Administrative Data Systems

1. Learning Objectives

After reviewing this chapter readers should be able to:

- Describe how the use of medical care administrative data systems has evolved to address fundamental social science questions in health and aging;
- Understand the elaborate set of control systems developed to assure that data protection standards are met; and
- Review advantages and disadvantages of research that relies upon administrative data, focusing on two major examples of longitudinal administrative data systems that exemplify the kinds of research that can be undertaken with this type of research resource.

Office of Behavioral and
Social Sciences Research

Department of Health
and Human Services

National Institutes
of Health

2. Introduction

Electronic data from health related business transactions, whether billing-related or clinical information, are increasingly used by public health and social science investigators in advancing our understanding of the determinants of health services utilization and the outcomes of services provided to different population groups.

Over the past 5 decades, data systems constructed from health related transactions have become increasingly complex, linking:
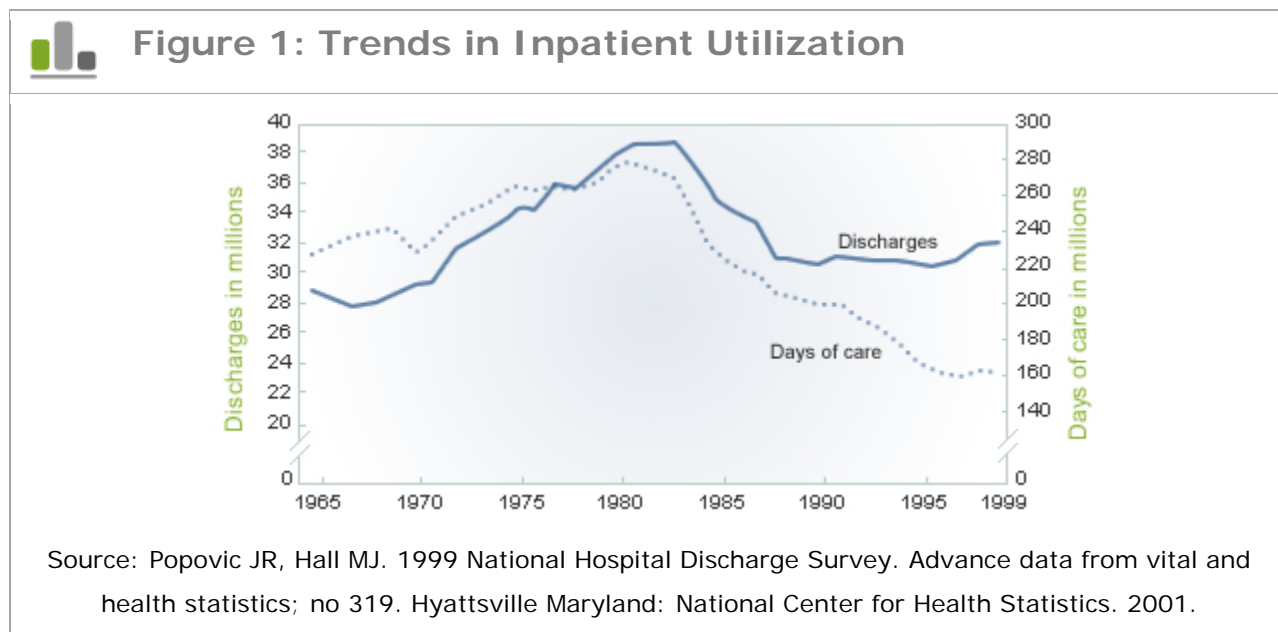
- Records about the same individuals over extended periods of time;
- Individuals' utilization hierarchicallyof medical service providers and the geographic areas in which those individuals are situated; and
- Surveys of the health and health behaviors of representative samples with medical care transaction data and vital statistics information, greatly enhancing the range of questions that can be addressed with these surveys.

3. History

One of the earliest systems available was the National Hospital Discharge Survey, created and maintained by the National Center for Health Statistics (NCHS).

Begun in 1965, it relied largely on medical record abstracting of discharges in sampled hospitals. By 1985 it was contributed to by most hospitals and states, who volunteered electronic data that was routinely generated by hospitals in a standardized format. The NCHS undertook other standing surveys of medical care utilization, such as the national ambulatory medical care survey (NAMCS) and the National Hospital Ambulatory Medical Care Survey (NHAMCS), early versions of which are all housed at the Inter-University Consortium for Political and Social Research (ICPSR). The annual summaries are accessible for research uses directly from the internet at the National Center for Health Statistics.

Figure 1 illustrates the kind of long-term data that can be derived from the National Hospital Discharge Survey. This figure shows how NHDS provides information on the changes in length of stay over a long period of time.

Figure 1: Trends in Inpatient Utilization

Source: Popovic JR, Hall MJ. 1999 National Hospital Discharge Survey. Advance data from vital and health statistics; no 319. Hyattsville Maryland: National Center for Health Statistics. 2001.

