Understanding the Context of Healthcare Utilization: Assessing Environmental and Provider-Related Variables in the Behavioral Model of Utilization

Kathryn A. Phillips, Kathleen R. Morrison, Ronald Andersen, and Lu Ann Aday

Objective. The behavioral model of utilization, developed by Andersen, Aday, and others, is one of the most frequently used frameworks for analyzing the factors that are associated with patient utilization of healthcare services. However, the use of the model for examining the context within which utilization occurs—the role of the environment and provider-related factors—has been largely neglected. Objective: To conduct a systematic review and analysis to determine if studies of medical care utilization that have used the behavioral model during the last 20 years have included environmental and provider-related variables and the methods used to analyze these variables. We discuss barriers to the use of these contextual variables and potential solutions.

Data Sources. The Social Science Citation Index and Science Citation Index. We included all articles from 1975–1995 that cited any of three key articles on the behavioral model, that included all articles that were empirical analyses and studies of formal medical care utilization, and articles that specifically stated their use of the behavioral model ($n = 139$).

Study Design. Design was a systematic literature review.

Data Analysis. We used a structured review process to code articles on whether they included contextual variables: (1) environmental variables (characteristics of the healthcare delivery system, external environment, and community-level enabling factors); and (2) provider-related variables (patient factors that may be influenced by providers and provider characteristics that interact with patient characteristics to influence utilization). We also examined the methods used in studies that included contextual variables.

Principal Findings. Forty-five percent of the studies included environmental variables and 51 percent included provider-related variables. Few studies examined specific measures of the healthcare system or provider characteristics or used methods other than simple regression analysis with hierarchical entry of variables. Only 14 percent of studies analyzed the context of healthcare by including both environmental and provider-related variables as well as using relevant methods.

Conclusions. By assessing whether and how contextual variables are used, we are able to highlight the contributions made by studies using these approaches, to identify variables and methods that have been relatively underused, and to suggest solutions to barriers in using contextual variables.

Key Words. Utilization behavior, behavioral model, context of healthcare utilization, methods

One of the most frequently used frameworks for analyzing patient utilization of healthcare services is the behavioral model developed by Andersen, Aday, and others. This conceptual framework uses a systems perspective to integrate a range of individual, environmental, and provider-related variables associated with decisions to seek care. The focus of our analysis is on the environmental and provider-related factors that influence utilization, which we refer to as "contextual" variables because they measure the context or milieu in which utilization occurs. Since the time that the behavioral model was first developed in 1968 (Andersen 1968), it has been extensively critiqued (Aday and Awe in press) and revised (Andersen 1995). However, its use for examining the context within which utilization occurs has not been reviewed.

The purpose of this study is to conduct a systematic review and analysis of the ways in which studies of medical care utilization that have used the behavioral model during the last 20 years have operationalized the context of healthcare. We conducted a structured literature review to examine whether studies have included environmental and provider-related variables and the methods used to analyze them.

Address correspondence and requests for reprints to Kathryn A. Phillips, Ph.D., Assistant Professor of Health Policy, Prevention Sciences Group, 74 New Montgomery #500, San Francisco, CA 94105. Dr. Phillips is also affiliated with the Institute for Health Policy Studies; Department of General Internal Medicine; and Department of Epidemiology and Biostatistics, University of California–San Francisco. Kathleen R. Morrison, M.P.H. is a Doctoral Student, Department of Health Services, School of Public Health, University of California–Los Angeles; Ronald Andersen, Ph.D. is Wasserman Professor, Department of Health Services, School of Public Health, University of California–Los Angeles; and Lu Ann Aday, Ph.D. is Professor of Behavioral Sciences and Management and Policy Sciences, School of Public Health, University of Texas–Houston. This article, submitted to *Health Services Research* on August 16, 1996, was revised and accepted for publication on August 18, 1997.

Environmental variables include characteristics of the healthcare de-livery system, external environment, and community—for example, health policies and the availability of providers. We examined these factors because they are increasingly recognized as having important yet poorly understood influences on utilization. For example, the extensive literature on small area variations demonstrates that utilization varies across areas, and numerous studies have shown that utilization patterns vary across managed care struc-tures (e.g., Miller and Luft 1994; Wennberg 1993). The role of environmental factors is a key policy issue, since researchers and policymakers are often interested in understanding the influence of health policies or organizations on utilization of healthcare services. The historical focus of utilization studies on individual factors, however, has resulted in a dearth of information on the role of environmental factors in influencing patient behavior.

Provider-related variables include patient factors that may be influenced by providers (e.g., whether individuals have a regular source of care), as well as provider characteristics that interact with patient characteristics to influence utilization (e.g., physician gender). As with environmental factors, provider-related variables are measures of the context within which utilization occurs, yet their influence has been relatively unexplored. For example, physician characteristics such as gender and training have been shown to influence utilization of preventive services (Lurie, Slater, McGovern, et al. 1993), and much of the variation in patient utilization rates may be due to physician decision making (Katz, Freund, Heck, et al. 1996; Wennberg, Barnes, and Zubkoff 1982). With the current trend toward shared decision making and the constraints on physician practices in a changing healthcare environment, it becomes increasingly critical to understand the influence of provider-related variables and the interaction of patient and provider factors.

Understanding the relationships among patient, provider, and envi-ronmental factors that influence utilization is particularly important, since these relationships are often of great interest from a programmatic and pol-icy perspective. For example, healthcare managers and policymakers want to know the relative contribution of patient, provider, and environmental factors to utilization. We therefore examined the methods used in studies that included contextual variables, focusing on those methods that facilitate the measurement and modeling of variables measured at different levels and the complex relationships between variables. We refer to these methods as "explanatory" methods, which we describe in detail in the following section.

We describe how the behavioral model ("model") has been used, using a structured literature review rather than judging the validity of the model or the

appropriateness of its use. The behavioral model is a framework for analysis rather than a mathematical model, and therefore it does not dictate the precise variables and methods that must be used. Clearly, the appropriateness of the inclusion of environmental or provider-related variables and the use of more complex methods will vary depending on the extent of prior research, the research question, the purpose of the study, and data availability. By assessing whether and how these variables and methods are used, however, we highlight the contributions made by studies using these approaches and identify variables and methods that have been relatively underused.

Understanding the factors that influence utilization is helpful in identifying reasons for differences in utilization, consumer satisfaction, and outcomes, and for formulating policies and programs that encourage appropriate utilization, discourage inappropriate utilization, and promote cost-effective care (Aday 1993). The results of this study help guide future research on utilization and outcomes by identifying the need for more creative and challenging conceptualization and innovative types of statistical analyses (Andersen 1995).

