

References

1. Health Resources and Services Administration: Estimates and projections of black and Hispanic physicians, dentists, and pharmacists to 2010. DHHS Publication No. HRS-P-DV 86-1. Rockville, MD, May 1986, table 28, p. 60.
2. U.S. Bureau of the Census: 1970 Census of population, detailed occupations of persons by race and sex. Vol. 1, pt. 1, sec. 2, U.S. Government Printing Office, Washington, DC, 1974, table 223, p. 65.
3. U.S. Bureau of the Census: Projections of the population of the United States by age, sex and race: 1983 to 2080. Series P-25, No. 952, U.S. Government Printing Office, Washington, DC, 1984, table 6, pp. 43 and 73.
4. U.S. Bureau of the Census: Projections of the Hispanic population: 1983 to 2080. Series P-25, No. 995, U.S. Government Printing Office, Washington, DC, 1986, table 2, pp. 39 and 43.
5. Keith, S.N., Bell, R.M., Swanson, A.G., and Williams, A.P.: Effects of affirmative action in medical schools. *New Engl J Med*: 1519-1525, Dec. 12, 1985.
6. Association of American Medical Colleges: Medical school admission requirements 1982-83 and 1987-88 editions, Washington, DC, 1981 and 1986.
7. Jolly, P., Taksel, L., and Baime, D.: U.S. medical school finances. *JAMA* 256: 1570-1580, Sept. 26, 1986.
8. Association of American Medical Colleges: Summary of information on underrepresented minorities, 1986-87. Washington, DC, March 1987.
9. U.S. Department of Education, Center for Statistics: Fall enrollment colleges and universities. Washington, DC, 1987.
10. U.S. Department of Education, Center for Statistics: Digest of education statistics 1985-86. Washington, DC, 1986, tables 20, 21, and 24, pp. 22, 23, and 26.
11. U.S. Department of Education, Center for Statistics: Digest of education statistics 1981. Washington, DC, 1981, tables 24, 25, and 26, pp. 30, 31, and 32.
12. Health Resources and Services Administration: Assessment of HCOP-supported retention programs in medical schools. Contract No. HRSA-232-82-0022. Rockville, MD, 1984, table 4.2, p. 62.
13. American Dental Association, Division of Educational Measurements: Annual reports on dental education, 1979 and 1986. Chicago, IL, 1979 and 1986.
14. American Bar Association: A review of legal education in the United States, Washington, DC, 1984.
15. American Bar Association: Law schools and bar admissions requirements. Washington, DC, 1986.
16. American Association of Engineering Societies, Engineering Manpower Commission: Engineering and technology enrollments, fall 1978 and fall 1985. Washington, DC, 1979 and 1986.

The Area Resource File —a Brief Look

HOWARD V. STAMBLER

Mr. Stambler is Director, Office of Data Analysis and Management, Bureau of Health Professions, Health Resources and Services Administration, 8-41 Parklawn Bldg., 5600 Fishers Lane, Rockville, MD, 20857.

Tearsheet requests to Mr. Stambler.

ALMOST ALL RESEARCHERS, policymakers, planners, and program managers reach a point where they need more "hard" data to develop needed analyses, to conduct research, to develop or evaluate programs, or simply to identify and understand recent developments or trends. On many such occasions, however, comparable, easily accessible data are not available. In such situations, the alternatives are to mount a major effort to obtain good data in the right form quickly or to look up a few facts in a reference document and make do with them. Since neither of these options is especially attractive, efficient, or cost-effective, the Bureau of Health Professions (BHPr) of the Health Resources and Services Administration (HRSA) decided on a different approach in the early 1970s and developed the Area Resource File

(ARF). The ARF, a computerized, county-based data system which consolidates a vast array of data into a consistent, current, and compatible set of files, has been maintained by staff in the Office of Data Analysis and Management since then.

The Area Resource File is a compilation from more than 200 sources of the most useful data for assessing the nation's health care resources. The data are merged and summarized at a county level, combined into one computerized file, and then carefully documented. By doing this, the ARF shortcuts many of the steps needed to use data. The file is used for a variety of different purposes and objectives. The general utility of the ARF is reflected in hundreds of uses by and users in government agencies, universities, associations, and consulting firms. This brief look at the ARF is

intended to help potential users to understand the ARF better in order to determine if they have a need for it.

