SPINE Volume 19, Number 18S, pp 2092S–2100S
©1994, J. B. Lippincott Company

Why Does Geographic Variation in Health Care Practices Matter?

(And Seven Questions to Ask in Evaluating Studies on Geographic Variation)

Ernest Volinn, PhD,*† Paula Diehr, PhD,‡§ Marcia A. Ciol, PhD,§ and John D. Loeser, MD*†‖

One of the most active fields in health services research is the study of "geographic variation," or disparities in rates of certain types of health care practices among large areas (such as countries or regions of a country) or small areas (such as countries or hospital market areas); "small area analysis" in particular has received much attention in journals and the popular press. Increasingly, data upon which to base studies of geographic variation are becoming available. This article poses questions to ask in applying studies on geographic variation to health care settings. Because findings from these studies may ultimately affect patient care, the questions are important for physicians as well as health services researchers. The questions are:

1) What events are to be analyzed?
2) What geographic units are to be analyzed?
3) How good are the data?
4) Are differences in rates due to chance alone?
5) Are high rates too high?
6) How is geographic variation to be explained?
7) What is the role of "presentation style" in explaining geographic variation?

[Key words: small area analysis, geographic variation, hospital market area, outcomes, lumbar spine surgery] **Spine 1994;19:2092S–2100S**

For more than 50 years, studies have found great disparities in rates with which certain types of health care have been provided from area to area.[20] Such disparities, referred to here as "geographic variation," have been documented among large areas (such as countries[35] or regions with a country[56]) and among small areas (such as counties[54] or hospital market areas[62]). Figure 1, for example, depicts rates of surgery for low back pain among hospital market areas in Washington State. These rates, which range from 79 to 489 per 100,000 adults, demonstrate widespread geographic variation. Most examples in this paper pertain to surgery for low back pain, but this practice has not been selected because of its exceptional variation. Rather, it is typical for health care practices in the U.S. In a study that examined variation in rates of diagnosis-related groups (DRGs) among small areas in Maine, 47% of all other DRGs exhibited more variation in rates than the DRG for spine surgery.[62]

The phenomenon of geographic variation raises the question of its underlying explanation. One explanation is that illness rates vary in a manner that corresponds to rates of associated health care practices, but this explanation has not gained support in the literature.[54,60] Another explanation is that the clinical judgments of physicians differ in different areas because, for practices that are highly variable, there is a lack of hard outcomes data. This latter explanation, less easily dismissed, has led to what has been prominently referred to as a "revolution" in health care,[45] the outcomes movement.

The analysis of variation in health care utilization among small areas, or "small area analysis," has been a particularly active field of research and remains so despite methodologic problems that, as will be discussed, have not yet been satisfactorily resolved. Data with which to perform small area analysis are available for 20 states in the U.S. (Table 1), and more states are expected to issue such data. Software to perform small area analysis may be purchased for each of the 20 states for which data are available, permitting analysis with relative ease. Aside from the U.S., data to perform small area analysis may be obtained for many countries, including Australia,[31] Britain,[35] Canada,[49] France,[53] The Netherlands,[52] New Zealand,[2] Norway,[35] Spain,[27] and Switzerland.[51] Articles on small area analysis have appeared with increasing frequency in general medical journals as well as health services journals and by now number in the hundreds (for reviews, see references 19, 39, 61). Results of these studies have received extensive press coverage (e.g., *Consumer Reports*, July 1992[12]) and have not escaped the notice of insurance companies

From the *Multidisciplinary Pain Center and Departments of †Anesthesiology, ‡Biostatistics, §Health Services, and ‖Neurological Surgery, University of Washington, Seattle, Washington.
Supported by grant no. HS 06344 from the Agency for Health Care Policy and Research.
Financial disclosure/device statement category 3, 7.
Accepted for publication May 6, 1994.

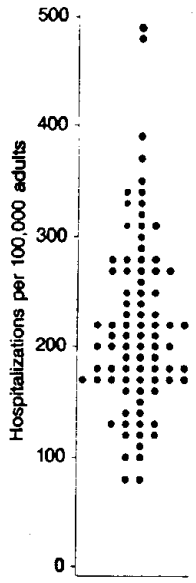

Figure 1. Surgery for low back pain age/sex adjusted rate per 100,000 persons age 20+, 86 hospital market areas, Washington State, 1989. This figure is based on hospital market areas (HMAs) as delineated by The Codman Research Group (Lebanon, NH). One HMA, Dallesport, was excluded; according to population estimates (Strategic Mapping, San Jose, CA), its age 20+ population was just over 500. The age 20+ population of the 86 included HMAs ranged from approximately 1,500 to more than 200,000.

and other payers (e.g., The Prudential Insurance Company of America placed an advertisement in *The New York Times* that reads: "Two patients with the same diagnosis who live in different communities shouldn't be treated differently."[38]).

