

Tools of the Trade:

COMPUTING SYNTHETIC ESTIMATES for LOCAL HEALTH DATA

Many different types of health-related data are readily available for the United States but are not available for Pennsylvania or county/local areas. This is especially true of morbidity and health prevention statistics. For example, the National Center for Health Statistics conducts several national surveys (National Health Interview Survey, National Health Examination and Nutrition Survey, National Medical Care Utilization and Expenditure Survey, et al.) that provide an abundant array of national health statistics from blood lead levels in the population to medical costs associated with musculoskeletal conditions. (An excellent reference for locating this material is the CATALOG OF PUBLICATIONS OF THE NATIONAL CENTER FOR HEALTH STATISTICS, U. S. Department of Health and Human Services, Public Health Service, National Center for Health Statistics, 3700 East-West Highway, Hyattsville, Maryland, 20782.) However, similar local data can be very difficult or impossible to locate among the output from these massive data collection efforts.

In those instances when local data are not available, synthetic estimates can be computed based on national rates. These estimates can be useful for planning purposes unless there is reason to believe that local rates would diverge widely from the national rates. Also, usually only national rates for the total population are available. However, prevalence of most health-related conditions varies considerably with age, and often with other factors such as sex and race. In these cases, a more precise synthetic estimate may be obtained by using age, age and sex, or age, sex and race specific prevalence or incidence rates.

Below is an example of how you can compute these synthetic estimates of local health data.

EXAMPLE:

Suppose you have national rates per hundred for the prevalence of hypertension but need state incidence rates, as follows. (Notice the substantial difference which race, sex, and age may make.)

Age (In Years)	White Male	White Female	Black Male	Black Female
25-34	12.2	3.2	13.4	5.8
35-44	15.2	9.9	33.2	17.4
45-54	28.6	20.1	29.3	42.9
55-64	29.7	24.4	45.7	34.2
65-74	32.7	35.0	32.1	40.0

Sum the Pennsylvania population estimates into a table having the same breakdown as the table of national rates, and use the results as follows:

1. National rate of white males age 25-34 with hypertension is 12.2 per hundred or 0.122.
2. July 1, 1986 Pennsylvania population estimate of white males age 25-34 is 865,124.
3. Synthetic estimate, using the national prevalence rate is $865,124 \times 0.122 = 105,545$ white males age 25-34 with hypertension in Pennsylvania.

Continue this process for all 20 cells of your table, until you have an estimate for each age and sex and race group, and sum them to get a total estimate.

It is important to qualify such estimates whenever you are forced to rely upon them. Completely documenting the source of the national data used to compute the local estimates should always be part of your report. Synthetic estimates should be used cautiously and sparingly but can be valuable when computed and used correctly to offer some insight into the characteristics of the local population you are studying, based on the national experience.