Primary data are not collected specifically for the ARF. Rather, it is a compilation of the most current data collected by others and converted into a more compatible and accessible form. The ARF normally has between 7,000 and 9,000 data elements at one time, with the precise number changing as older data are removed and new data are added. For many important types of data, time-series are maintained as far back as the data are available.

Contents of the File

Major types of data in the ARF follow:

Health professions. The ARF includes data on physicians (MDs) by specialty and activity by county for 1975, 1979, 1983, and 1985. In many cases, it is the only source of these data, as the American Medical Association does not always publish them. There are also physician specialty data by age, sex, and location of medical school. The latest available data for other health professions are also in the file, although in many instances the estimates by county are from early inventories and are somewhat dated. Health professions data amount to roughly 1,300 data elements, about one-sixth of the ARF.

Health facilities. Hospital characteristics and utilization data amount to about a third of the ARF, or between 2,500 and 2,700 data elements. All of the ARF data are at the county level (although aggregable to larger geographic levels), with most estimates being totals for short-term, general hospitals. Time-series data back to 1969 are for numbers of beds, patient days, admissions, outpatient visits, and many similar measures. Data on full-time and part-time personnel of hospitals in more than 30 health occupations are included, and generally they are the only sound data on the geographic distribution of such personnel. Also included are nursing home data on numbers of homes, beds, and residents; these are updated whenever new data are available. National Health Service Corps' sites and staffing data, and BHP health manpower shortage area designation data are also maintained.

Vital statistics. The ARF contains data on numbers of births and infant mortality since 1970. Begin-

'The Area Resource File is a compilation from more than 200 sources of the most useful data for assessing the nation's health care resources. The data are merged and summarized at a county level, combined into one computerized file, and then carefully documented.'

ning in 1981, the number of low-birth weight infants by race and sex were added. Mortality data are kept by cause, age, sex, and race, and they include a breakout of infant mortality by these factors. In addition, the National Cancer Institute's 1980 compilation of 10 years of data on cancer mortality by body site, sex, and race is on the file, along with total estimates for two prior decades—1970s and 1960s.

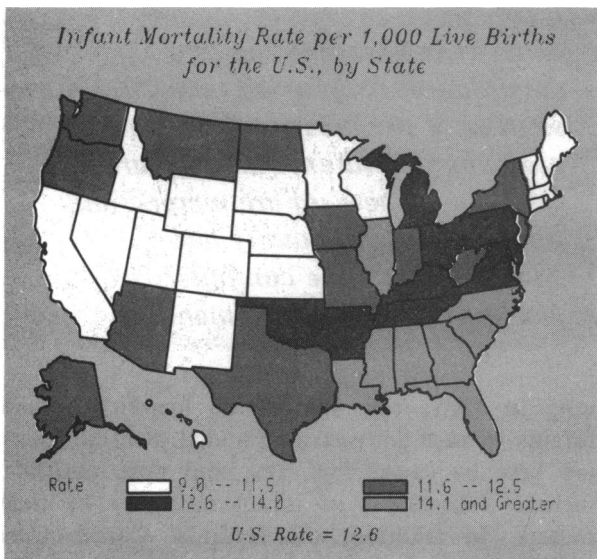
Expenditures. Medicare expenditures by county are maintained back to 1967. In addition, a weighted index of prevailing charges is available to provide a deflator for such expenditures.

Demographic data. Detailed age-, sex-, and race-specific population estimates are included for 1970, 1975, and 1980, as are single-year estimates of the total population through 1985. In the file, from the 1980 Decennial Census, is information on housing units, households and families, urban and farm population, and educational level of the population.

Economic data. The ARF has time-series data on per capita income since 1959. Income distribution measures, including mean income by race, come from the 1980 census and are for 1979. Other measures include more than 20 statistics on poverty populations in 1969 and 1979, annual unemployment and employment statistics beginning in 1960, and 1980 patterns of commuting to work.

Health professions education. For the schools of health professions, numbers of students and graduates and tuition statistics are in the file; typically they cover the 5 most recent years and 2 earlier, not consecutive, years.

In addition to the data in these major categories, there are other county measures such as land area, climate data, and crime rates, which have proved to be useful for a variety of studies. There



Source: *Area Resource File System ODAM 9/86*

are also a number of county classification variables, such as metropolitan area codes and county group codes, which make it relatively simple to aggregate county data into other major geographic groupings.