Small area analysis became popular at a time when "cost containment" became a major goal of health care reform. If rates are considerably higher in some areas than in others, there is at least a possibility that high rates may be reduced to save on health care costs. As a booklet issued by the Agency for Health Care Policy and Research states,[1] "unexplained variations raise serious questions about the quality, appropriateness, and cost effectiveness of health care—questions now being addressed." Our purpose, however, is to point out that:

- Categories in studies of geographic variation (e.g., the practices of interest or the areas among which the practices are found to vary) may not always be the same as those used by health care practitioners. For this reason, findings from these studies may not be immediately applicable to actual health care settings.
- Increasingly, as data on the local level become available, the practices of individual physicians will come under scrutiny and will be compared with those in other areas. An obvious practical use of small area analysis, for example, would be to target physicians in an area with a high rate of a certain practice with an intervention intended to lower this rate. An understanding of how small area analyses are conducted will enable physicians to influence decisions that ultimately may affect their patients; on the other hand, unless physicians are actively involved, these decisions may be made by health care policy analysts and others at least somewhat removed from clinical decision making.

We thus pose seven questions to be asked in evaluating studies on geographic variation and, more specifically, small area analyses.

■ Question 1: What Events Are to Be Analyzed?

Hospital discharge datasets issued by states in the U.S. contain several variables on individual patients. Aside from principal procedure and principal diagnosis, most datasets contain fields with which to list two or more secondary procedures and diagnoses.

Some studies, however, may not use such detailed information on diagnosis and procedure; less refined categories may increase the number of cases in an analysis, particularly in areas with small population sizes. But in overlooking detailed information on procedure and diagnosis, different types of events may be aggregated and analyzed together as though they were a single type of event. For instance, all nonsurgical hospitalizations may be analyzed as one type of event and all surgical hospitalizations as another.[34] Other studies pertain to specific diagnosis-related groups (DRGs).[62] An example of a DRG is "back and neck procedure, age <70, without complication or comorbidity" (DRG 215). Although more specific than all nonsurgical or surgical hospitalizations, studies based on DRGs still place heterogeneous events in the same category. For instance, the DRG for spine surgery includes all locations within the spine. In addition, cases of spine surgery may be placed on a continuum from less elective (cauda equina syndrome, progressive paralysis, certain cases of malignancy, infection, and trauma) to more elective (surgery primarily intended to relieve the symptom of back pain). When the relevant physicians learn that they are practicing in an area with a high rate of spine surgery, they may want to know in some detail why the rate

Table 1. States with All-Payer Hospital Discharge Data

Arizona	New Hampshire
California	New Jersey
Colorado	New York
Florida	Oregon
Illinois	Pennsylvania
Iowa	Rhode Island
Maine	South Carolina
Maryland	Vermont
Massachusetts	Washington
Nevada	Wisconsin

Source: National Association for Health Data Organization (NAHDO), 254B North Washington Street, Falls Church, VA 22046. This list pertains to states that maintain all-payer hospital discharge data available for small area analysis. Variables and data release policies differ among states. For further information contact NAHDO.

was high. What proportion of the operations is attributable to, respectively, less elective cases and more elective cases? Physicians are also likely to be concerned about outcomes, particularly if an alternative procedure or more conservative therapy is an option. But outcome studies usually concern specific procedures performed for specific reasons in specific regions of the spine.[13,43] Studies of outcomes may not be clinically relevant unless diagnosis and procedure are specified in greater detail than a DRG.

Algorithms may be necessary to select specific types of cases from all cases in a dataset. An algorithm of this nature defines a process in which secondary as well as primary procedures and diagnoses may be used.[42,55] For instance, various combinations of procedures and diagnoses were used to select cases of what here is termed "surgery for low back pain," among them cases whose a) principal procedure was discectomy, laminectomy, or fusion, b) principal diagnosis was herniated lumbar disc, and c) secondary diagnoses did not indicate major trauma, infection, or neoplasm.[7] Cases resulting from such a process, compared with all of those within a DRG, provide a more homogeneous basis for an analysis. The selection of cases is a process that requires forethought and should be done with care, especially in analyses that may have practical implications.

■ Question 2: What Geographic Units Are to Be Analyzed?

The basis for analyzing geographic variations is usually where patients reside, as distinct from where they receive treatment. If site of treatment were to be used instead of residence, major treatment centers would have inflated rates, and it would not be possible to discern utilization rates for areas with no treatment center. Furthermore, it would not be possible to age and sex adjust rates or to ascertain how other characteristics of areas in which patients reside affect rates (explanations of geographic variations will be discussed Question 6). The patient's postal code of residence (in the U.S., zip code of residence, which is included in most state datasets) is usually sufficient to identify residence in studies of geographic variation.