## DESCRIPTION OF THE BEHAVIORAL MODEL AND IMPLICATIONS FOR OPERATIONALIZATION

Figure 1 depicts the model used in our analysis, adapted from Andersen (1995).

Environmental Variables. The initial behavioral model, described in a 1968 research monograph, included three major categories: (1) predisposing factors (demographic, social structural, and attitudinal-belief variables); (2) enabling factors (family resources and characteristics of the community); and (3) need factors (perceived and evaluated illness) (Andersen 1968). Categories for environmental variables ("societal determinants" and the "health services system") were added shortly after the initial research monograph in a widely cited journal article (Andersen and Newman 1973).

We define environmental variables in this analysis as: (a) healthcare delivery system characteristics, (b) external environment factors, and (c) community-level enabling variables:

- Healthcare delivery system characteristics are the policies, resources, organization, and financial arrangements influencing the accessibility, availability, and acceptability of medical care services (e.g., physician supply).

Figure 1: Utilization Model
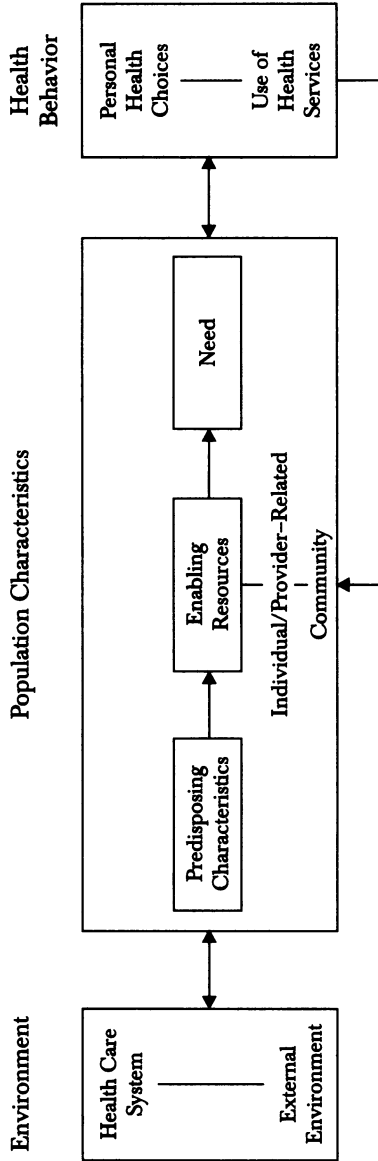
Adapted from: R. M. Andersen, 1995. "Revisiting the Behavioral Model and Access to Medical Care: Does It Matter?" *Journal of Health and Social Behavior* 36 (March): 1–10. © March 1995 by the American Sociological Association, Washington, D.C. Adapted by permission.

- External environmental factors reflect the economic climate, relative wealth, politics, level of stress and violence, and prevailing norms of the society.
- Community-level enabling variables include attributes of the community where the individual lives that enable the individual to obtain services (e.g., availability of physicians in the community) (Andersen and Davidson 1996).[1]

These variables are often measured at the aggregate level, for example, the percentage of the population that is urban within a state. However, they can also be measured at the individual level when they identify the context in which the individual "lodges," for instance, whether a patient lives in an urban or a rural area (which is a proxy for more specific measures such as availability of services).[2]

Provider-Related Variables. In contrast to environmental variables, provider-related variables are not included as a separate category in the model.[3] However, similar to the environmental variables discussed above, they measure the context within which utilization occurs. We define provider-related variables in our analysis as:

- patient factors that may be influenced by providers and which enable patients to obtain services (e.g., whether individuals have a regular source of care, the convenience of obtaining care, previous use of services, and the out-of-pocket price of services); and
- provider characteristics that interact with patient characteristics to influence utilization (e.g., specialty or gender of physician).[4]

Implications for Operationalization. The inclusion of contextual variables has implications for the methods used to operationalize the model. We therefore examined the methods used in studies that included contextual variables, focusing on those methods that facilitate the measurement and modeling of variables with different units of measurement and the complex relationships between variables. For brevity, we call these methods "explanatory methods" since they contribute to a better "explanation" of the role of contextual variables.

The inclusion of contextual variables has four key implications for the methods used to operationalize the model. First, environmental variables are often measured at the aggregate level (such as a county or HMO) while other variables in the model are measured at the individual level. Therefore, analytical techniques that take different levels into account ("contextual," "multilevel," or "hierarchical" models) may be used to better specify the

relationships among variables at different levels. (Bryk and Raudenbush 1992; Gatsonis et al. 1993; Iversen 1991). (The Appendix defines technical terms used in the article.)

Second, the components of the model are postulated to vary depending on the type of utilization being studied (Andersen 1995). This is particularly germane to provider-related variables, since patient-related factors may influence whether individuals seek care, while provider-related variables may play a more important role in influencing how much care is provided. One approach that addresses these issues is the use of two-part or multi-part equations. This approach, which was used, for example, in the RAND Health Insurance Experiment, allows for better understanding of the factors that influence initial use versus amount of use, and improves the robustness of estimates (Duan et al. 1984).

Third, the model is depicted with feedback loops ("simultaneity," "endogeneity," or "reciprocity") such that individual characteristics influence utilization and vice versa. In particular, provider-related variables such as a regular source of care often have a reciprocal relationship with utilization and therefore structural equation models or simultaneous equations (analyzed using, for example, two-stage least squares methods or LISREL) are necessary to obtain unbiased and consistent estimates. Feedback loops were added to the model more recently (Andersen 1995), although early versions of the model noted the possibility of feedback loops (Aday and Andersen 1974).

Fourth, contextual variables often have complex relationships with other variables and indirect as well as direct associations with utilization. Although utilization studies often use simple regression analysis with computer-generated stepwise or simultaneous entry of variables to analyze the correlates of utilization behavior, these methods are less useful when analyzing these complex relationships. When the goal is explanation rather than prediction, simple regression analysis may be less useful since it does not separate out the independent influence of variables or take into account the causal ordering of variables, while methods such as path analysis and theory-driven hierarchical entry of variables may be more useful.