Many types of data in the ARF are updated annually but at different times of the year. To make these data available quickly, the ARF is updated and distributed twice yearly along with its 130-page technical documentation and its 220-page user documentation.

Typical Uses of ARF Data

A large number of questions, issues, and problems related to health resources can be addressed using the ARF. Some of the issues that can be investigated, using only the ARF, are these:

- the relationship of medical specialists to the availability and use of specialized units in local hospitals,
- the diffusion of physicians into areas with low physician to population ratios, and the determination of the relative impact of such factors as economic conditions, location of medical schools and graduate medical education programs, hospital facilities, climate, and other factors on changes in physician availability,
- analysis of geographic and area characteristics in relation to dental school tuition,
- the geographic incidence of infant mortality in relation to low birth weight, minority and poverty status, income levels, and physician resources, and

- the relative demand by geographic region for various associated health professionals (such as medical technicians, medical record administrators, and audiologists) and how hospitals are adjusting staffing levels following the change to payment by diagnosis-related groups while controlling for regional patterns, hospital use, and area characteristics.

The preceding are a few of the topics that can be or have been examined with just the ARF. Less technical uses of the ARF might include typifying an area and identifying similar areas throughout the nation to get a comparison panel, that is, to examine how well an area or county is doing compared with similar ones. Often, simple cross-tabulations, rankings, or trend tables can provide answers or illustrate a point.

A few examples of the special tabulations and analyses recently developed by the Bureau of Health Professions using the ARF are listed:

- profiles of counties along the U.S.-Mexico border, for the House Appropriations Committee;
- maps for a "Report on Health Professions Trends and Developments" for the HRSA Administrator;
- rankings of county population densities and their characteristics for a HRSA initiative in Primary Care Actions in Frontier Areas;
- tabulations and maps of Health Care Financing Administration-listed rural hospitals for the Office of the Secretary of the Department of Health and Human Services;
- county characteristics and hospitals of Indian Health Service hospital sites for the Indian Health Service; and
- "nursing home beds and aged population" for the Senate Committee on Aging.

Perhaps one of the most significant and extensive recent uses of the Area Resource File has been to prepare tabulations, analyses, and maps for the Secretary's Task Force on Black and Minority Health. For this effort, each county in the United States was examined for significant concentrations of minority population; "key counties" with relatively high proportions of blacks, Hispanics, Asian Pacific Islanders, or American Indians were identified; and a wide variety of relevant factors examined for each of these counties compared with other counties in particular States and the United States as a whole. Among the characteristics examined for these key counties were poverty rates, Aid to Families with Dependent Children

recipients per 100,000 population, population density, infant mortality, hospital use, urban-rural residence, physician to population ratios, and health professions schools. (The results of these analyses are described in detail in "Cross-cutting Issues on Minority Health," Volume II, "Report of the Secretary's Task Force on Black and Minority Health.")

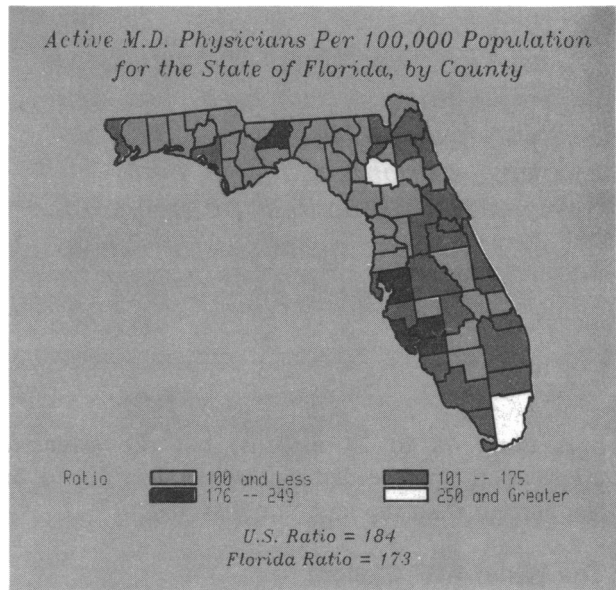
Few analytic tools are as flexible in displaying geographic patterns as carefully developed maps. As a geographic database, the ARF is easily adapted to mapping and to use with various types of software. Two examples are shown—the U.S. infant mortality rate by State and active physicians per 100,000 population by county for the State of Florida (figs. 1 and 2).