A basic problem in small area analysis is how to delineate small geographic units. Different approaches to this problem may lead to seemingly discordant results. Different approaches used to assign hospitals—or the events that take place within them, i.e., hospitalizations—to small areas illustrate how discordant results may occur.

The ideal unit—one that would not be problematic—would be a small area with one or more hospitals to which all of those hospitalized from the area were admitted but none from outside the area were admitted. Under such circumstances, events could be unambiguously assigned to areas. For a particular area, the rate of a specific type of hospitalization would be simply: number of hospitalizations within the area/population of the small area.

Unfortunately (for the purposes of small area analysis), circumstances are often far more complex, especially in urban areas. There may be many hospitals clustered close together with service areas that overlap, and a high proportion of patients may travel across boundaries of areas in which they reside to hospitals elsewhere. Under these circumstances, how are events that take place within hospitals to be assigned to geographic areas?

The "plurality method" was devised to deal with this problem. This is the most commonly used method to delineate small areas in the literature. Briefly, in this method, "hospital market areas," i.e., small area units of analysis, are formulated on the basis of travel patterns of patients from geographic subunits, hereafter referred to as zip codes. If a "plurality" of patients residing in a particular zip code travel to hospital A (i.e., a higher percentage than travel to any other hospital), then the zip code is assigned to the hospital market area of hospital A. In urban areas where hospital markets overlap, a plurality of patients may be far less than a "majority" (i.e., $< 50\%$). This has led to a method whereby, in an urban area, travel patterns of residents from each zip code are correlated with travel patterns of all the other zip codes; contiguous zip codes are then consolidated into hospital market areas on the basis of the "fit" of travel patterns.[9]

Hospital market areas are usually based on travel patterns for all hospitalizations and not a specific type of hospitalization. Yet, a hospital may be renowned throughout many areas for one type of service and lack renown even within its own area for another type. The important point is that hospital market areas as constituted in many studies, while they may reflect travel patterns for all hospitalizations, may not reflect travel patterns for a specific type of hospitalization. Thus, from the perspective of those concerned with events that take place within an area's hospitals, the use of areas based on all hospitalizations may be problematic.

Can this problem be overlooked in order to proceed with the analysis? To explore this question, we used 86 hospital market areas in Washington as small area units of analysis; these had been delineated on the basis of travel patterns for all hospitalizations.[9] At the same time, however, we examined a specific type of hospitalization, surgery for low back pain. Cases of this specific type were selected with an algorithm[7] from a statewide database compiled in 1989.[58] We found that patients were unlikely to stay within the hospital market area in which they resided for surgery for low back pain. A majority stayed in only 16 of 86 hospital market areas, and a plurality stayed in only 23 hospital market areas. The implications of this incongruency between areas based on all hospitalizations and a specific type of hospitalization are:

• The small area rate of a specific type of hospitalization may be high compared with the rates of other small areas, but physicians outside the area may be primarily responsible for the high rate.
• Socioeconomic characteristics of a small area may be used to explain variation (e.g., mean per capita income), but medical supply factors strictly within a small area (e.g., physician supply per capita or hospital beds per capita) may not constitute adequate indicators of access and thus may be misleading as explanations of small area variation.

A small area unit of analysis based on the specific type of hospitalization under consideration may be preferable to one based on all hospitalizations. In formulating such a unit of analysis, it would probably be necessary to cumulate several years of data to discern stable travel patterns. Alternatively, hospital market areas based on all hospitalizations may be retained but aggregated into larger areas such that, for the specific type of hospitalization, at least a plurality of patients are hospitalized in the areas (now larger areas) in which they reside (a method similar to the one used in reference 29).

■ Question 3: How Good Are the Data?

This question[47] is important for any study that may affect patient care, including studies of geographic variation. In one study, an area with a high rate of a particular procedure (coronary-artery bypass surgery) was identified, but physicians to whom the high rate was attributed re-examined the original data source and found that the data were flawed (on many records, the zip code of the hospital was entered into the field reserved for the zip code of the patient's residence); the erroneous result was then corrected.[8]

Data accuracy should be considered for the numerator of an area's rate (in most instances, an annual count of hospitalizations) as well as the denominator (area populations). Regarding the numerator, these questions may be asked:

• Which hospitals contribute to the data? In the U.S., many discharge datasets are restricted to nonmilitary, short-stay hospitals. Among the hospitals not included are Veterans' Affairs Hospitals. Unsupplemented discharge data may yield undercounts.
• Are only those patients residing in the small areas of analysis included in the data? Patients from elsewhere (e.g., another state) may travel to the areas of analysis for treatment. Out-of-area patients should be excluded.
• Are patients who reside in the areas of analysis and travel elsewhere for treatment included? Patients who should be included in the analysis may travel to areas other than those in the analysis for treatment. For instance, Portland, Oregon, a major treatment center, is just across the border of Washington State; unless Oregon data are specifically included, patients who reside in Washington but travel to Oregon for treatment would not be included in analyses of Washington. This is a problem for many states.[37] Data from adjacent states may be available to correct for out-of-state travel (Table 1). If such data are not available, border areas may be excluded from the analysis, particularly if they are located near out-of-area treatment centers.
• Have there been recent changes in procedure or diagnosis codes used for case selection? As new diagnoses are discovered and procedures developed, preexisting coding categories may be subdivided or otherwise revised. Even though the title of *The International Classification of Diseases, Ninth Revision, Clinical Modification* has remained the same, it is occasionally updated. For instance, two procedure categories, "excision or destruction of intervertebral disc" and "spinal fusion," have been revised in recent years.
• To what extent do diagnoses and procedures in the discharge data represent the actual condition of the patient and the hospital care received? Errors in judgment, documentation, and coding may be made at various points in the process of creating patient discharge data[21,48] (Figure 2). Yet, discharge data have been independently evaluated, and, in general, the coding of invasive procedures was sufficiently accurate to conduct analyses.[46] In two studies, there was 93–97% agreement between the discharge data and the medical record on whether lumbar spine surgery has been performed.[17,42] Noninvasive procedures may be less reliably coded because coding requirements may be ambiguous (e.g., the discectomy may be coded but not the accompanying CT scan or MRI[6]). Given the multistep process required to create patient discharge data, an independent evaluation of accuracy may be advisable.

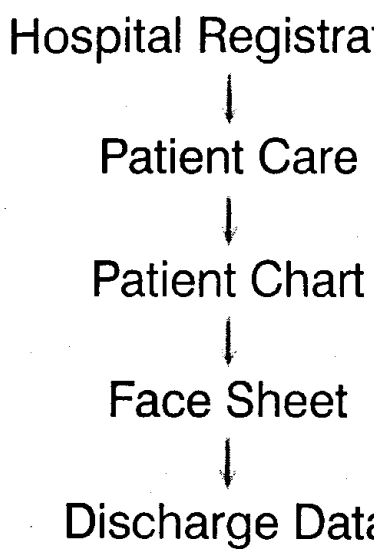

Figure 2. The process of creating hospital discharge data.

Regarding the denominator of the small area rate, questions include:

- How reliable is the population estimate? The gold-standard is based on a census count which, of course, becomes less reliable with each passing intercensus year. In the U.S., commercial firms update census estimates for small areas, but the updated estimates of various firms may be discordant.[4] States may also issue updated population data for counties and other small area units.
- Can the patient's residence be matched to a postal code or other unit with a definable population? Postal codes, created for the purposes of mail delivery, may subdivide or otherwise change. This may happen after populations of zip codes have been estimated and the data distributed. To reconcile possible inconsistencies, it is necessary to obtain the appropriate directory of postal codes.

■ Question 4: Are the Differences in Rates due to Chance Alone?

A small area analysis yields a distribution of rates, from highest to lowest. The simplest way to describe the distribution, still commonly used in the literature, is the "extremal quotient," the highest rate in the distribution divided by the lowest. Rates of surgery for low back pain depicted in Figure 1, for instance, vary from 84/100,000 to 489/100,000, or almost sixfold.

The extremal quotient, however, is probably not the best way to describe a distribution of small area rates. In any distribution of rates, the range may be surprisingly large due to chance alone, especially if the number of areas is large or the expected number of relevant hospitalizations in some areas is small (e.g., < 1 per year);[14] readmissions of the same patient for the same type of hospitalization in the same year may inflate rates and further affect the range.[14]

A better way to describe the distribution is with a summary statistic that characterizes the dispersion of the individual rates within it. The coefficient of variation and the systematic coefficient of variation have been used for this purpose,[35] but a computer simulation has shown that both of these statistics have some unfortunate properties.[14,15] Based on analysis of Washington State data, our recommendation is the CVA (coefficient of variation from analysis of variance), which is a modification of the coefficient of variation.[16] The CVA is not influenced by the prevalence of the specific type of hospitalization under consideration, and it allows for the comparison of different types of hospitalizations. The CVA also has a statistical test associated with it by which to determine whether more than chance variation has occurred.[16] However, unless the expected number of relevant hospitalizations in the smallest area is large (usually > 5), the critical values for the significance test must be determined by computer simulation.[14]

Provided that more than chance variation has occurred, a further issue is which rates are significantly lower or higher than the mean comprised of all rates in the analysis. Otherwise put, the issue is which individual rates depart more from the mean than would be expected by chance alone. The designation that a rate is lower or higher than the mean assumes that confidence levels have been adjusted for "multiple comparisons" (i.e., as the rates of more areas are compared to the mean, the likelihood of a significant departure from the mean by chance alone increases.)[10,pp. 47–51;53]

■ Question 5: Are High Rates Too High?

