Identifying Integrated Regions for Health Care Delivery

J. DALE TALIAFERRO, PhD, and W. W. REMMERS

THE analysis of health care delivery in any geographic area requires assumptions regarding the existence of health care delivery regions. The bases on which these regions are delineated should be appropriate to the particular purposes of the analysis. These purposes also should dictate the data to be used in the analysis as well as influence the selection of analytical methods.

The methods available for determining health care regions can be divided into three categories on the basis of the relationship of each method to the relevant data.

Administrative bases. Health care delivery regions based on administrative regions are determined arbitrarily without use of data on the health care phenomenon under consideration. This choice usually is a result of administrative expedience rather than a result of rational processes. An example is the selection of political boundaries to define health care regions.

Ecological bases. When ecological bases for regionalization are used, the health care regions are a function of the users of the health services. Thus, the regions are derived or "discovered" by observation of the use of services within the area.

Optimization bases. When optimization bases are used for the regionalization of health services, data about the health care phenomenon of interest, combined with the theoretical considerations imposed by the analyst, determine the regions. With this method, regions are thus prescribed rather than discovered. The objective is to optimize health care delivery on some selected set of criteria.

These three methods of regional delineation are the most common ways health care regions are formally and informally defined. The regions are thus either imposed from without, determined from existing behavior patterns, or prescribed on the basis of relevant data and of theoretical improvements for the delivery system.

The analysis and planning of health care usually entails the use of what we define as ecological bases for determining regions. Use of ecological bases is especially likely when various public and private organizations act as facilitators and planners rather than as the actual deliverers of service. Health organizations often make regions conform to political jurisdictions and occasionally to some kind of optimization base. Regardless of the system used, however, it is important to know the existing patterns of patient behavior with respect to the distribution of persons and services in the area. These patterns could be determined by an analytical system in which some kind of reliable indicator of utilization of health services was used

---

Dr. Taliaferro, formerly director of Social Systems Research for the Kansas Regional Medical Program, Health Services and Mental Health Administration, is now employed by the Greater Kansas City Mental Health Foundation. He is currently serving as director of the diagnostic and rehabilitation unit of the Jackson County Jail in Kansas City, Mo. Mr. Remmers is the mathematical systems consultant of the Kansas Regional Medical Program.

Tearsheet requests to Dr. J. Dale Taliaferro, Jackson County Jail Diagnostic Rehabilitation Unit, P.O. Box 15458, Kansas City, Mo. 64105.

to define integrated health service regions. The system we present is an attempt to modify some existing analytical techniques so as to produce a more desirable kind of regionalization procedure.

Indicators of Use of Health Services

Complete data on the health service utilization of each person in an area would be an ideal information base for application of a regionalization system. In reality, considerably fewer data are available. Data on a single phenomenon, however, would be sufficient for most purposes if the phenomenon was a sufficiently good indicator of general patterns of utilization of services.

An indicator of such use should have the properties of availability, quantifiability, reliability, and validity. Under the present system of health care delivery in most of the United States, a large proportion of the health care services are delivered by private practitioners, and data on services so delivered are not readily available. Another major portion, however, is delivered through hospitals, and information about these services is readily available and reasonably uniform.

With respect to the place of hospitalization, origin of patients by residence, number of admissions, and number of days of hospitalization, quantification presents no problems. With the methods described in this paper, hospital activity by patient origin is used as the data base. Because the data are collected by hospitals from their billing records, we believe they can be considered reliable. The hospitalization data have only face validity as an indicator of general patterns of utilization of services. Since, however, hospital choice is affected by the choice of physician, the physician's relationship to the hospital, referral patterns, local and area loyalties, and general trade patterns, we consider that there is sufficient reason to use hospitalization patterns as an indicator of health service utilization.