Although the actual and potential direct uses of the ARF alone are important, the ARF is equally important as a supplemental data source. When microdata have been collected or are available, other variables are often needed to complete an analysis. Typically, population variables and data to control for differences in important aspects of the environment are needed—such as the wealth of the area, its population density and population composition, and the health resources, such as physicians and hospitals. A major value of the ARF is the ready availability of these data, thus avoiding the necessity of securing tapes from diverse sources, evaluating the data for consistency, and merging them into a cohesive, coherent file. When the data make it clear to the researcher, in the middle of an analysis, that there is a need to control for another variable, that variable may very well be on the ARF and ready to use. In essence, the ARF is a data library and has many of the uses and advantages of a library.

ARF Availability and Distribution

The ARF is distributed in three forms through the National Technical Information Service (NTIS) of Springfield, VA, as a mainframe computer tape, on diskettes for use on a personal computer, and in the form of hard copy profiles (ordering information is available from the author). While attempts are continually underway to make the ARF data widely available, the limited staff of the Bureau of Health Professions cannot give special assistance to the general public beyond making the data available through NTIS.

The mainframe tape contains approximately 7,000 data elements combined into records of about 32,000 bytes. Paper documentation alone or



Source: Area Resource File System ODAM 9/86

on a second computer tape with electronic documentation is available.

In contrast, the ARF data on diskettes are much less extensive. In total, the four diskettes now available for each State and its counties have about 2,000 data elements. They are structured to make it easy to extract useful data into tables using Lotus 1-2-3. About once a year updated diskettes are added to the series, making the data more current and extending the time-series.

For those who do not have access to a computer, the ARF Profiles in paper copy or microfiche are excellent reference documents that provide detailed health measures at the State and county level. The Profiles are published as two reports for each State. "Selected Geographic Resources from the Bureau of Health Professions' Area Resource File (ARF)" has a basic profile of each county in the State. It is a 2-page summary of more than 150 measures of health resources. It also contains a ranking profile, which gives the county's relative rank against other counties in the State on more than 30 key measures of health resources and population characteristics. "Geographic Trends in Resources from the Bureau of Health Professions' Area Resource File" has four 1-page tables for each county. The first has trend data on MD physicians, the second has population and vital statistics trends, and the third has trends for all health resources. The fourth table summarizes activity, age, and sex distribution of physicians by specialty. New Profiles replace the older

'The ARF's usefulness and value stem from its filling a clear need. The ARF is cost-effective because it is far less expensive to obtain, merge, and document the hundreds of useful data sources in one systematic effort than to have hundreds of users do this repeatedly.'

ones every 18 to 24 months, but the extensive preparation time needed causes the Profiles to be less current than the tape and diskettes.

The Wider ARF System

The Area Resource File is maintained to support the analytic programs of the Bureau of Health Professions and the Health Resources and Services Administration. To support the ARF and increase its usefulness, a large number of other data files and supporting systems have been developed. Immediately supporting the ARF are four other major files:

1. facility-level hospital file,
2. county-level hospital file,
3. demographic forecast file, and
4. MD-DO file.

Also, more than 100 source data files have had standard headers added to the records to permit rapid merging. An access system has been developed that asks the user questions and then prepares the tabulation requested. It is presently a preprocessor that generates SAS programming code, but use of the SAS Application Facility to make a menu-driven access system is being considered. Most ad hoc programming is also done in SAS, and SAS is extensively used to develop mainframe graphics. An extensive graphics system supports development of custom charts and maps.

Logic of the ARF Approach

The ARF's usefulness and value stem from its filling a clear need. The ARF is cost-effective because it is far less expensive to obtain, merge, and document the hundreds of useful data sources in one systematic effort than to have hundreds of users do this repeatedly. The ARF provides opportunities as well. Many of the requests to BHP for

data require rapid response. The ARF allows the Bureau to respond to requests that would otherwise be impossible to fill if data had to be sought or merged first. The ARF also minimizes errors because its data transfers are carefully and consistently checked and routines developed with a consequent reduction in errors. The final benefit of the ARF is that users have come to recognize that its existence leads to more and better analyses of the nation's health care resources than before its development. In addition, because analysis becomes less difficult with use of the ARF, there are more analyses being done now than ever before.

The Federal Government recognizes, as does private industry, that information is a major resource that needs to be protected, managed, and used. The ARF does this for the Federal health resource programs with the assistance and cooperation of dozens of public and private data collection entities. It is a resource that should be considered for any data effort in the area of health resources.