Statistics reports. However, this also means that our closed population files will not always contain the most up-to-date population data.

As stated early in this report, population figures are produced by estimation and projection methods. This is not an exact science, and published population figures

are, of necessity, simply educated estimates of an ever-changing population landscape. There may also be some degree of error in the numerators of rates calculated by the State Center for Health Statistics, even when there is a complete count. Misclassification of race on vital records, small numbers of events for some geographic areas and population groups, and other factors may result in errors in counting health events. So rates and other health measures should be considered to have a margin of error rather than being precise calculations.⁴

References

1. North Carolina State Demographics Unit. Certified July 1 North Carolina Resident Population Estimates 2001-2005. http://demog.state.nc.us/frame_start.html
2. National Center for Health Statistics. Estimates of the July 1, 2000-July 1, 2005, United States resident population from the Vintage 2005 postcensal series by year, county, age, sex, race, and Hispanic origin, prepared under a collaborative arrangement with the U.S. Census Bureau. <http://www.cdc.gov/nchs/about/major/dvs/popbridge/popbridge.htm>. August 16, 2006.
3. Buescher PA, Gizlice Z, Jones-Vessey KA. Discrepancies between published data on racial classification and self-reported race: evidence from 2002 North Carolina live birth records. *Public Health Reports* 2005; 120: 393-398.
4. Buescher PA. Problems with rates based on small numbers. *Statistical Primer*, No. 13. State Center for Health Statistics, April 1997. <http://www.schs.state.nc.us/SCHS/pdf/primer12.pdf>.



State of North Carolina
Michael F. Easley, Governor

Department of Health and Human Services
Carmen Hooker Odom, Secretary
www.ncdhhs.gov

Division of Public Health
Leah Devlin, D.D.S., M.P.H., State Health Director

Chronic Disease and Injury Section
Marcus Plescia, M.D., M.P.H., Section Chief

State Center for Health Statistics
Paul A. Buescher, Ph.D., Director
www.schs.state.nc.us/SCHS/

The N.C. Department of Health and Human Services does not discriminate on the basis of race, color, national origin, sex, religion, age or disability in employment or the provision of services.

350 copies of this public document were printed at a cost of \$130.09 or 4¢ per copy. 5/07

Department of Health and Human Services
State Center for Health Statistics
1908 Mail Service Center
Raleigh, N.C. 27699-1908