For analyzing the adequacy of regional services, the planning of new services, or the planning of training programs to help health personnel meet service needs, hospital data are a reasonable choice for use in defining regions. The final test is the utility of the derived regions for the various purposes. The validity of these data as an indicator of service utilization can be tested by sampling private practitioners, nursing homes, and so forth to determine the mutual correspondence in their service patterns. There is enough apparent validity and utility in data on hospital patients' origins to merit experimentation with methods of regionalization and with the resulting regions based on these data.

Definition and Procedures

Definition of regions for health services planning and analysis should provide areas that are integrated and show some level of independence. Integration in this context means that the regions should evidence considerable intraregional dependence with regard to the provision and utilization of services. A community or county is dependent on another when it has services available and depends on the other for clients to provide adequate utilization of its available services. Obviously, there is a reciprocal dependence for provision of these services. Regional independence in this context means that interregional service linkages should be fewer and weaker than intraregional. A region that is successfully defined on an ecological basis will have intraregional interactions which are quantitatively and, in the most desirable case, qualitatively distinguishable from interregional interactions.

For exact conformity with the concept of an ecological basis, an analytical system will probably have to be used that permits regional boundaries to be defined anywhere in the area. Data acquisition and analysis, however, are generally easier if the area first is divided into small units and these units are then assigned to regions. Census tracts, townships, zip code zones, and counties are the units that are frequently used as a basis for data collection and storage.

The system presented here has been tested in the determination of regional patterns within Kansas, the county being the chosen unit. In Kansas there are 105 counties of fairly equal size. The county is the geographic unit within the State for which at present the most relevant information is already assembled. The county unit also has further utility because numerous health agencies are organized on county lines, for example, county health departments. Therefore, use of counties as the basic units for the building of regions should not greatly compromise the ecological concept and should facilitate analysis of an area for planning or evaluation of regional health programs.

Three procedures were applied to the Kansas hospitalization data to derive integrated health care regions. In all three, a matrix providing the county of origin and the county of hospitalization for all Kansas patients discharged from hospitals in one calendar year was used as the basic data (*1*). The degree of dependence of one county on another for services is indicated by the proportion, rather than the absolute number, of patients from one county who are served in the facilities of another. The basic data were thus converted to a matrix, which indicated for each county the fraction of its hospital discharges that came from each Kansas county or from out of the State. Thus, an individual entry in the matrix (M) is denoted by m_{ij}, which indicates the fraction of the patients from county I who are discharged from hospitals in county J. The matrix M was subjected to three analytical systems for partitioning the counties into regions.