METHODS: STRUCTURED REVIEW

Data Sources. A research librarian conducted a computerized search in June 1995 of the Science Citation Index (SCI) and Social Science Citation Index (SSCI) for all articles that cited one or more of the following key articles on the behavioral model:

1. Andersen, R. 1968. *A Behavioral Model of Families' Use of Health Services* (Research Series #25). Chicago: Center for Health Administration Studies, The University of Chicago.
2. Andersen, R., and J. F. Newman. 1973. "Societal and Individual Determinants of Medical Care Utilization in the United States." *Milbank Quarterly* 51: 95–124.
3. Aday, L. A., and R. Andersen. 1974. "A Framework for the Study of Access to Medical Care." *Health Services Research* 9 (fall): 208–220.

The first reference is the original research monograph describing the model, and the second and third references—the first key journal articles—are frequently cited when the model is used.

All articles found in either database were considered; articles were from the period 1974–1995 (SCI) or 1975–1995 (SSCI). Both indexes were searched because a preliminary search indicated that a sizable number of articles were found in only one index. The SCI includes over 3,200 journals across a wide range of scientific and technological disciplines, and the SSCI includes 1,400 social science journals and selected social science articles from more than 3,200 journals (medical journals are included in both databases, although books and non-English journals are not included).

Exclusion Criteria. After identifying 395 studies for possible inclusion, we used the following criteria for exclusion:

1. Because this study focuses on analytical approaches, we excluded studies that were not empirical analyses (e.g., reviews, letters, and conceptual papers without data analyses) ($n = 51$).
2. Only studies that focused specifically on patients' formal medical care utilization were included. Studies that were not of formal medical care (e.g., self-care, informal home care, social services, non-prescription medication), or that were not studies of medical care utilization (e.g., school problems, service awareness, health insurance coverage, perceived need) were excluded ($n = 101$).[5]
3. We excluded studies that did not specifically state in the text that they were using the behavioral model ($n = 104$). Many studies cite the model for descriptive purposes rather than using the model as a framework for their analysis, and therefore they are not relevant for the purpose of this study.

We first screened search results for inclusion using the abstract; screening was conducted by the first author with concurrence by the second author.

When the abstract was not available or sufficient information was not included in it, we screened the full study.

   Coding Methods and Criteria. We piloted the coding process by having the first two authors independently code ten studies, selected at random, that had been excluded because they were not studies of medical care utilization. All included studies were coded independently by the first two authors, with disagreements resolved by consensus. We included 139 articles and excluded 256 articles (both lists are available on request from the first author). Variable categories are described in Table 1. We further analyzed studies that included environmental *and* provider-related variables *and* that used explanatory methods to assess whether they had different characteristics from those of the other studies.

RESULTS

The first column in Table 1 shows the characteristics of the total sample. Forty-five percent of the studies included environmental variables. Among studies using environmental variables, the most frequently used were urban/rural location and variables measured at the individual level rather than the aggregate level.

   Fifty-one percent of studies included provider-related variables. The most frequently used provider-related variable was whether an individual had a regular source of care. Only two studies examined provider characteristics using data obtained from providers (i.e., physician specialty, type of practice, and quality of the patient-physician interaction) (Mirowsky and Ross 1983; Nichol, Stimmel, and Lange 1995). We did not find any studies that included the out-of-pocket price of services.

   Forty-five percent of studies that included either environmental or provider-related variables also used explanatory methods. By far the most frequently used method was hierarchical entry of variables. Only 14 percent of studies using explanatory methods analyzed feedback loops, and we found no studies that explicitly used methods such as hierarchical linear models for analyzing multilevel data.

   Thirty-eight percent of the studies collected primary data. Of studies using secondary data, the most common source was the National Health Interview Survey. Two-thirds of studies used data that were national, state, or regional in scope. Older adults were studied more often than would be expected based on their percentage of the population (possibly because of

Table 1:   Results

Characteristics	Total Sample N=139(%)	Studies with Environmental and Provider-Related Variables and Explanatory Methods N=20(%)
Number of Studies with Environmental Variables	N=62 (45)	na
Of studies measuring environment, number measuring		
Urban or rural location	39 (65)[†]	10 (50)[†]
Region	18 (29)	7 (35)
Supply (e.g., physician-to-population ratio)	16 (26)	9 (45)*
Population density (e.g., percent urban)	10 (16)	4 (20)
Healthcare system characteristics	8 (13)	4 (20)
Other (e.g., Medicaid policies, experimental site)	16 (26)	5 (25)
Of studies measuring environment, number measuring variables at the		
Individual level (e.g., whether patient lives in urban area)	37 (60)[†]	9 (45)[†]
Individual and aggregate level	17 (27)	6 (30)
Aggregate level (e.g., percent urban population)	8 (13)	5 (25)*
Number of Studies with Provider-Related Variables	N=71 (51)	na
Of studies measuring provider-related variables, number measuring		
Regular source of care	33 (46)[†]	7 (35)[†]
Type of provider or location of care (e.g., private office)	24 (34)	7 (35)
Previous use of healthcare	21 (30)	6 (30)
Convenience of care (e.g., waiting time)	16 (23)	4 (20)
Provider characteristics (e.g., specialty)	2 (03)	1 (05)
Out-of-pocket price of services	0 (00)	0 (00)
Other (e.g., physician recommendation)	19 (27)	2 (10)
Number of Studies with Explanatory Methods (among those with environmental *or* provider variables)	N=62 (45)	na
Number using		
Hierarchical entry of variables	36 (58)[†]	7 (35)[†*]
Path analysis	11 (18)	3 (15)
Methods to examine feedback loops	9 (14)	5 (25)
Two-part or multipart models	8 (13)	6 (30)*
Multilevel methods	0 (00)	0 (00)
Other	2 (03)	0 (00)
Other Study Characteristics		
Data Sources	N=139 (100)	N=20 (100)
Primary data[§]	53 (38)	4 (20)
Secondary data	86 (62)	16 (80)

continued

Table 1:   *Continued*

Characteristics	Total Sample N=139(%)	Studies with Environmental and Provider-Related Variables and Explanatory Methods N=20(%)
National Health Interview Survey (NHIS)	18 (13)	1 (05)
National Medical Expenditure Survey (NMES, NMCUES)	7 (05)	4 (20)*
Scope of Data	*N*=139 (100)	*N*=20 (100)
National/multiple states	64 (46)	11 (55)
State/region/multiple counties	30 (22)	5 (25)
Local (city or county)	45 (32)	4 (20)
Population	*N*=139 (100)	*N*=20 (100)
Elderly/Middle-aged adults	61 (44)†	7 (35)†
Women	14 (10)	3 (15)*
Children	9 (06)	2 (10)
Specific racial/ethnic groups	11 (08)	1 (05)
Other (adults not otherwise specified, veterns, etc.)	69 (50)	9 (45)
R^2 Values	*N*=65	*N*=13
Range	<.01–.87	<.01–85
Median	.19	.20

*$p < .05$, Chi-square tests comparing studies with environmental and provider variables and explanatory methods vs. studies without all three criteria (not shown).