3. History

> **Medicare and Medicaid were initially enacted in 1965; like other insurance companies, Medicare used a third party payment system, where providers billed Medicare for covered services rendered to eligible individuals.**

Following the example of some other private insurers, Medicare soon computerized bills to better track expenditures. These data began to be used in the 1970's to describe the cost experience of Medicare beneficiaries and to track the utilization and costs of the new expanded benefit for End Stage Renal Disease (ESRD) (Klar, 1979; Pine, Gornick et al., 1981; Eggers, 1984). The availability of Medicaid data for research purposes has been far more sporadic and largely determined by whether states computerized their systems for their own purposes. In general, the largest states have for decades had computerized Medicaid administrative claims data, leading to extensive research that has ranged from evaluations of health policy changes to pharmaco-epidemiology (Zechnich, Greenlick et al., 1998; Dickson and Plauschinat, 2008).

Medicare data have been important in another development that has marked the historical progress in the use of medical administrative data for social science research: the emergence of databases that specifically combine survey or clinical data with administrative claims. Some examples include:

- **The Surveillance, Epidemiology, End Stage Results (SEER)** a population-based registry of all newly diagnosed cancers in selected areas around the country, linked to Medicare claims files for all eligible individuals for nearly two decades. Cancer epidemiologists and health services researchers have examined the influence of pre-existing co-morbidities on cancer treatments, documented racial and ethnic differences in treatment, and used the SEER definitions of cancer stage and progression as a means of creating "valid" claims based indicators that could then be applied to all Medicare beneficiaries (Cooper, Yuan et al., 1999; Etzioni, Riley et al., 2002; Gibbs, McLaughlin et al., 2006; Wright, Barlow et al., 2007).

- ***The NIA-funded Longitudinal Survey on Aging (LSOA),*** which obtained consent from respondents to match responses to Medicare claims, greatly facilitating research on the inter-relationship between functional decline and health care utilization (Mor, Wilcox et al., 1994; Wolinsky, Stump et al., 1995; Wolinsky, Krygiel et al., 2002).

- ***The Medicare Beneficiary Survey (MCBS)*** uses the Medicare enrollment files as a sampling frame to select beneficiaries to participate in successive waves of 3-year panel studies and all their claims data are linked to the longitudinal survey (Chulis, Eppig et al., 1993; Eppig and Chulis, 1997).

3. History

Other surveys still underway, such as the Health and Retirement Survey and planned (e.g. the renewed) National Long Term Care Survey, have adopted a similar approach of adding Medicare claims data to the survey data, thereby greatly enriching the value of individuals' responses about their health and social support. Finally, NCHS has linked various vital statistics files such as birth and death records and health surveys linked to death certificates.

> **Large non-profit health maintenance organizations, such as Kaiser Permanente, Group Health Cooperative, and the Health Insurance Plan of New York, have used their own data systems to improve their approach to health care delivery and management (Densen, 1959; Marks, 1980).**

For example, researchers at Kaiser examined drug utilization practices in an HMO vs. community setting and compared their own hospital use rates to local and national standards, designing claims-based systems to adjust for case mix acuity (Greenlick and Darsky, 1968; Hurtado and Greenlick, 1971). The Health Insurance Plan of greater New York (HIP) not only used its data to examine out-of-plan use by members but implemented one of the most ambitious randomized clinical trials of mammography for breast cancer screening that had ever been attempted (Densen, 1959; Greenfield, Densen et al., 1978; Shapiro, Venet et al., 1985).

The most recent embodiment of both the HMO and Blue Cross plans' efforts to use their data for management and research purposes is the "Marketscan" database maintained by Thompson Health Care, formerly MedStat (Adamson, 2008). This is a system of databases combining medical claims data from multiple large insurers, managed care plans, employers, and Medicare beneficiaries that is used for management research and "benchmarking" of efficiency and performance across plans, as well as by researchers addressing policy questions and examining adverse outcomes associated with prescription medicines (Holzer, Engelhart et al., 1997; Gibson, Mark et al., 2006).

Exercise 1: Research Contributions

Review the contributions each resource has made to answering social science research questions and drag the resource which most appropriately matches the contribution.

LSOA	Birth and Death Records	Health and Retirement Survey

SEER	MCBS
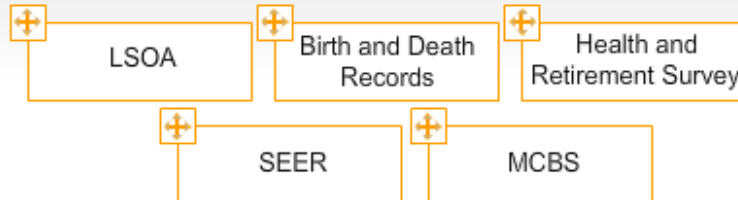
Contribution	Resource
Relatively definitive biologically determined diagnosis information. Can locate consumption behavior prior to or following event.	
First time there was a nationally representative survey sample matched to Medicare claims early on in design phase.	
Solved problem that managed care does not provide claims data. Utilization data that can be used to compare managed care and non managed care.	
Data on wealth and income to serve as predictor of future patterns of Medicare consumption.	
Has been used individually and demography but more recent linkage has improved use in social science research.	

4. Features

> **Since administrative data are by definition created for some purpose other than research, their 'input' features must be understood for researchers to properly use them.**

Level of the data. By their very nature, medical insurance based administrative data exist at the level of **the "event"**: hospitalization, physician visit, pharmacy prescription, etc. This is the level at which services are paid for following a utilization event. In some cases, the level of the data relate to an individual patient and might include repeated assessments of some aspect of the patient's clinical condition; for instance the results of annual screening and/or physical exams.