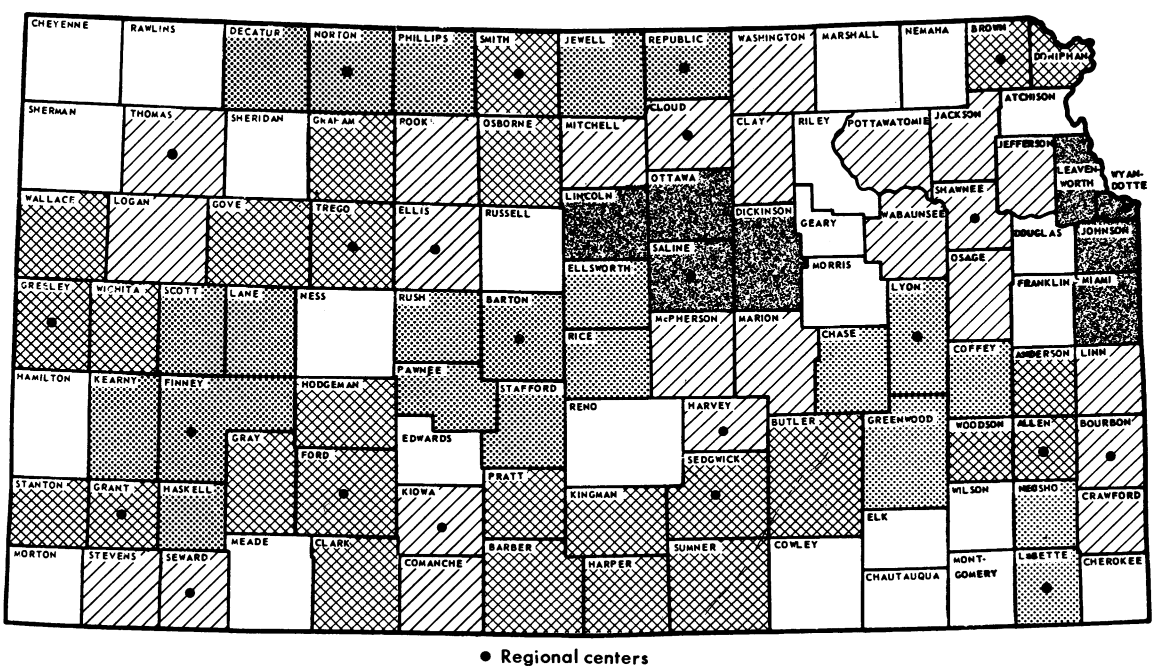
Definition by Regional Centers

In the first analytical system, a region is defined by identifying a regional center and the surrounding units that relate to it. This has been supported because it defines integrated regional areas for general economic and social planning (*2*). With this routine the counties are partitioned in the following steps:

1. The number of counties served by each county above a given cutoff level, x, is identified. Thus, if $m_{ij}>x$, county J serves county I. Services of a county to itself ($m_{ii}>x$) are included.
2. Every county that serves two or more counties as a center is identified, and the counties it serves are indicated.
3. Each county served by more than one other county is identified and assigned to the region of whatever county provides hospitalization to the largest proportion of its patients. Thus, if county I is served by counties J and K and $m_{ij}>m_{jk}$, county I will be considered to be in the region of county J and not in that of county K.
4. Any county that was a center but whose service number is reduced to less than 2 by step 3 is reassigned to another region.
5. Any center assigned to another center is identified, and all the counties that it serves are assigned to that other center.

The result of application of this routine to the data from the 1970 Kansas Hospital Association Hospital Patient Origin Study is shown in figure 1. The cutoff level used for this regionalization was 0.10. Each identified region is shown in the figure by continuing the same tone across adjacent counties. The blank counties were not assigned to a center and therefore are not defined as part of

Figure 1. Health service regions for Kansas, derived from county data by a regional center analysis



NOTE: Intercounty connections are based on service to 10 percent or more of a county's patients. Blank areas are not assigned to regions. Circles indicate regional centers.

any region. Twenty-three counties in Kansas thus were not assigned to any of the 25 defined regions.

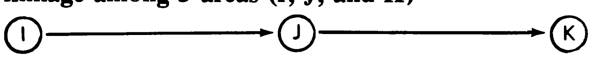
This general approach has several weaknesses for the definition of integrated health service areas. One is lack of clarity about the meaning of the unassigned counties. The unassigned counties may or may not actually be isolated (that is, independent one-county regions). Also, some patterns of service relationships across counties are meaningful that are not revealed.

These other kinds of relationships, however, can be described by means of a directed graph or digraph (*3*). The matrix *M*, or parts of it, can be represented as a digraph. Figures 2 and 3 show subgraphs of a digraph representing *M*. The circles *I*, *J*, and *K* represent counties, the lines indicate service relationships among them, and the arrows indicate the direction of those relationships. In this case the arrows indicate the direction of patient flow or of dependence for hospitalization. There will thus be a line from *I* to *J* with an arrow pointing toward *J* if $m_{ij} > x$.

The three-county chain shown in figure 2 represents a relationship in which health service affairs in county *I*, *J*, or *K* are likely to affect the other two. An increase or decrease of services for any of the three would be expected to affect the other two. The kind of link shown in this figure will not be indicated by the regional center approach. Figure 3 presents a case in which one county depends upon two others for services and thus serves to link the three counties together. The situations represented in figures 2 and 3 are not recognized by the regional center routine just described.