† Percentages may add up to >100% because studies can have multiple variables or methods.

‡ Several variables fall into more than one model category: for example, travel time could be considered an environmental or provider-related variable. We coded variables such as travel time as provider-related variables when they were measured at the individual level. We considered out-of-pocket price to be a provider-related variable since it is an individual level variable that reflects characteristics of the provider and source of care; variables measuring the average charges in an area were coded as environmental.

§ Data were considered to be primary data if the authors stated that they collected data, if they cited themselves when discussing data sources, or if no reference was made to secondary data sets.

the availability of Medicare data or because older adults use more services), while relatively few studies were found of women, children, and specific racial/ethnic groups. Studies covered a range of utilization topics, including mental health services, preventive screening, dental care, and general outpatient services.

The median R^2 reported was 19 percent (range = <1%–87%, n = 65).[7] Studies that calculated the change in R^2 when environmental and/or

provider-related variables were added found that these variables increased the R^2 by a median of 13 percent (range = 2%–38%, $n = 7$), accounting for 19 percent of these studies' total R^2 (range = 8%–72%). Environmental and provider-related variables therefore may account for an important amount of the variation associated with utilization. However, few studies reported the change in R^2 for these specific sets of variables, and studies varied in the order in which these groups of variables were entered, so these results must be interpreted with caution.

The second column in Table 1 reports results for the 14 percent ($n = 20$) of studies that included environmental and provider-related variables and that used explanatory methods. Of particular interest is that these studies were more likely to include aggregate-level environmental variables (e.g., the supply of health services) and to analyze multipart models.

Table 2 describes the studies in the second column in Table 1 in more detail. As with the total sample, these studies covered a range of topics. What distinguishes these studies, however, is the conclusions they were able to draw by using contextual variables and explanatory methods. Several studies used linked databases to examine how individual utilization was associated with environmental variables such as state Medicaid policies or physician supply. Studies frequently used methods such as multipart models, which enabled them to separately assess the predictors of the probability of use and the amount of use. Studies also used methods such as hierarchical entry of variables and feedback models to assess the causal relationships among variables, often with results that contradicted other studies' findings.

DISCUSSION

Our review revealed that approximately half of the studies using the behavioral model included environmental or provider-related variables or used explanatory methods, although only 14 percent used all three. Most of the studies that included environmental variables measured urban/rural location or region, which are imprecise proxies for more specific measures such as the supply of services or access to care. Although many studies included some provider-related variables, few studies included data on provider characteristics. Most of the studies employing explanatory methods used hierarchical entry of variables, while other methods were infrequently used.

These results confirm that contextual variables using explanatory methods are less frequently examined in using the behavioral model. This can

Table 2: Description of Studies Including Environmental and Provider Variables and Using Explanatory Methods (N = 20)

First Author	Date	Research Question	Key Conclusion from Use of Contextual Variables or Explanatory Methods
Andersen	1978	To provide framework for understanding differential access and suggest ways to achieve equity	By using path analysis, found that physician density and usual source had direct effects on utilization.
Barrilleaux	1992	To examine effects of differences in state Medicaid programs on access and use of services	By combining individual and state-level data, was able to examine effects of different state Medicaid policies on utilization.
Cafferata	1987	To examine impact of marital status and living arrangements on use of formal services	By using path analysis, was able to show that effects on use of health services previously attributable to marital status are due instead to living arrangements.
Cohen	1993	To examine how fees affect use of physician and hospital services under Medicaid	By using multipart models, was able to examine factors that influence decisions to seek care, selection of site, and how much and what type of care (with the latter made by physicians).
Cunningham	1995	To examine differential effects of personal and community resources on access to care for the U.S. and Native American populations	By linking data from the Area Resource File on health manpower shortage areas, was able to demonstrate that access problems for Native Americans are obscured by national data.
Fillenbaum	1993	To examine concomitants of prescription and nonprescription drug use	By using hierarchical entry of variables, was able to demonstrate that prescription drug use was predominantly explained by health status and use of the healthcare system as well as having Medicaid and Medigap insurance.
Fosu	1989	To examine factors influencing use of clinics, drug vendors, and traditional healers in Ghana	By using hierarchical entry of variables, determined that prior contact with the healthcare system rather than current need for care was the key predictor of utilization.

Table 2:   Continued

First Author	Date	Research Question	Key Conclusion from Use of Contextual Variables or Explanatory Methods
McAuley	1984	To examine predictors of use of in-home care	By using path analysis, was able to determine that utilization of care is based on a hierarchy of need, with functional status intervening between physical and mental health and utilization.
Nichol	1995	To investigate factors that predict prescribing of multiple psychotropic medications	By examining specialty of the physician, was able to determine that patients who visited psychiatrists were more likely to obtain psychotropic medicines.
Padgett	1993	To examine the effect of insurance benefit changes on utilization of child mental health services	By using a two-part model, was able to separately examine the factors predicting the probability of use and amount of use of mental health services.
Padgett	1994	To examine factors affecting ethnic differences in utilization of outpatient mental health services	By examining environmental variables, found that ethnic congruity (the percentage of the county that was white, Latino, and African American) was associated with utilization by different ethnic groups.
Padgett	1994	To examine factors affecting ethnic differences in women's utilization of outpatient mental health services	By examining environmental variables, found that the absence of ethnic providers was a barrier to use of services.
Potvin	1995	To examine the contribution of health services utilization variables to recency of mammography	By linking databases with individual and claims data, found that patterns of previous utilization were associated with the recency of mammography.
Reisine	1987	To analyze predisposing, enabling, and need factors for utilization of dental services	By using hierarchical entry of variables, demonstrated that predisposing factors have an indirect effect on dental utilization, i.e., they affect level of need, which has a direct effect on utilization.