Studies of geographic variation may be used to identify areas with low utilization rates or those that are medically underserved. For instance, among women age 50–70, annual mammography is not as common in British Columbia as it is in Washington State, and early detection of invasive breast cancer is correspondingly less common.[24] Asthma mortality is largely preventable, but rates are highly variable among small areas of New York City.[3] These findings suggest that in some areas certain types of health services may be increased to improve the quality of care.

Given the current emphasis on cutting health care costs, however, high rates are more likely to attract attention than low rates. But even if rates are significantly high, the question still is, Are they too high?

This is a question that has been studied. Consensus panels composed of the relevant medical experts were convened to formulate criteria by which to evaluate the "appropriateness" of certain procedures. Procedures performed in areas with a wide variation in rates were then evaluated according to those criteria. Although in some instances procedures that were deemed to be inappropriate were performed more frequently in high rate areas than in moderate or low rate areas, at most such procedures explained little of the variation in rates.[5,28,50] From these studies, it may be inferred that rates in high rate areas are not necessarily too high.

Rather than inappropriate or excessive health care that is being supplied in some areas and not in others, the explanation most commonly offered for variability in rates is that physicians have different opinions about practices for which outcomes data are lacking or insufficient. Large-scale outcome studies now under way for many of these practices[44] may eventually serve as a basis for improved guidelines on appropriateness. Physicians on the local level, aware that outcome studies are for the most part conducted at major research institutions and other settings unlike their own, may also suspect that findings on outcomes do not pertain to their patients. Even after guidelines based on outcome studies have been disseminated, physicians on the local level often do not change their practice styles.[32]

For this reason, data on the local level are useful in changing physician practice style.[25] Besides utilization

rates, the issue for physicians is outcomes on the local level. Precollected data such as state discharge data (Table 1) and established methods[13] may be used to discern certain types of outcomes on the local level. These include rates of:

- Additional health care that ensues from the treatment of interest (re-operation, re-hospitalization, transfusion, discharge to nursing home).
- For surgical procedures, severe complications (pneumonia, infection, myocardial infarction, pulmonary embolus, death).

Many outcomes that are useful in clinical decision making cannot be ascertained from precollected data (quality-of-life, pain, disability, patient satisfaction, return-to-work, and impairment). Precollected data, however, allow outcomes such as those listed to be tracked for long periods at modest expense.[63]

Thus, the rate of a practice (high, medium, or low) does not in itself determine appropriateness; in many cases, outcomes are also important in making this judgment. Furthermore, outcomes on the local level will be most relevant to practitioners, especially if their outcomes can be compared to outcomes in other areas.

■ Question 6: How Is Geographic Variation To Be Explained?

This is an important question because of its policy implications. For instance, one hypothesis that may be tested in studies of geographic variation is that the supply of surgeons induces the demand for surgery or, more formally, the number of surgeons per area population is directly related to the area rate of surgery. If the supply of surgeons affects rates even after considering disease rates, an implication may be that the supply of surgical specialists is excessive. (Testing this hypothesis will be discussed shortly. Here we only note that it has received support in some studies,[22,30] but results are inconsistent.[40,49,54])

In most analyses of variation, the first step is to age and sex adjust crude utilization rates. This is done because some medical conditions are more common in certain age groups than others (while herniated disc is most common among those of working age, spinal stenosis is more common among the older population) and among one sex as opposed to the other (osteoporosis is more common among women). Small area rates may vary simply because of differences in the age and sex composition among small areas and corresponding differences in the prevalence of the relevant medical conditions. With sex and age adjustment, small area rates are modified to those that would obtain if all areas in the analysis had the same age and sex composition as a "reference population," (often an entire state or country). Depending on data that are available, different techniques may be used to adjust rates.[10,pp 47–51]

Numerous studies have attempted to explain variation in small area rates. The amount of variation accounted for wide ranges from less than 15% to more than 50%.[19,39] Common problems in these studies include:

- The ecological fallacy. Hypothetically, socioeconomic variables such as income and education affect health care utilization but, with the exception of payer (if government assistance such as Medicaid is the payer, low income is indicated), most datasets omit socioeconomic data on the individual level. In some analyses, aggregate data on socioeconomic characteristics are used instead of data on individuals, such as the proportion of the area's population receiving food stamps or mean per capita income.[34,54] Such analyses, referred to as "ecological analyses," remain only plausible and not compelling because results obtained with aggregate level data may not be generalizable to variables constructed with individual level data.[11] For instance, the proportion of a population receiving food stamps may directly relate to health care utilization,[54] but those individuals who receive food stamps are not necessarily the ones who utilize health care.
- Spurious correlation. Utilization rates have population size as their denominator, and many variables used to explain variation also consist of rates with population size as a denominator (e.g., hospital beds or surgeons per capita). In regression or correlation analyses, rates with the same denominator on both sides of the equation pose the possibility of "spurious correlation," and resulting coefficients may be artifactual rather than indicative of actual relationships.[18] To obviate this problem, it is necessary to control for population size (for techniques on how this is done, see reference 26).
- Large number of potential explanatory variables. In small area analyses, the number of geographic units is usually small but the number of variables used to explain variation may be large. If this is the case, there is a high probability in regression analyses that some results will be significant due to chance alone.
- Testing the effect of supply factors. Supply factors such as physicians per capita or hospital beds per capita may induce demand for medical services, but testing this hypothesis is not straightforward. First, there is the problem of cause-and-effect: physicians may locate themselves and hospitals may be built in areas where demand is great; even if a relationship is found between supply and utilization, the demand may not have been caused by the supply but, rather, may have preceded it. In addition, the locus of small area analysis is the small area in which patients reside; on this basis, supply factors are entered into an analysis. Yet (as discussed in "Question 2) for specific

types of hospitalizations, most patients from a small area may travel to different areas for treatment. While supply may induce demand, it may be the supply of areas outside those in which patients reside that induces the demand.

In studies that explain little of the variation in rates, the residual variation may be attributed to "physician practice style factor."[59] Yet, this has remained an elusive concept to operationalize. Specific types of hospitalizations are rare events in small areas and rates may be driven up by one or a few physicians whose volumes are anomalously high compared with other physicians in the area. In ascribing a high rate to "practice style," the question is whether the particular style pertains to a minority of physicians or to a community standard that evolves among most physicians in the area. The constituents of practice style factor also lack definition. It may refer to the training of physicians, their predisposition either to act immediately or wait watchfully, or other unspecified characteristics. The concept must be clarified if it is to be rigorously tested in an explanation of geographic variation.

■ Question 7: What Is the Role of "Presentation Style" in Explaining Geographic Variation?

"Presentation style" complements physician practice style but has received less attention in explaining variation in rates. We use the term to refer to how people express their symptoms, their expectations for care, and the disability they associate with symptoms. There may be variability in whether people with a given symptom present themselves to physicians or other health care providers at all. Among small areas of Britain, for instance, the prevalence of back pain varied relatively little, although health care seeking because of back pain was highly variable.[57] Once health care has been sought, presentation style may be particular to a locale.[23] Residential areas to a large extent may be segregated on the basis of race, ethnicity, or class,[33] and because the expression of symptoms may also vary according to these characteristics,[41,65] area rates of health care utilization may be affected.

■ Conclusion

The study of geographic variation is a "hot topic" in health care research. As is characteristic of hot topics,[36] it has attracted active researchers whose findings appear frequently in major journals. They have conducted methodological as well as substantive research. The search for an explanation of geographic variation has opened up new fields, most notably the study of outcomes. These are studies that, furthermore, have practical implications for health care policy. As suggested, however, the study of geographic variation is still a developing field. Important questions within it remain to be answered.

Acknowledgment

The authors thank Sandra Coke for excellent clerical and computer graphic assistance.