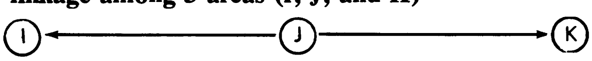
Although in many cases knowing where the service area of a center is, is useful information, defining an integrated health service region requires the recognition of other kinds of connections or relations. Other kinds of service patterns can also be used to identify regions.

Figure 2. Digraph illustrating a 1-way service linkage among 3 areas (*I*, *J*, and *K*)



NOTE: Patients travel from one area to another in the direction of the arrow for hospital services.

Figure 3. Digraph illustrating a 2-way service linkage among 3 areas (*I*, *J*, and *K*)



NOTE: Patients travel from one area to another in the direction of the arrow for hospital services.

Definition by Digraph Weak Components

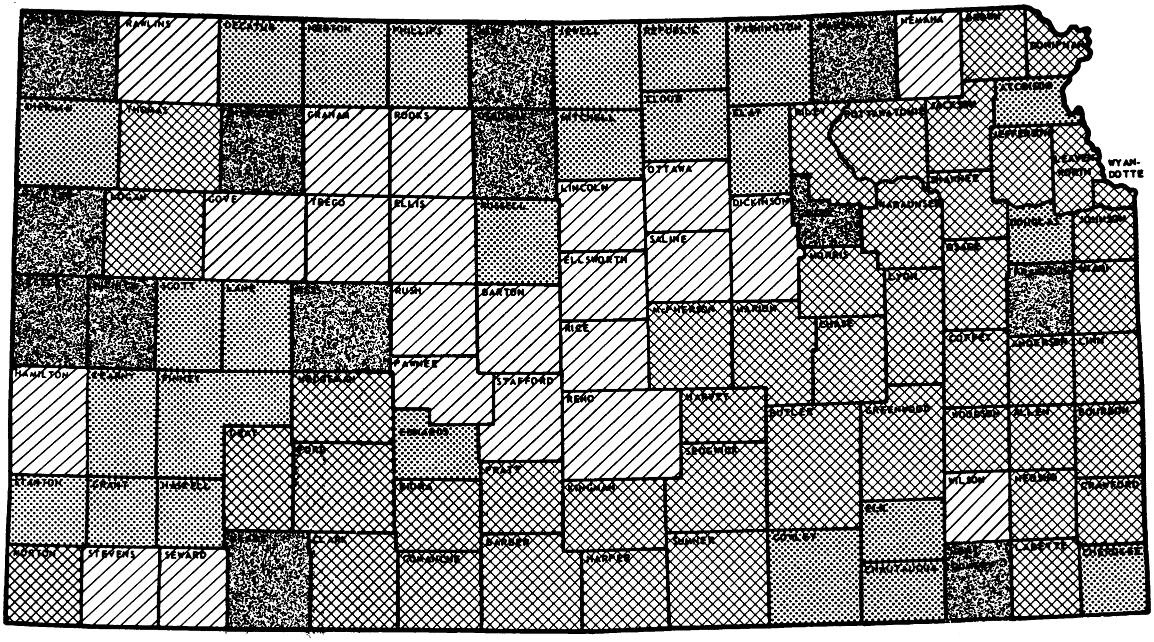
Following the definitions of graph theory (*3*), an adjacency or matrix of connectedness such as the matrix *M* already described can be partitioned into sets called "digraph weak components." An element of a weak component, in this case a county, has a path or semipath connecting it to every other element of the component and no path connecting it to any element not in the component. A path consists of a set of lines and points that can be traversed from one element to another by following the lines only in the direction of the arrows. A semipath is a set of lines and points that can be traversed from one element to another without regard to the direction of the arrows. In figure 2 there is a path from *I* to *K* and a semipath from *K* to *I*, but no path from *K* to *I*. Strong components are defined by paths rather than semipaths, but they are not useful in the present case as they require a kind of mutual or circular dependence not common in hospital service patterns. Analysis of the matrix *M* using strong components produced 103 components from the 105 Kansas counties, with a cutoff level of $x = 0.10$.