Table 2: *Continued*

First Author	Date	Research Question	Key Conclusion from Use of Contextual Variables or Explanatory Methods
Romeis	1991	To examine factors associated with healthcare utilization by female veterans	By using two-part models, found that the predictors of female veterans' and nonveterans' initial use and amount of use of services were similar.
Stoller	1982	To examine whether factors exert a joint influence on utilization, their relative impact, and their separate contributions	By using two-part models and provider-related variables, found that the physician/population ratio and inconvenience in getting to a physician's office were significant in predicting initial contact by not volume of visits.
Wan	1981	To determine how area characteristics and personal factors affect amount and type of utilization	By including both environmental and individual characteristics, found that both influenced the amount and type of utilization.
Wan	1987	To examine the relationships among health status, social support, and utilization	By using both recursive and non-recursive models, found that there is a complementary relationship between utilization of physician visits and social services.
Wan	1989	To examine the effect of a managed care demonstration on utilization	By using a structural equation model, found that there is a complementary relationship between ambulatory and hospital utilization.
White-Means	1989	To examine the determinants of ER use for nonemergency medical conditions	By using a two-part model (with correction for selection bias), found that, among African Americans and whites, the predictors of any utilization and the predictors for site of care are different.

lead to biased and misleading results and to large amounts of unexplained variance that hamper the understanding of utilization behavior. The exclusion of relevant variables is a key explanation given for the typically low amount of variance explained by studies using the behavioral model (Mechanic 1979).

Barriers and Solutions

Although examining contextual factors has a number of potential benefits, a number of barriers have to be overcome: (1) a lack of data, (2) analytical difficulties, and (3) model conceptualization.

   Lack of Data. A key barrier to the inclusion of contextual variables is the lack of these variables in data sets that also include individual-level utilization data. Over half of the studies rely on secondary data, most frequently national surveys conducted by the federal government that include only limited contextual variables. For example, the National Health Interview Survey (NHIS) generally includes only environmental variables, such as urban location and region, that do not allow the analysis of specific health system characteristics or policies, and it generally does not make available area identifiers below the SMSA level. The NHIS also does not obtain any data from providers, and it includes only limited data on the interaction of patients and providers. One example of a new national survey that does obtain data from individuals as well as providers and insurers (usually employers) is the Medical Expenditures Panel Survey (MEPS), which replaces the National Medical Expenditures Survey (NMES). This survey can be a rich source for conducting more complex studies of utilization that link individual and contextual variables.

   One solution to the lack of contextual variables is to merge databases that include detailed patient-level utilization data with databases that include environmental or provider characteristics. Increasing demand for data, spurred in part by the recent emphasis on outcomes and effectiveness research, has prompted several reviews of database linkages (Agency for Health Care Policy and Research 1991; Paul, Weiss, and Epstein 1993). Examples of linkable databases that include contextual variables include the Area Resource File, Census Bureau data, American Medical Association databases, American Hospital Association databases, and InterStudy data on HMO penetration. The use of the NHIS as the sampling "nucleus" for other surveys, such as the MEPS and the National Survey of Family Growth, will provide more opportunities to perform linked analyses. An example of the linking of databases is found in a study by Phillips, Kerlikowske, Baker, et al. (1998) on screening mammography utilization. This study linked individual-level

utilization data from the National Health Interview Survey with mammography facility data from the 1992 National Survey of Mammography Facilities, county-level data on 1990 HMO market share (Baker 1995); and county-level data on the supply of primary care providers (the 1991 Area Resource File).

There are a number of obstacles, however, to linking databases. Linkages require personal identifiers such as social security numbers, although confidentiality restrictions often preclude the release of these identifiers. Therefore data managers rather than the end users must often perform the linkages, creating another obstacle for researchers (Paul, Weiss, and Epstein 1993). Confidentiality restrictions may also preclude the release of data when the sample size is insufficient (for example, patients in small counties when data are being linked at the county level), creating missing data problems. Data to be linked may be from different time periods or have incompatible file structures, and the computational resources necessary to create and analyze the large databases that result may be substantial (Paul, Weiss, and Epstein 1993). Efforts to link databases can be furthered by the use of common identifiers, the development of approaches for releasing identifiers without compromising confidentiality, and the design of national surveys that provide estimates at state or local levels to facilitate linkages with data at those levels. Although the efforts to address these issues are ongoing, researchers may not be aware of linkage possibilities or they may not have the requisite technical expertise.

As with environmental variables, researchers may not include provider-related variables because they are not readily available. For example, we did not find any studies in our sample that included the out-of-pocket price of services, which is a key variable from an economic perspective. Out-of-pocket price data often are not collected, and the specification and measurement of relevant prices is an additional difficult problem (Broyles and Rosko 1988). Similar problems hinder the analysis of patient knowledge of insurance coverage and provider knowledge of reimbursement policies.[8] Although many studies in our sample examined the effects of patient income and insurance coverage on utilization, these effects may be mediated by whether patients are knowledgeable about their coverage. Similarly, although several studies examined whether physicians recommended services such as mammography screening, these recommendations might have varied based on the physicians' knowledge of patients' coverage and the relevant reimbursement policies. It will be important for future studies—particularly given the proliferation of different types of managed care plans—to collect and

validate data on out-of-pocket costs and on knowledge of insurance coverage and reimbursement policies.

Another gap is the lack of studies examining provider characteristics. Mirowsky's study of the reciprocal effects of satisfaction and utilization provides an example of ways in which collecting data from both patients and providers can add to the understanding of utilization (Mirowsky and Ross 1983). By using data from both providers and patients, they were able to test their hypotheses in two different data sets and from two different perspectives that increased the robustness of their findings.

We did find, however, that one provider-related variable—previous utilization of health services—was included in 30 percent of studies and that its use has become more frequent over time. Numerous studies have found that previous utilization of services is a strong predictor of current utilization and that its inclusion increases the amount of variance explained (e.g., Pottick et al. 1995). Therefore, its inclusion can be particularly useful when the focus of the analysis is on prediction. However, several issues introduced by the inclusion of previous utilization have been largely neglected in discussions of the behavioral model. These issues include the conceptualization and interpretation of measures of previous utilization, confounding with other variables, and autocorrelation in time-series data (see, for example, Greene 1993; Stump, Johnson, and Wolinsky 1995; Wolinsky, Stump, and Johnson 1995). For example, it is unclear whether high prior use reflects greater illness levels, a greater predisposition to seek care, provider factors, or other factors. Future research needs to examine these issues in greater depth.[9]

Analytical Difficulties. Other key barriers to the use of contextual variables are analytical, particularly (1) the analysis of both aggregate-level and individual-level data, and (2) feedback loops. First, environmental variables are often measured at an aggregate level, for example, state-level policies, while utilization is generally measured at an individual level. The typical approach to dealing with variables at multiple levels is to aggregate or disaggregate all variables to the same level and then to use standard analytic methods. This approach, however, creates statistical problems (e.g., loss of power and autocorrelation) and problems of interpretation.