An additional level is that of **the provider organization**: the hospital, the nursing home, etc. Such data are also repeated periodically in the form of annual cost reports, inspections, etc. Medicare/Medicaid certified health care maintains and periodically updates "Provider of Service" files for many different types of providers. Cost reports, required of Medicare providers, contain data relevant to staffing, organization, services, and business practices which can then be linked to individual Medicare beneficiaries served by the patient.

Finally, all the levels of data above can be aggregated up to the level of **a geographic area**, such as a zip code, county, Hospital Service Area, Hospital Referral Region or state (Cooper, Yuan et al., 2001; Wennberg, Fisher et al., 2004a).

4. Features

Data Linkage. A key aspect of administrative data is whether the events or repeated measurements on individuals can be linked with some form of **identifier**. In some cases linkage might be restricted to certain types of medical events, e.g., hospital stays, whereas in other cases all types of services covered under the insurance policy can be linked. In instances where the identifiable data necessary to link individual billing claims are not available, it is still possible to construct meaningful provider- or area-level summaries in order to compare providers or areas in terms of the rates of hospitalization or even the use specific procedures.

Level of Clinical Detail. In the U.S. most inpatient claims contain data fields for up to 10 different diagnoses of procedures codes, using the International Classification of Diseases version 9. Physician visits and other out-patient bills often contain less detailed diagnostic information, particularly if the visit was a "well" visit. However, these claims often contain considerable detail regarding the nature of the service provided, via different **codes, scales, and assessment tools.**

- **"Current Procedural Technology" (CPT) codes**, created and maintained by the American Medical Association , are used by providers and other medical offices to bill for procedures ranging from out-patient surgery to laboratory tests, etc.
- Medicare's Resource Based - Relative Value Scale (RBRVS) can be linked to these codes to characterize the content and payment of physician visits, though the claims data are often coded to reflect incentives built into the billing system and not to perfectly reflect the clinical validity of patients' conditions (Foster, 1985; Sorrentino, 1989).
- The Minimum Data Set for nursing home resident assessment (MDS), which was mandated by Congress under the Nursing Home Reform Act, requires nursing homes to conduct a comprehensive clinical assessment that becomes the basis for a care plan which is updated at least quarterly and has been computerized since 1998 (Morris, Nonemaker et al., 1997; Mor, 2004).
- The Outcome and Assessment Information Set (OASIS) is used to assess all users of Medicare home health care when enrolled in the service, with a re-assessment performed at discharge and at least every 60 days thereafter while they remain in service (Sangl, Saliba et al., 2005).

MDS and OASIS, in particular, contain extensive clinical and functional information about each individual receiving care from nursing homes or home health agencies, which can be used to complement the diagnostic information contained in hospital and other medical claims.

4. Features

Representativeness of the Population. Most insurance-based claims data are **not population-based**. Even Medicare data does not include beneficiaries who have elected Medicare Managed Care, since these organizations tend not to submit utilization data, and data that are submitted have not, to date, been integrated with the "fee for service" Medicare claims.

- Since the proportion of Medicare beneficiaries served by MCOs varies by region of the country, it is important to keep this in mind as we examine the meaning of large geographic variation in health care utilization among Medicare beneficiaries (Wennberg, Fisher et al., 2004a).
- Medicaid only covers those who are eligible and since eligibility criteria vary across states, making inter-state comparisons runs some risk of bias.
- Medicaid, like private insurance coverage, may not be continuous for any given individual as they move in and out of eligibility for Medicaid and gain, lose, or switch private health insurance coverage. Since there are nearly 50 million uninsured individuals in the country, including some 14% of those aged 55 to 64, administrative data covering their use of health services is obviously not population-based.

e-Source
Behavioral & Social Sciences Research

Exercise 2: Uses of Administrative Data

In this exercise, for each question below, please review the accompanying statement and select whether you think it is **appropriate** or **inappropriate** use of administrative data as presented.

Question: Level of Data

A research study is interested in understanding how many people have multiple hospitalizations by region. To accomplish this, the research team analyzes administrative data which provides use per thousand population.

[+ Appropriate] [− Inappropriate]

Question: Level of Data

A study is investigating cardiac events per individual to determine rates per hospital.

[+ Appropriate] [− Inappropriate]

Question: Data Linkage

A researcher is investigating rate of hospitalization based on zip code data. The researcher first determines the heterogeneity of individuals within the zip codes selected.

[+ Appropriate] [− Inappropriate]

Question: Level of Clinical Detail

A researcher would like to understand suboptimal use of Emergency Rooms for non-Emergency care. To do this, it is sufficient to consider administrative data which provides ICD codes and procedure codes.

[+ Appropriate] [− Inappropriate]

Question: Representativeness of Population

A researcher wanted to examine the impact of state guidelines on utilization, but decided not to use data provided by a survey such as the LSOA, even though the survey used a national sample.

[+ Appropriate] [− Inappropriate]

5. Accessing Data

Administrative medical data is necessarily "identifiable," particularly if it contains identifying information about the individual beneficiary or insurance subscriber. As such, privacy considerations are paramount and extensive protections are built into any request by researchers to use "identifiable" data.

Companies and their contractual agents engaged in electronic data interchange (EDI) must protect the confidentiality of their customers; any use of identifiable data must be specifically consented to by that individual unless it is approved for research purposes by a specially constituted Privacy Board. Following is a summary of the means by which researchers can safely access administrative/clinical data.