The partitioning of Kansas counties into regions as defined by digraph weak components is shown in figure 4. The cutoff level for connections between counties was $x = 0.10$. It is logical that changes in health services in any county with relationships such as those depicted in figures 2 and 3 will affect services in the other two counties. In the weak component method these relationships are recognized and made a part of the analysis. With the long semipaths that the definition of regions by this method permits (which results, for example, in a region that reaches from Doniphan County in the northeast to Clark County on the southern border), there is no guarantee that the defined region will be the kind of integrated one that is desirable for health care planning and analysis. Yet increasing the cutoff level *x* to reduce the chains results in the creation of a large number of isolated counties.

Computer Routine "Region"

The two approaches to regionalization described have some utility in viewing regional patterns, but they need to be modified if they are to produce regions with the particular characteristics

Figure 4. Health service regions for Kansas, derived from county data by a digraph weak component analysis



NOTE: Intercounty connections are based on service to 10 percent or more of a county's patients.

desired. A third system affords a special routine for defining integrated health care regions based on data on the origin of hospital patients.

"Region" provides for partitioning geographic units into integrated health care regions through a computerized routine and combines features of both of the two regionalization systems already described. The following steps are carried out, using the previously described matrix M of fractions of each county's patients who are hospitalized in each other county:

1. The number of counties served by each county, above a minimum utilization level x, is determined for each county as in the definition by regional center.

2. With the county unit having the largest number of service provision connections as a center, the county units are partitioned into sets as follows:

a. Region set (R). The region set contains the county having the regional service center and all units that are 1-connected or 2-connected to that county. A unit J is 1-connected to the center I if m_{ij} or $m_{ji} > x$, that is, if there is a line connecting I and J. A unit K is 2-connected to I if m_{ij} or $m_{ji} > x$ and m_{jk} or $m_{kj} > x$, that is, if there is a line connecting I and J, as well as a line connecting K and J.

b. Outside set (Q). The outside set consists of all units that are not l-connected or 2-connected to I.

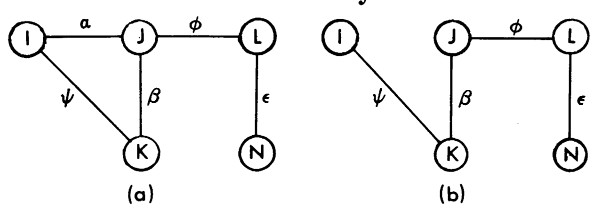
Figure 5a shows examples of the members of these sets. With unit I as the regional service center, if $\alpha > x$ and $\psi > x$, then J and K respectively are in the regional set R. If $\alpha > x$ and $\phi > x$, L is in the set R. The unit N is not 1- or 2-connected to I and is in the set Q.

3. Each member of the regional set R is considered for membership in the region around I in the following manner:

a. I is a member of the region.

b. Each 1-connected or 2-connected unit has its internal (within R) connections compared with its external (outside R) connections. Consider the unit J either 1-connected as in figure 5a or 2-

Figure 5. Digraphs illustrating 2 kinds of connectivity



NOTE: In (*a*), areas I and J are 2-connected by lines β and Ψ, through area K. In (*b*), areas I and J are 1-connected by line α, in addition to being 2-connected.

connected as in figure 5b. The strength λ of the connection of J to the set R is the value of its strongest 1- or 2-step connection to I. The strength α of the 1-connection for J in figure 5a is the maximum of m_{ij} and m_{ji}. In figure 5b there is no 1-connection of J to I. The strength of the 2-connection is determined by the point K, which provides the maximum value of ω the 2-connection strength. The value ω is the minimum of the two values making up the 2-step connection; ω is the minimum of ψ and β, where ψ is the maximum of m_{ik} and m_{ki} and β is the maximum of m_{jk} and m_{kj}. The strength of connection λ of J to the region is then the maximum of α and ω in the situation represented in figure 5a or ω in the situation represented in figure 5b.