To solve these problems, multilevel models have been developed that allow for the simultaneous examination of variables measured at different levels. These models, which are also called contextual analysis, hierarchical linear regression analysis, and mixed (random and fixed-effects) models, offer a more robust and efficient means for analyzing contextual effects (Bryk and Raudenbush 1992; Duncan, Jones, and Moon 1996; Gatsonis

et al. 1993; Iversen 1991; Kreft 1995). These methods, however, require more complex statistical expertise and computer programming (requiring specialized software, such as HLM and ML3).[10]

Although multilevel models have been discussed in the statistical literature for some time, their use in health services research is more recent (Gatsonis et al. 1993). Multilevel models have been used for studying practice variations and for profiling hospitals and providers (e.g., Gatsonis et al. 1993; Geller, Burns, and Brailer 1996), as well as for analyzing longitudinal data (e.g., Crystal and Sambamoorthi 1996). In our sample, virtually all of the studies were conducted with the individual as the unit of analysis, and any aggregate-level environmental variables were disaggregated to the individual level. We did not find that any of the studies explicitly discussed the use of multilevel models to examine variance components (beyond simply adjusting standard errors for multiple levels). Therefore, not only should future studies collect relevant contextual variables, they should also consider analyzing these variables using multilevel models.

A second analytical barrier is created by the presence of feedback loops. An example illustrates a reason why feedback loops are an issue and ways in which they can be analyzed. The most commonly used provider-related variable—having a regular source of care—is usually assumed to have a one-directional relationship with utilization. However, the relationship may be two-directional; heavy users are more likely to have a usual source of care, and therefore studies that assume a one-directional relationship can be biased and misleading. Kuder compared the results from three models: (1) a model based on the assumption that the relationship is one-directional (i.e., a single-equation model with usual source as a determinant of visits); (2) a model based on the assumption that the relationship is one-directional but reversed (i.e., a single-equation model with usual source as a result of visits); and (3) a model based on the assumption that a two-directional relationship exists between usual source and utilization, analyzed using a two-stage least squares method (Kuder and Levitz 1985). He found evidence that the relationship between usual source and utilization is two-directional, which is an important finding for assessing the effects of programs and policies to increase access.

In general, methods other than simple regression analysis may be infrequently used because they can require separate or more complex statistical computer software. Further, methods derived from economics, such as two-stage least squares and multipart models, may be less familiar to researchers who use the behavioral model and they can present difficult and controversial analytical issues (e.g., the choice of instrumental variables and correction for

selection bias). The lack of analysis of feedback loops, in particular, may be due partly to their only recent explication in the model, and future research will be needed to expand and test their conceptualization.

Model Conceptualization. The lack of environmental variables may reflect confusion over the model's conceptualization, as there is some overlap between environmental and enabling variables and researchers may not be aware that there is a distinct environmental component in the model. Because the model has become well known for its "predisposing, enabling, and need" components, it appears that the environmental component may be overlooked by many researchers.

One explanation given by authors for the lack of environmental measures is that their study focuses on populations where the environment varies little, for example, individuals in one city or members of a group HMO. When environmental variables are not included because little variance is expected, this should be explicitly stated. However, two-thirds of the studies used data sources that were not local in scope. Further, many research questions of interest require analyses of different environments and organizational structures; therefore, future research should be designed to collect and analyze the appropriate environmental variables.

An explanation for the lack of provider-related variables is that, although the behavioral model explicitly includes provider-related variables, it is a model of patient utilization and it does not explicitly incorporate provider characteristics per se. Future research should elucidate the conceptual framework used to examine the role of provider characteristics and provider-patient interactions. Finally, it would be useful for more studies to examine both conceptually and statistically the contribution of environmental and provider-related variables to the understanding of utilization behavior.

Limitations

Our findings should be considered in light of several limitations. First, although we used two computerized databases and three citations to locate studies as well as two reviewers to screen studies, we may have missed or inappropriately excluded some studies. Second, our study covers formal medical care utilization only, and therefore our results may not be generalizable to studies using the model for other topics. Third, some of the studies were difficult to code, due to lack of information in the articles, or were subject to different interpretations. Finally, it was beyond the scope of this study to judge the quality of the studies or the appropriateness of their operationalization. Future research could examine these issues.

FUTURE RESEARCH AND CONCLUSIONS

This study has pointed out several promising areas for future research by in-cluding environmental and provider-related variables and by using explana-tory methods. Many critical health policy issues require an understanding of the context of utilization behavior and of interactions among individuals, providers, and the healthcare environment. By assessing whether and how contextual variables are used, we highlighted the contributions made by stud-ies using these approaches, identified variables and methods that have been relatively underused, and suggested solutions to barriers in using contextual variables. A better understanding of the context of utilization behavior will require studies that are designed to examine contextual factors; the linking of databases with patient, provider, and environmental variables; the devel-opment of methods and software to allow more sophisticated analyses; and the incorporation of variables and methods from other disciplines such as economics. Our study suggests that the measurement and analysis of the context of utilization can improve the understanding of utilization behavior and thereby can contribute to more effective health programs and policies.

ACKNOWLEDGMENTS

APPENDIX

Definition of Terms

Endogenous variables: Variables that are jointly determined from within the model. These are in contrast to exogenous variables, which are predetermined from outside the model (Greene 1993).

Contextual analysis: An extension of multiple regression analysis that includes interac-tions of variables from different levels (Iversen 1991).

Multilevel analysis/hierarchical linear regression: An extension of multiple regression analysis that analyzes variables from different levels simultaneously in order to account for autocorrelation in data from a hierarchical data structure (Bryk and Raudenbush 1992; Gatsonis et al. 1993).

Hierarchical entry of variables: The entry of sets of independent variables based on an a priori theory about the expected relationships; this is in contrast to simultaneous entry

or computer-generated stepwise entry (Cohen and Cohen 1983). (Note that the term "hierarchical" is used both to indicate hierarchical entry of variables and hierarchical linear models.)

Multipart models: Using two or more interdependent regressions to analyze, for example, (1) the probability of use versus no use, and (2) among users, the amount of use (Duan et al. 1984); they may or may not include a correction for selection bias and censored data.