While there have always been protections in the use of identifiable data, under the privacy rules promulgated pursuant to HIPAA (the Health Insurance Portability and Accountability Act, passed in 1996), these were more systematized.

Data Use Agreements. Whether working with Medicare, Medicaid, or individual medical care insurance companies, researchers requesting to work with individually-identifying information must complete and conform with standardized data use agreements. Such agreements have to stipulate the precise data elements required and how these are needed to achieve the goal of the research, and describe, in detail, data security arrangements that restrict whom and how identifiers are accessed and used.

> "The User agrees to establish appropriate administrative, technical, and physical safeguards to protect the confidentiality of the data and to prevent unauthorized use or access to it. The safeguards shall provide a level and scope of security that is not less than the level and scope of security established by the Office of Management and Budget (OMB) in OMB Circular No. A-130, Appendix III--Security of Federal Automated Information Systems, which sets forth guidelines for security plans for automated information systems in Federal agencies."

Working within the framework of a data use agreement, researchers can link medical care utilization events for a given patient to demographic and pre-existing clinical characteristics longitudinally as well as hierarchically.

> Reporting the results of analyses of event- and person-level files based upon identifiable data are restricted by a requirement that no fewer than 10 individuals be included in any table cell published, in order to minimize the possibility of individuals' identities being uncovered.

The Centers for Medicare and Medicaid Services (CMS) has an elaborate procedure for processing data use requests, relying upon a subcontractor, the Research Data Assistance Center (ResDAC) at the University of Minnesota, to review, modify, and clarify data use requests before they are submitted to CMS.

5. Accessing Data

De-identified individual level data. Researchers can apply to use data files that will generally mask the patient's location or the provider's identity. This aims to minimize the potential of an individual study subject's personal information being used to identify that individual, leaving the user with only regional or state indicators of provider location. In most cases the precise dates of service utilization are also excluded, which offers the researcher only an indicator of the year in which the utilization event occurred.

A perfect example of a de-identified data base that has great utility for social science and public health research is the Hospital Cost and Utilization Project, maintained by the Agency for Health Research and Quality (AHRQ). Beginning with various state based samples, it has grown to a nationally representative sample of non-federal short-stay hospitals all reporting data which was first compiled in the 1970's (McCarthy, 1981). Authors of a recent report prepared under contract from AHRQ note that these data can be used for numerous applications ranging from injury and disease surveillance to comparative reporting on hospital quality performance on measures such as post-surgical infection or infection rates. An extensive bibliography of research studies based upon these data exists on line at the AHRQ web site.

De-identified data, which are available about each discharge from either the National Inpatient Sample or from all hospitals in various states, still contain:

- Demographic information;
- Up to 10 diagnoses;
- Insurance coverage; and
- Admission and discharge location.

While multiple hospitalizations of the same individual cannot be linked, given the proper data use agreement it is possible to:

- Aggregate all discharges of a given type or with a given condition or outcome to the level of the hospital and, in so doing, make comparisons across hospitals;
- Report on racial and ethnic disparities in hospital treatment given to minority and white hospitalized patients.

These data have also been used to contribute to the new National HealthCare Quality and Disparities Report prepared on behalf of AHRQ to examine inter-state differences in the disparities in health care receipt among white and minority patients.

5. Accessing Data

Aggregated Provider or Area Level Data. The Health Resources Service Administration (HRSA) has for many years maintained and updated data elements related to health care providers and delivery resources in the Area Resource File (ARF). These data are used in much the same way as are census data, compiled to the level of the county, including information on numbers of hospitals, long term care facilities, physicians, nurses, and other health care professionals.

- **Census-based population data** can be used to standardize the level of health care resources to create, for example, the number of hospital beds per 1000 population.
- **More detailed clinical data** (such as from the HCUP data base or measures constructed from Medicare billing data events) can be aggregated to provide numerators for the construction of population-based measures of health care utilization such as discharges of persons 65+ per 1000 population in one part of the country vs. another or Medicare costs in the last year of life per Medicare beneficiary (Wennberg, Freeman et al., 1989; Fisher, Wennberg et al., 1994).
- **Data from hospital discharge records** have been aggregated to construct a wide array of inpatient quality indicators on domains ranging from mortality for medical and surgical admissions, to area-wide measures of utilization relevant to characterizing the rate of "ambulatory care sensitive" hospitalizations such as asthma attacks for children.

5. Accessing Data

Semi-customized tables from on-line resources. Policy analysis groups, state and local governments, as well as advocacy groups are increasingly seeking to access "meta-versions" of research data that heretofore have been the exclusive province of those with the technical and statistical skills to analyze large surveys or all types of Medicare and/or Medicaid claims data. All these sources of "meta-data" allow the user to customize a data request that results in a multi-way table indicating the proportion of the population with selected characteristics (e.g., age, gender, employment, insurance status, etc.) that were hospitalized, visited a physician or spent over a certain amount for their health care services in the last several months. Examples include:

1. Data created for research purposes:
   - The AHRQ-sponsored Medical Expenditure Panel Survey;
   - The National Institute on Aging-funded Health and Retirement Survey;
   - The Robert Wood Johnson Foundation-funded Center for Health Systems Change consumer and physician surveys.
2. Query systems to help researchers decide on sample sizes and other aspects of the appropriateness of their research question before actually having to invest in the data acquisition:
   - The National Cancer Institute-maintained SEER- Medicare match data (Surveillance Epidemiology and End Result cancer registry);
   - The Kaiser Family Fund, which has created web sites that make available all manner of information about Medicaid policies, service use, demographics, and poverty and health statistics on a state-by-state basis;
   - The Annie E. Casey Foundation's "KIDSCOUNT" interactive data center, which, like the Kaiser Family Fund, allows the user to either focus in great detail on a given state or to compare states on a wide variety of parameters relevant to kids in the US;
   - The Dartmouth Atlas web site, which allows users to compare health regions on a variety of different parameters, specifically emphasizing regional heterogeneity in the intensity of health services use among similar types of populations.

Recent expansions of available aggregated data have permitted comparisons of hospitals on many of the same parameters that they've used to compare regions, exposing for the first time substantial intra-regional variation in the intensity of care processes that are provided in one place versus another.

Exercise 3: Importance of Data Security

Below are several statements with the following choices. Please select which choice is appropriate for each statement:

| Data Use Agreements (individually identifying data) | De-Identified Individual Level Data | Semi-customized tables from on-line resources | Aggregated Provider or Area Level Data |

Determining sample size during development of a research study.

Examining changes in hospital beds per capita in US counties from 1960 to 2000.

Estimating the prevalence of home care use in the population of community dwelling elderly.

Examining the effect of racial minority status and income on the likelihood of receiving appropriate medical treatment for cardio-vascular disease and whether observed disparity is related to where care is provided or whether the patient is minority.

6. Advantages and Disadvantages

Administrative health data are clearly more like demographic data available from the census than the rich, nuanced information that comes from direct observation. On the one hand, administrative health data contain clinical detail, chronological sequencing, powerful linkages between clinical and geographic detail, and health care utilization data that are situated in place, time, and person.

> **On the other hand, virtually completely missing is information about the social roles and functions that so often influence patients' decisions about health care utilization, regardless of clinical situation.**

Advantages. The primary advantage of administrative data systems is that in many cases they are population-based or at a minimum based upon a clearly defined population. With the correct linkages, population surveys and even population-based registries can greatly enhance the available information about clinical or social characteristics of the population. Here are some examples.

- *The National Long Term Care Survey* was conducted periodically over two decades using the Medicare/Medicaid population base as a sampling frame (Manton, Vertrees et al., 1990; Manton, Woodbury et al., 1993).
- *The Longitudinal Survey on Aging*, conducted during the 1980's, matched individual respondents' Medicare beneficiary numbers to their surveys (Mor, Wilcox et al., 1994; Wolinsky, Krygiel et al., 2002).
- *The SEER/Medicare match files*, created by the National Cancer Institute in collaboration with CMS, has repeatedly shown the value of integrating two population based data sources designed for totally different purposes (Nattinger, Gottlieb et al., 1992; Polsky, Armstrong et al., 2006).
- *The Medicare Beneficiary Survey (MCBS).* Integrating primary survey data collection with Medicare beneficiary administrative records and claims, the MCBS identified beneficiaries who'd switched to Managed Care companies that weren't required to submit administrative medical claims data, which has allowed numerous comparisons

of fee for service and managed care (Chulis, Eppig et al., 1993; Eppig and Chulis, 1997).

In the "run up" to the introduction of Medicare's Part D prescription drug act, the MCBS was one of the only sources of population-based data which could be used to estimate how beneficiaries would behave with universal drug coverage (Gilman and Kautter, 2008).

With the proper arrangements and assurances, these data can also often be encoded as belonging in markets which can be easily located in states or other political boundaries which influence the manner in which health care is delivered (Roos, Mustard et al., 1993). This ability to aggregate individuals' utilization experience makes it possible to estimate the effect of being served by a provider, or residing in an area, characterized by high rates of aggressive medical care, for example (Wennberg, Fisher et al., 2004b). To the extent that individuals' health care experience varies as a function of geographic region, as Wennberg and his colleagues have repeatedly shown, having "only" a nationally representative sample of individuals, even if one is able to locate them geographically, displaces the ability to construct contextual variables which are only possible if individual level data about the population of events, for example hospitalizations, are aggregated to the level of the area or the provider (Wennberg, Freeman et al., 1989; Wennberg, Fisher et al., 2004b).

6. Advantages and Disadvantages

Disadvantages. The principle disadvantage of administrative data is that one is limited to the data elements that were introduced, almost always for a totally different purpose. This has become particularly pertinent as hospitals and other providers are being compared based upon the "outcomes" of care (Bradley, Herrin et al., 2006; Werner and Bradlow, 2006). For example, for years researchers have been frustrated with hospitalization claims from Medicare or hospital discharge abstracts; while they allow for up to 10 diagnoses in the record, this listing often only addresses those diagnoses that were pertinent to the immediate management of the patient and not necessarily the pre-existing co-morbidities that might affect discharge outcomes like mortality or re-hospitalization (Jencks, Daley et al., 1988; Iezzoni, 2002). Researchers would like to have more clinical and social information to better "adjust" for unmeasured biases that undermine the validity of quality performance comparisons.