The strength of the outside connection θ of J is the maximum value of a connection between J and an element of Q. If L is that element of Q, then $\theta =$ the maximum of m_{jl} and m_{lj}.

The county unit J is a member of the region around I if $\lambda > \theta$. If $\lambda < \theta$, J is transferred to the set Q.

c. The process described in step *b* is repeated for each element of the set R (except I) until there are no changes on one complete pass through the set.

d. Geographic continuity for the region is checked. The matrix G, where $g_{ij} = g_{ji} = 1$ if county I and J are adjacent geographically or $g_{ji} = g_{ij} = 0$ if they are not, is used. There should be a path from each county J to the center I for all counties in the region. If not, the county is removed from the region and placed in set Q. The region is then recomputed.

4. Counties assigned to regions are so labeled, and all other counties are considered through steps 2 and 3 until all counties have been assigned to a region.

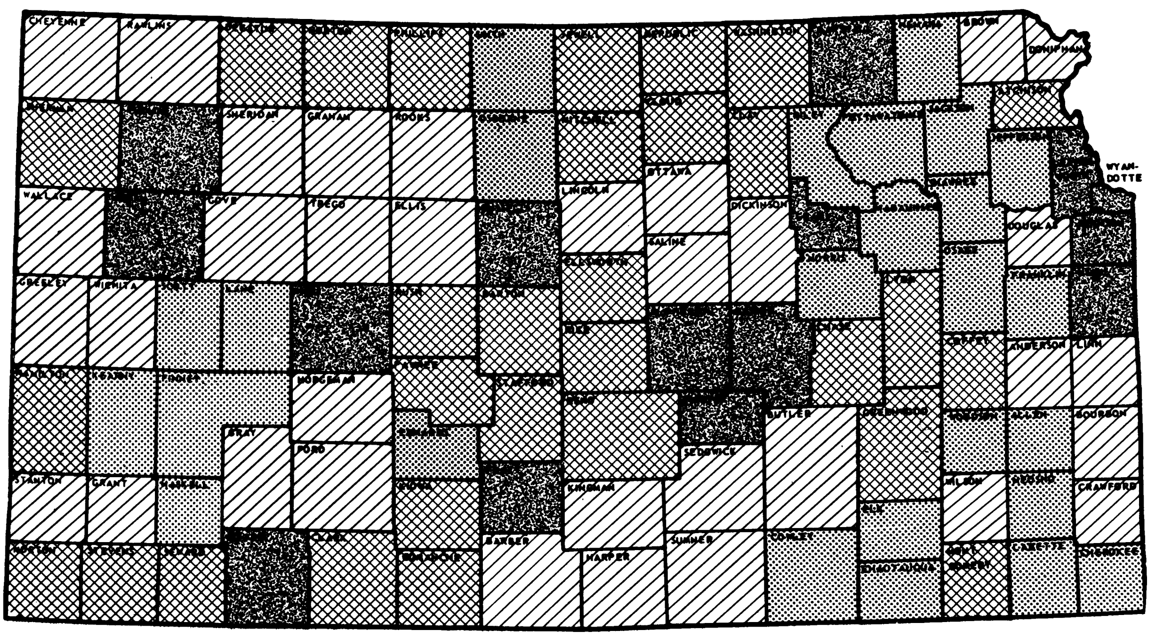
5. All isolates (regions consisting of one county) are examined for sufficient independence. There is a preset value w for the fraction of patients who must be served within the county before a county can stand as a region. Unless $m_{ii} > w$, the county I will be assigned to the geographically adjacent region containing the county that serves the largest fraction of patients from county I.

The results of application of this system to the data previously considered, using the computer routine called REGION, is shown in figure 6. The cutoff value $x = 0.10$ was used for this analysis, and $w = 0.65$ was used for the isolate determinations.