Path analysis: Regression models that use standardized betas to estimate direct and indirect effects (Cohen and Cohen 1983).

Structural equation modeling: A system of equations that specify the theoretical and mathematical relationships between causes (exogenous variables) and effects (endogenous variables) (Cohen and Cohen 1983).

Two-stage least squares with instrumental variables: The first stage involves the creation of an instrument; the second stage then uses that instrument for estimation. An instrument is a variable that is both correlated with the independent variables in a feedback loop and uncorrelated with the error term (Greene 1993).

Simultaneous equation models: Models in which dependent variables are jointly determined; for example, demand and supply models (Greene 1993).

NOTES

1. Community-level enabling variables could be the same as delivery system characteristics or external environmental variables with the distinction being that the level of measurement is the community.
2. Note, however, that variables *measured* at the aggregate level may be *analyzed* at either the individual or aggregate level, or both.
3. Other models derived from the behavioral model, however, have depicted provider-related variables as a separate category (e.g., Kronenfeld 1978, 1980; Shortell, Richardson, LoGerfo, et al. 1977).
4. Other individual enabling variables such as income are not included as provider-related variables since income, for example, is not influenced by the provider. In addition, although provider-related variables are reflective of organizational characteristics and therefore can be measured at the aggregate level, for the purposes of this analysis we defined provider-related variables as those measured at the individual level in order to be able to categorize variables in our analysis as either environmental or provider-related.
5. We focused on utilization because, although the model has been used for a variety of topics and has been expanded recently to include health outcomes (Andersen 1995), it is primarily used to study medical care utilization.
6. See Aday and Awe (in press) for an overview of topics covered by studies using the behavioral model.

7. Another 20 studies did not report R^2 although they used ordinary least squares regression analysis.
8. We thank an anonymous reviewer for this point.
9. We thank an anonymous reviewer for this point.
10. The concepts underlying multilevel models are described in detail elsewhere (e.g., Duncan, Jones, and Moon 1996, and the Summer 1995 issue of the *Journal of Educational and Behavioral Statistics*).

REFERENCES

Aday, L. A. 1993. "Indicators and Predictors of Health Services Utilization." In *Introduction to Health Services,* edited by S. Williams and P. Torrens. Ch. 3. Albany, NY: Delmar Publishers Inc.

Aday, L., and R. Andersen. 1974. "A Framework for the Study of Access to Medical Care." *Health Services Research* (fall): 208–20.

Aday, L. A., and W.C. Awe. in press. "Health Services Utilization Models." In *Handbook of Health Behavior Research, Vol. I. Determinants of Health Behavior: Personal and Social,* edited by D. Gochman. New York: Plenum Publishing Corporation.

Agency for Health Care Policy and Research. 1991. Report to Congress: The Feasibility of Linking Research-related Data Bases to Federal and Non-federal Medical Administrative Data Bases. Washington, DC: Government Printing Office.

Andersen, R. M. 1968. *Behavioral Model of Families' Use of Health Services,* Research Series no. 25. Chicago: Center for Health Administration Studies, University of Chicago.

————. 1995. "Revisiting the Behavioral Model and Access to Medical Care: Does It Matter?" *Journal of Health and Social Behavior* 36 (March): 1–10.

Andersen, R., and L. A. Aday. 1978. "Access to Medical Care in the U.S.: Realized and Potential. *Medical Care* 16 (4): 533–46.

Andersen, R. M., and P. L. Davidson. 1996. "Measuring Access and Trends." In *Changing the U.S. Health Care System,* edited by R. Andersen, T. Rice, and G. Kominski. San Francisco: Jossey-Bass.

Andersen, R., and J. F. Newman. 1973. "Societal and Individual Determinants of Medical Care Utilization in the United States." *Milbank Memorial Fund Quarterly* 51 (1): 95–124.

Baker, L. 1995. County-level Estimates of HMO Enrollment and Market Share, Working Paper.

Barrilleaux, C. J., and M. E. Miller. 1992. "Decisions Without Consequences: Cost Control and Access in State Medicaid Programs." *Journal of Health Politics, Policy and Law* 17 (1): 97–118.

Broyles, R. W., and M. D. Rosko. 1988. "The Demand for Health Insurance and Medical Care: A Review of the Literature." *Medical Care Review* 45 (2): 291–338.

Bryk, A. S., and S. W. Raudenbush. 1992. *Hierarchical Linear Models: Applications and Data Analysis Methods.* Newbury Park, CA: Sage Publications.

Cafferata, G. L. 1987. "Marital Status, Living Arrangements, and the Use of Health Services by Elderly Persons." *Journal of Gerontology* 42 (5): 613–18.

Cohen, J. W. 1993. "Medicaid Physician Needs and Use of Physician and Hospital Services." *Inquiry* 30 (2): 281–92.

Cohen, J., and P. Cohen. 1983. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, MI: Lawrence Erlbaum Associates, Inc.

Crystal, S., and U. Sambamoorthi. 1996. "Functional Impairment Trajectories among Persons with HIV Disease: A Hierarchical Linear Models Approach." *Health Services Research* 31 (4):470–88.

Cunningham, P. J., and L. J. Cornelius. 1995. "Access to Ambulatory Care for American Indians and Alaska Natives: The Relative Importance of Personal and Community Resources." *Social Science and Medicine* 40 (3): 393–407.

Duan, N., W. G. Manning, C. N. Morris, and J. P. Newhouse. 1984. "Choosing Between the Sample-Selection Model and the Multi-part Model." *Journal of Business & Economic Statistics* 2 (3): 283–89.

Duncan, D., K. Jones, and G. Moon. 1996. "Health-related Behaviour in Context: A Multilevel Modelling Approach." *Social Science and Medicine* 42 (6): 817–30.

Fillenbaum, G. G., J. T. Hanlon, E. H. Corder, T. Ziqubu-Page, W. E. Wall, Jr., and D. Brock. 1993. "Prescription and Nonprescription Drug Use Among Black and White Community-Residing Elderly." *American Journal of Public Health* 83 (16): 1577–82.

Fosu, G. B. 1989. "Access to Health Care in Urban Areas of Developing Societies." *Journal of Health and Social Behavior* 30 (3): 398–411.

Gatsonis, C., S. L. Normand, C. Liu, and C. Morris. 1993. "Geographic Variations of Procedure Utilization: A Hierarchical Model Approach." *Medical Care* 31 (5): YS54–YS59.

Geller, S. E., L. R. Burns, and D. J. Brailer. 1996. "The Impact of Nonclinical Factors on Practice Variations: The Case of Hysterectomies." *Health Services Research* 30 (6): 729–50.