References

1. Agency for Health Care Policy and Research. AHCPR: Purpose and Programs. Washington, DC: U.S. Department of Health and Human Services, Public Health Service, September, 1988;1.
2. Brown LJ, Barnett JR. Influence of bed supply and health care organization on regional and local patterns of diabetes related hospitalization. Soc Sci Med 1992;35:1155–1170.
3. Carr W, Zeitel L, Weiss K. Variations in asthma hospitalizations and deaths in New York City. Am J Public Health 1992;82:59–65.
4. Chapman J. Cast a critical eye: Small area estimates and projections sometimes can be dramatically different. Am Demograph 1987;9:30–33.
5. Chassin MR, Kosecoff J, Park RE, et al. Does inappropriate use explain geographic variations in the use of health care services? A study of three procedures. JAMA 1987;258:2533–2537.
6. Cherkin DC, Deyo RA. Non-surgical hospitalization for low back pain: Is it necessary? Spine 1993;18:1728–1735.
7. Cherkin DC, Deyo RA, Volinn E, et al. Use of the International Classification of Diseases (ICD-9-CM) to identify hospitalizations for mechanical low back problems in administrative databases. Spine 1992;17:817–825.
8. Cherry JK, Carmichael DB, Shean FC, et al. Inaccurate data in "Solving the Medical Care Dilemma" (Letter). N Engl J Med 1988;319:800.
9. The Codman Research Group. Small Area Analysis: Concepts, Methodology, and Applications. Lebanon, NH: The Codman Research Group, Undated Publication.
10. Colton T. Statistics in Medicine. Boston, MA: Little, Brown, and Company, 1974.
11. Connell FA, Diehr P, Hart LG: The use of large data bases in health care studies. Ann Rev Public Health 1987;8:51–74.
12. Consumer Reports 57:435–448, July, 1992.
13. Deyo RA, Cherkin DC, Loeser JD, et al. Morbidity and mortality in association with operations on the lumbar spine. J Bone Joint Surg 1992;74-A:536–543.
14. Diehr P, Cain K, Connell F, et al. What is too much variation? The null hypothesis in small-area analysis. Health Serv Res 1990;24:741–771.
15. Diehr P, Cain KC, Kreuter W, et al. Can small-area analysis detect variation in surgery rates? The power of small-area variation analysis. Med Care 1992;30:484–502.
16. Diehr P, Cain K, Ye Z, et al. Small-area analysis: Methods for comparing several DRG's. Med Care 1993;31(Supplement):YS45–YS53.
17. Einstadter D, Kent DL, Fihn SD, et al. Variation in the rate of cervical spine surgery in Washington State. Med Care 1993;31:711–718.
18. Firebaugh G, Gibbs JP. Using ratio variables to control for population size. Sociol Methods Res 1986;15:101–117.
19. Folland S, Stano M. Small area variations: A critical review of propositions, methods, and evidence. Med Care Rev 1990;47:419–465.
20. Glover JA. The incidence of tonsillectomy in school children. Proc Roy Soc Med 1938;31:1219–1236.