Results

The basic concept of determining health care regions on the basis of regional centers and their

Figure 6. Health service regions for Kansas, derived by use of the computer routine REGION



NOTE: Intercounty connections are based on service to 10 percent or more of a county's patients. Single-county regions provide incounty hospitalization for 65 percent or more of the county's patients.

dependent areas has been systemized to produce the regions for Kansas shown in figure 1. Expansion of the definition to incorporate more complex kinds of service linkages was accomplished by using a routine of digraph weak components in analysis. The results of such analysis, presented in figure 4, reveal that this routine permits retention of the regions defined by the regional center system of analysis; either counties are added or regions are combined. The inclusion of these more complex relationships, however, results in some large regions that cover too much area and so many centers that the high level of internal integration desired for planning and analysis cannot be achieved.

The results of analysis of the data from the Kansas Hospital Association's 1970 Patient Origin Study for Kansas residents by using the analytical system operationalized in the computer program REGION are displayed in figure 6. This map shows the integrated health care regions that were derived for Kansas. These integrated regions can be used in planning and analysis of health care services, particularly when hospital services are part of the considerations. They can also be used to anticipate difficulties and special considerations in which planning or analysis requires use of other regional boundaries. Use of the arbitrary geographic unit of a county is as consistent with the desired goals as possible since the pertinent data are available only on a county basis. The basic units for analysis could be reduced appropriately if more precise data on the patients' origins were available. The outside boundary of the area—the State's boundary—is arbitrary and produces some distortion. In the analysis presented, the border counties that use out-of-State facilities to a significant extent must also be identified. For analysis of an area for which the data are available, units outside the area that are significant could be included.

The potential general usefulness of the analytical routine operationalized by the program REGION and its applicability to other kinds of data are apparent. Further evaluation, however, should be carried out by making comparisons with other systems, including those based on "gravity" models (*4*) and on models based on multivariate systems (*5–7*). Also, the generalizability of the routine should be tested by determining the correlation between the regions it produces and those derived from other kinds of health service utilization.

REFERENCES

(*1*) Kansas Hospital Association: 1970 patient origin study. Topeka, Kans., June 1971.

(*2*) Friedman, J.: The concept of a planning region—the evolution of an idea in the United States. *In* Regional development and planning, by J. Friedman and W. Alonso. Massachusetts Institute of Technology Press, Cambridge, Mass., 1964.

(*3*) Harary, F., Norman, R. Z., and Cartwright, D.: Structural models: An introduction to the theory of directed graphs. John Wiley & Sons, Inc., New York, 1965.

(*4*) Carrothers, G. A. P.: An historical review of the gravity and potential concepts of human interaction. J Am Inst Planners 22: 94–102 (1965).

(*5*) Berry, B. J. L.: Approaches to regional analysis: A synthesis. Ann Assoc Am Geograph 54: 2–11, March 1964.

(*6*) Thompson, J. H., Sufrin, S. C., Gould, P. R., and Buck, M. A.: Toward a geography of economic health: The case of New York State. Ann Assoc Am Geograph 52: 1–29, March 1962.

(*7*) Berry, B. J. L.: A method for deriving multi-factor uniform regions. Przeglact Geografrzny 33: 263–279 (1961).

TALIAFERRO, J. DALE (Greater Kansas City Mental Health Foundation), and REMMERS, W. W.: *Identifying integrated regions for health care delivery. Health Services Reports, Vol. 88, April 1973, pp. 337–343.*

Before health services can be planned and made available to consumers, the regions for their delivery have to be defined. Any health service facility planning to develop, contract for, offer, or evaluate health services requires some knowledge about the population to be served.

Regions for health planning or analysis can be defined by (*a*) arbitrarily dividing the area under consideration into units, (*b*) devising a system idealized on some variable such as the travel time to the service, or (*c*) observing the actual behavior of the inhabitants of the area in seeking services.

A systematic analysis of the observed data is necessary to convert raw data on patients' origins into regional definitions. The mathematical procedures of graph theory provide a way of operationalizing a set of rules so as to define highly integrated regions of health care utilization.