A related problem is the absence of contextual information related to the "culture" of a particular region of the country that might provide some "explanation" of the large observed differences in service use patterns. Wennberg and his colleagues have frequently reported the high level of regional variation in any number of different measures of standardized health services use (Wennberg, Fisher et al., 2004a). However, assembling the contextual information about this culture, whether at the level of the facility or a region, requires supplemental data that are not readily available from secondary data.

The second disadvantage of secondary data systems is that the quality and consistency of data reported may be influenced by the purposes for which it is collected. While difficult to determine, there is some evidence that with the introduction of hospital-based prospective case-mix adjusted reimbursement, some of the increase in the average case mix acuity of Medicare beneficiaries admitted to hospital was due to "up-coding" (Hsia, Ahern et al., 1992). Similar complaints have been levied against the introduction of case-mix reimbursement in nursing homes (Feng, Grabowski et al., 2006). Likewise, since the advent of public reporting of quality measures, providers also have an incentive to "under-code" complexities and complications that might result in their being identified as having a quality problem. Great care has to be taken in using the variables in these data systems that are subject to manipulation, whether to maximize reimbursement or minimize quality problem reporting.

Exercise 4: Advantages and Disadvantages

Below are several examples of how administrative data may be used. Please identify whether these uses are advantages or disadvantages.

Clearly defined population	Advantage	Disadvantage

Defines social characteristics of group	Advantage	Disadvantage

Serves as function of geographic region	Advantage	Disadvantage

Outcomes predefined by data	Advantage	Disadvantage

Key cultural and/or personal information and information preferences, social importance omitted	Advantage	Disadvantage

Data collected for a specific purpose	Advantage	Disadvantage

Do not have response rate variation	Advantage	Disadvantage

7. Illustrations

As noted in the preceding sections, the use of administrative data in social science-based research about health care systems has become increasingly prevalent in the field today.

> **Most major health care policy studies rely upon one source of administrative data or another.**

It is rare for a month to go by without one of the major medical journals including the results of an observational study based in part, or completely, on administrative data drawn from health care insurers, vital statistics, or state-wide or national hospital discharge data. A brief look at the scientific agenda for AcademyHealth in any recent year will reveal that the preponderance of presentations relies upon one or more types of administrative health care data.

In this section two research programs are described in terms of:

- The **intellectual contributions** that they've made to their respective fields;
- The **methodological work** they've advanced to test the quality of the data; and
- The **meaningfulness of the measures** they've constructed based upon administrative health data.

7. Illustrations

The Dartmouth Atlas Project (DAP).

Summary: The Dartmouth Health Care Atlas emerged after nearly a decade of work by Jack Wennberg and his colleagues documenting regional variation in health care utilization, particularly of discretionary surgical procedures (McPherson, Wennberg et al., 1982; Wennberg, 1985; Wennberg, Roos et al., 1987). DAP research has essentially defined the field of small area variation, moving from comparing the pattern and level of health care use across regions to comparing the treatment provided to Medicare patients by individual hospitals and, most recently, to examining patients' "loyalty" to primary care physicians (Welch, Miller et al., 1993; Goodman, Mick et al., 2003; Wennberg, Fisher et al., 2004a; Bynum, Bernal-Delgado et al., 2007). While many contributions have been based only on analyses of the Medicare claims data, some work has linked a national sample of physicians surveyed about approaches to diagnosing common clinical case presentation with aggregated estimates of regional variation in Medicare spending. In several such studies, the researchers found that physicians' reporting a higher willingness to use expensive diagnostic tests were significantly more likely to be located in high cost regions, thus contributing to the notion that "medical cultures" vary and contribute to the large observed differences in treatment patterns (Sirovich, Gottlieb et al., 2005; Sirovich, Gottlieb et al., 2006).

Intellectual Contributions: The intellectual focus of much of the research conducted by the group over the last several decades has been based upon the premise that providers have a vested interest in inducing demand for the health care services that they provide (Fisher, Wennberg et al., 2000). They pioneered the approach of contrasting regional variation in the per capita utilization of discretionary and non-discretionary services on the assumption that the larger the discretion, the more likely providers and the patients they serve in the high discretionary use areas are responding to something other than clinical "need" (Weinstein, Lurie et al., 2006). A recent report from the DAP project displays variation in the rates of surgery among Medicare beneficiaries across 300+ hospital referral regions for 4 types of surgery at varying levels of discretion.