21. Green J, Wintfield N. How accurate are hospital discharge data for evaluating the effectiveness of care? Med Care 1993;31:719–731.
22. Holahan J, Berenson RA, Kachavos PG. Area variations in selected Medicare procedures. Health Affairs 1990;9:166–175.
23. Horton CF. Women have headaches, men have backaches: Patterns of illness in an Appalachian Community. Soc Sci Med 1984;19:647–654.
24. Katz SJ, Larson EB, LoGerfo JP. Trends in the utilization of mammography in Washington State and British Columbia. Med Care 1992;30:320–328.
25. Keller RB, Soule DN, Wennberg JE, et al. Dealing with geographic variations in the use of hospitals: The experience of the Maine Medical Assessment Foundation orthopaedic study group. J Bone Joint Surg 1990;72A:1286–1293.
26. Kronmal RA. Spurious correlation and the fallacy of the ratio standard revisited. Journal of the Royal Statistical Society, A 1993;156:379–392.
27. Lardelli P, Masa J, Maderuelo A, et al. Infant, neonatal, postneonatal, and perinatal mortality in Spain, 1975–1984, interregional and interannual differences. Soc Sci Med 1991; 33:613–620.
28. Leape LL, Park RE, Solomon DH. Does inappropriate use explain small-area variations in the use of health care services? JAMA 1990;263:669–672.
29. Legnini MW, Thompson JD, DerSimonian R, et al. Changes in Connecticut hospital use rates: Have small area variations been affected? Inquiry 1990;27:51–60.
30. Lewis CF. Variations in the incidence of surgery. New Engl J Med 1969;281:880–884.
31. Loeser JD, Van Konkelenberg R, Volinn E, et al. Small area analysis of lumbar spine surgery in South Australia. Aust NZ J Surg 1993;63:14–19.
32. Lomas J. Words without action? The production, dissemination, and impact of consensus recommendations. Ann Rev Public Health 1991;12:41–65.
33. Massey DS, Denton NA. Trends in the residential segregation of Blacks, Hispanics, and Asians: 1970–1980. Am Soc Rev 1987;52:802–825.
34. McLaughlin CG, Normolle DP, Wolfe RA, et al. Small-area variation in hospital discharge rates: Do socioeconomic variables matter? Med Care 1989;27:507–521.
35. McPherson K, Wennberg JE, Hovind OB, et al. Small-area variations in the use of common surgical procedures: An international comparison of New England, England, and Norway. N Engl J Med 1982;307:1310–1314.
36. Merton RK. Three fragments from a sociologist's notebooks: Establishing the phenomenon, specified ignorance, and strategic research materials. Ann Rev Soc 1987;13:1–28.
37. Miller ME, Welch WP. State border crossing for Medicare hospital admissions. Med Care 1992;30:1053–1058.
38. New York Times, A11, July 13, 1993.
39. Paul-Shaheen P, Clark JD, Williams D. Small area analysis: A review of the North American literature. J Health Polit Policy Law 1987;12:741–809.
40. Peterson ME, Hollenberg JP, Szatrowski TP, et al. Geographic variations in the rates of elective total hip and knee athroplasties among Medicare beneficiaries in the United States. J Bone Joint Surg 1992;74-A:1530–1539.
41. Rainwater L. The lower class: Health, illness, and medical institutions. In: Deutscher I, Thompson EJ, editors. Encounters with the Poor. New York, NY: Basic Books, 1968;259–278.
42. Ramirez LF, Thisted R. Using a national health care data base to determine surgical complications in community hospitals: Lumbar discectomy as an example. Neurosurgery 1989;25:218–225.
43. Ramirez LF, Thisted R. Complications and demographic characteristics of patients undergoing lumbar discectomy in community hospitals. Neurosurgery 1989;25:226–230.
44. Raskin IE, Maklan CW. Medical treatment effectiveness research: A view from inside the Agency for Health Care Policy and Research. Evaluation and the Health Care Professions 1991;14:161–186.
45. Relman AS. Assessment and accountability: The third revolution in medical care. New Engl J Med 1988;319:1220–1222.
46. Roos LL, Mustard CA, Nicol JP, et al. Registries and administrative data: Organization and accuracy. Med Care 1993;31:202–212.
47. Roos LL, Roos NP, Cageorge SM. How good are the data? Reliability of one health care data bank. Med Care 1982;20:266–276.
48. Roos LL, Sharp SM, Wajda A. Assessing data quality: A computerized approach. Soc Sci Med 1989;28:175–182.
49. Roos NP. Hysterectomy: Variations in rates across small areas and across physician practices. Am J Public Health 1984;74:327–335.
50. Roos NP, Roos LL, Henteleff PD. Elective surgical rates—do high rates mean lower standards? Tonsillectomy and adenoidectomy in Manitoba. N Engl J Med 1977;297: 360–365.
51. Santos-Eggimann B, Paccaud F, Gutzwiller F. Coronary arteriography rates in Switzerland-How do they vary? Soc Sci Med 1989;28:115–120.
52. van Doorslaer EKA, van Vliet RCJA. "A built bed is a filled bed?" An empirical reexamination. Soc Sci Med 1989; 28:155–164.
53. Viel JF, Richardson ST. Lymphoma, multiple myeloma and leukemia among farmers in relation to pesticide exposure. Soc Sci Med 1993;37:771–777.
54. Volinn E, Mayer J, Diehr P, et al. Small area analysis of surgery for low-back pain. Spine 1992;17:575–581.
55. Volinn E, Turczyn KM, Loeser JD. Theories of back pain and health care utilization. Neurosurg Clin North Am 1991; 2:739–748.
56. Volinn E, Turczyn KM, Loeser JD. Patterns in low back pain hospitalizations: Implications for the treatment of low back pain in an era of health care reform. Clin J Pain 1994; 10:64–70.
57. Walsh K, Cruddas M, Coggon D. Low back pain in eight areas ofBritain. J Epidemiol Comm Health 1992;46:227–230.
58. Washington State Department of Health, Comprehensive Hospital Abstract Reporting System, 711 South Capitol Way, Olympia, WA 98504, Public Use Tape for 1989.
59. Wennberg JE. Dealing with medical practice variations: A proposal for action. Health Affairs 1984;3:6–32.
60. Wennberg JE. Population illness rates do not explain population hospitalization rates. Med Care 1987;25:354–359.
61. Wennberg JE, Gittelsohn AM. A Small Area Approach to the Analysis of Health System Performance. Washington, D.C., U.S. Government Printing Office, 1980 (DHHS Publication No. (HRA) 80-14012).

62. Wennberg JE, McPherson K, Caper P. Will payment based on diagnosis-related groups control hospital costs? N Engl J Med 1984;311:295–300.
63. Wennberg JE, Roos N, Sola L, et al. Use of claims data systems to evaluate health care outcomes: Mortality and reoperation following prostatectomy. JAMA 1987;257:933–936.
64. Wennberg JE. Small area analysis and the medical care outcome problem. AHCPR Conference Proceedings, Tuscon, Arizona, 1987, edited by L. Sechrest, E. Perrin, J. Bunker. Washington, D.C., U.S. Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research, 1990;177–206.
65. Zola IK. Socio-Medical Inquiries: Recollections, Reflections, and Reconsiderations. Philadelphia: Temple University Press, 1983;86–108.

Address reprint requests to

Ernest Volinn, PhD
Multidisciplinary Pain Center
University of Washington, RC-95
Seattle, WA 98195