Greene, W. H. 1993. *Econometric Analysis*. 2nd ed. New York: Macmillan Publishing Company.

Iversen, G. R. 1991. *Contextual Analysis*. Newbury Park, CA: Sage Publications.

Katz, B. P., D. A. Freund, D. A. Heck, R. S. Dittus, J. E. Paul, J. Wright, P. Coyte, E. Holleman, and G. Hawker. 1996. "Demographic Variation in the Rate of Knee Replacement: A Multi-Year Analysis." *Health Services Research* 31 (2): 126–40.

Kreft, I. G. G. 1995. "Hierarchical Linear Models: Problems and Perspectives." *Journal of Education and Behavioral Statistics* 20 (2): 109–13.

Kronenfeld, J. J. 1978. "Provider Variables and the Utilization of Ambulatory Care Services." *Journal of Health and Social Behavior* 19 (1): 68–76.

———. 1980. "Sources of Ambulatory Care and Utilization Models." *Health Services Research* (1): 4–19.

Kuder, J. M., and G. Levitz. 1985. "Visits to the Physician: An Evaluation of the Usual-Source Effect." *Health Services Research* 20 (5): 580–96.

Lurie, N., J. Slater, P. McGovern, J. Ekstrum, L. Quam, and K. Margolis. 1993.

"Preventive Care for Women: Does the Sex of the Physician Matter?" *The New England Journal of Medicine* 329 (7): 478–82.

McAuley, W. J., and G. Arling. 1984. "Use of In-Home Care by Very Old People." *Journal of Health and Social Behavior* 30 (3): 398–411.

Mechanic, D. 1979. "Correlates of Physician Utilization: Why Do Multivariate Studies of Physician Utilization Find Trivial Psychosocial and Organizational Effects?" *Journal of Health and Social Behavior* 20 (4): 387–96.

Miller, R., and H. Luft. 1994. "Managed Care Plan Performance Since 1980: A Literature Analysis." *Journal of the American Medical Association* 271: 1512–19.

Mirowsky, J., and C. E. Ross. 1983. "Patient Satisfaction and Visiting the Doctor: A Self-Regulating System." *Social Science and Medicine* 17 (18): 1353–61.

Nichol, M. B., G. L. Stimmel, and S. C. Lange. 1995. "Factors Predicting the Use of Multiple Psychotropic Medications." *Journal of Clinical Psychiatry* 56 (2): 60–66.

Padgett, D. K., C. Patrick, B. J. Burns, H. J. Schlesinger, and J. Cohen. 1993. "The Effect of Insurance Benefit Changes on Use of Child and Adolescent Outpatient Mental Health Services. *Medical Care* 31 (1): 96–110.

Padgett, D. K., C. Patrick, B. J. Burns, and H. J. Schlesinger. 1994. "Ethnicity and the Use of Outpatient Mental Health Services in a National Insured Population." *American Journal of Public Health* 84 (2): 222–26.

———. 1994. "Women and Outpatient Mental Health Services: Use by Black, Hispanic, and White Women in a National Insured Population." *Journal of Mental Health Administration* 21 (3): 347–60.

Paul, J., K. Weiss, and R. Epstein. 1993. "Data Bases for Variations Research." *Medical Care* 31 (5): YS96–YS102.

Phillips, K. A., K. Kerlikowske, L. C. Baker, S. W. Chang, and M. L. Brown. 1998. "Factors Associated with Women's Adherence to Mammography Screening Guidelines." *Health Services Research* 33 (1): 29–54.

Pottick, K., S. Hansell, E. Gutterman, and H. R. White. 1995. "Factors Associated with Inpatient and Outpatient Treatment for Children and Adolescents with Serious Mental Illness." *Journal of the American Academy of Child and Adolescent Psychiatry* 34 (4): 425–33.

Potvin, L., J. Camirand, and F. Beland. 1995. "Patterns of Health Services Utilization and Mammography Use Among Women Aged 50 to 59 Years in the Quebec Medicare System." *Medical Care* 33 (5): 515–30.

Reisine, S. 1987. "A Path Analysis of the Utilization of Dental Services." *Community Dental and Oral Epidemiology* 15 (1): 119–24.

Romeis, J. C., K. N. Gillespie, K. S. Virgo, and K. E. Thorman. 1991. "Female Veterans' and Nonveterans' Use of Health Care Services." *Medical Care* 29 (8): 932–36.

Shortell, S. M., W. C. Richardson, J. P. LoGerfo, P. Diehr, B. Weaver, and K. E. Green. 1977. "The Relationships Among Dimensions of Health Services in Two Provider Systems: A Causal Model Approach." *Journal of Health and Social Behavior* 18 (2): 139–59.

Stoller, E. P. 1982. "Patterns of Physician Utilization by the Elderly: A Multivariate Analysis." *Medical Care* 20 (9): 1080–89.

Stump, T. E., R. J. Johnson, and F. D. Wolinsky. 1995. "Changes in Physician Utilization over Time Among Older Adults." *Journal of Gerontology: Social Sciences* 50B (1): S45–S58.

Wan, T. T. 1987. "Functionally Disabled Elderly: Health Status, Social Support, and Use of Health Services." *Research on Aging* 9 (1): 61–42.

———. 1989. "The Effect of Managed Care on Health Services Use by Dually Eligible Elders." *Medical Care* 17 (8): 983–1001.

Wan, T. T., and J. H. Broida. 1981. "Factors Affecting Variations in Health Services Utilization in Quebec, Canada." *Socioeconomic Planning Sciences* 15 (2): 231–42.

Wennberg, J. E. 1993. "Future Directions for Small Area Variations." *Medical Care* 31 (5, Supplement): YS75–YS80.

Wennberg, J. E., W. A. Barnes, and M. Zubkoff. 1982. "Professional Uncertainty and the Problem of Supplier Induced Demand." *Social Science and Medicine* 16 (7): 811–24.

White-Means, S. I., and M. C. Thornton. 1989. "Nonemergency Visits to Hospital Emergency Rooms: A Comparison of Blacks and Whites." *Milbank Quarterly* 67 (1): 35–57.

Wolinsky, F. D., T. E. Stump, and R. J. Johnson. 1995. "Hospital Utilization Profiles Among Older Adults over Time: Consistency and Volume Among Survivors and Decedents." *Journal of Gerontology: Social Sciences* 50B (2): S88–S100.