7. Illustrations

Figure 2: Variation in Discharge Ratio

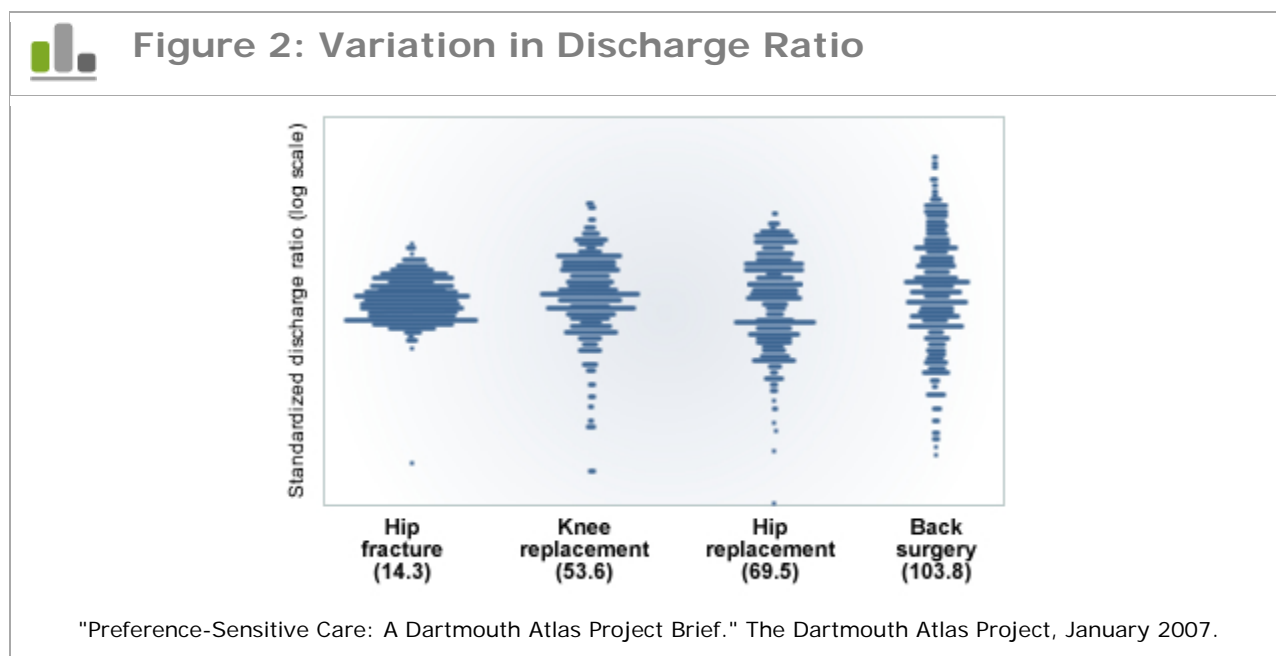"Preference-Sensitive Care: A Dartmouth Atlas Project Brief." The Dartmouth Atlas Project, January 2007.

Figure 2 presents these data, clearly revealing the increasing degree of variation as a function of the level of discretion in the surgical procedures in question.

For example, hip fracture repair rates per 1000 Medicare beneficiaries is relatively similar across the 306 HRRs. On the other hand, back surgery is extremely variable, with some areas manifesting rates that are almost 8 times as great as areas with the lowest rates (Weinstein, Lurie et al., 2006). Earlier studies clearly documented the relationship between the number of medical specialists (or orthopedic surgeons) and the average level of Medicare health care spending per capita in an area, adding further weight to the argument that supply stimulates demand for care (Welch, Miller et al., 1993; Fisher and Welch, 1999).

The creation of variables that characterize the style of medical treatment in one part of the country or another has worked to define an entire field of research and to continuously push the boundaries of epidemiological, health services, and social science research.

In particular, the DAP website is a great source of data for researchers interested in understanding the intra and inter-regional differences in various health care consumption measures that might shed light on a whole new phenomenon (Wennberg, Fisher et al., 2005; Adamson, 2008). The possibilities, particularly once the data are linked to other geographically rooted data, are limitless. Here are some examples of research data structures that DAP created:

- *Hospital Referral Regions (HRRs)*. One of the first major methodological advances made by the group, these determined the boundaries of a region based upon the proportion of Medicare beneficiaries who received cardio-vascular surgeries, then drew them to maximize that proportion across all regions of the country, all of which was done with Medicare hospital claims data, aggregated to region based upon the zip code of the beneficiary (McPherson, Wennberg et al., 1982).
- *Hospital Service Areas (HSAs)*. Using the same approach to conducting a "patient origin" study, these are considerably smaller than tertiary referral regions (HRRs) (Fisher, Welch et al., 1992). HSAs are smaller groups of hospitals which have the advantage of being a more coherent market for general hospital care but result in considerably more overlap among patients and physicians.
- *Primary care service areas*. DAP researchers constructed even smaller units of measure by undertaking another patient origin study that assigned zip code clusters to those beneficiaries who received a preponderance of primary care (based upon claims data containing information on type and location of service) from physicians located in those areas (Goodman, Mick et al., 2003). In all these market definitions, the patient origin studies revealed that between 65% and 85% of the relevant type of service was provided by physicians or hospitals located in the area.

7. Illustrations

Methodological Challenges: Another methodological challenge that the DAP investigators have embraced relates to determining the appropriate unit of analysis for their studies.

- Data from other sources (physician surveys, SEER cancer diagnosis and stage, etc.) have been matched in related variation in responses of physicians and treatment patterns to HRR level aggregates (Polsky, Armstrong et al., 2006; Sirovich, Gottlieb et al., 2006).

- Claims data have been aggregated to the level of the HRR and then analyses were conducted based upon the 306 HRRs (Wennberg, Fisher et al., 2004a).

- Claims have been aggregated to the level of the hospital then linked to information from the American Hospital Association files or other measures constructed by CMS, used to rank the hospitals in terms of the quality of care provided (Fisher and Welch, 1999; Wennberg, Fisher et al., 2004b).

- Data have been used to examine racial disparities, e.g., the likelihood of receiving recommended treatments or receiving more discretionary treatments (Skinner, Weinstein et al., 2003). Of particular interest are their more recent studies seeking to determine how much of the observed racial disparities in treatments is attributable to an individual's race or to the provider or part of the country from which it comes (Baicker, Chandra et al., 2004).

7. Illustrations

Brown Long Term Care Facts on Care in the U.S. (LTCFocUS) Data Base

Summary: This database was stimulated by the mandated implementation and subsequent computerization of a uniform resident assessment instrument (MDS) for all nursing home residents (regardless of funding source) in all Medicare/Medicaid nursing facilities in the country. These data made it possible, for the first time, to document the changing population of nursing home users and how it varies across the country. Since the MDS includes identifiers, it can be linked to Medicare and Medicaid claims and because the facility providing care is indicated, individuals' assessment data can be linked to characteristics of the facility and all the other residents it serves. The data base includes:

- All MDS assessments (required upon admission and at least quarterly thereafter);
- Medicare enrollment records and Medicare inpatient, outpatient, home health and hospice claims;
- Nursing home level inspection, staffing and ownership data collected annually;
- Area Resource File (ARF) data characterizing the demographic, economic, and health care delivery features of the county in which the facility is located; and
- Surveys of state Medicaid agencies regarding various policies, including payment rate, that affect nursing homes in that state (Teno, Feng et al., 2008; Grabowski, Feng et al., 2008; Mor, 2004).

To make the assessment data more useful in and of themselves as a source of research, investigators at Brown and elsewhere:

- Matched MDS to Medicare data at the patient level;
- Linked each patient to the appropriate nursing home provider; and
- Geo-coded provider's address to a level even more detailed than zip code, county, and state.

While the MDS was initially designed as a clinical assessment tool to structure the care planning process, it was rapidly adapted as the basis for determining case-mix acuity and

reimbursement, and as the basis for measuring, monitoring, and ultimately publicly reporting the quality of care offered by nursing home providers (Morris, Hawes et al., 1990; Mor, 2004).

Intellectual Contributions

Brown's LTCFoCUS data base is a much newer development than the Dartmouth Atlas Project, but has many of the same features; it is population-based, geographically-oriented, and able to address clinical as well as policy questions pertinent to long term care in the US. From the early days of the data base several novel intellectual contributions, ranging from clinical issues to estimating the impact of states' policies to understanding the implications of racial disparities in care, have already emerged. These are briefly summarized in the paragraphs below.

The Epidemiology of Nursing Home Drug Use

The Brown LTCFoCUS database initially allowed for the detailed study of the epidemiology of drug use among nursing home residents because detailed drug data were now available, since some states had mandated their collection using National Drug Codes to enumerate all prescription and most non-prescription drugs. Once these data were merged with Medicare claims, it was possible to estimate the effect of different drug classes on residents' experience of adverse events like hospitalization due to stroke, congestive heart failure, or hip fracture. This required matching individual level MDS assessment data with Medicare claims and enrollment records and sequencing them by dates of occurrence to establish residents' exposure to drugs and subsequent risk of an outcome like hospitalization or death (Gambassi, Landi et al., 1998; Pedone, Lapane et al., 2004; Liperoti, Gambassi et al., 2005).

Research in this area also focused on common geriatric conditions and quality problems in the nursing home context such as pain, poly-pharmacy and psychotropic drug use. Underuse of analgesics has been identified in numerous studies of hospital and ambulatory care, but an important contribution was the documentation of inadequate control of nursing home residents' pain with medications (Bernabei, Gambassi et al., 1998; Won, Lapane et al., 1999). Investigators examined the effects of ACE inhibitors on hospitalization for congestive heart failure in a population whose use of these drugs had heretofore not been systematically investigated (Gambassi, Forman et al., 2000; Gambassi, Lapane et al., 2000). Given the salience of the use of psychotropic agents to control residents' disruptive behavior, considerable effort has gone into characterizing and documenting the prevalence of adverse events

associated with the use of various classes of these medications (Sgadari, Lapane et al., 2000; Liperoti, Mor et al., 2003; Liperoti, Gambassi et al., 2005 Liperoti, Onder et al., 2007).

7. Illustrations

The Effect of Policies and Markets on Providers and Residents' Outcomes

Since 1999 investigators at Brown University have linked data acquired via surveys of Medicaid officials and reviews of state web sites to facility-level information that summarizes state and federal inspections of nursing homes and resident-level MDS and Medicare claims records. Their aim is to test the effect of changes in state Medicaid policies on the strategic management choices nursing facility administrators make and how those decisions ultimately affect the lives of the residents served (Mor, Zinn et al., 2007; Grabowski, Feng et al., 2008; ZInn, Feng et al., 2008). This framework motivated a National Institute on Aging-funded Program Project that is now extending these data for more than a decade and linking this information with a specially designed survey of a large random sample of nursing homes. Figure 3 below shows a heuristic that motivates this research, which is made possible precisely by linking individual and health service utilization event-level data for a defined population to information about the providers of care, and the state policy and market contexts in which they operate.

These data permit rigorous testing of states' policies' impact by taking advantage of the "natural experiments" operating in states around the country. Since federal Medicare policies governing payment of nursing homes and home health providers have also changed over the observation, it is also possible to compare changes in federal policies with related state-based Medicaid policy changes, thereby better informing our understanding of the impact of federal policy changes.

Figure 3: Policies and Market Effects
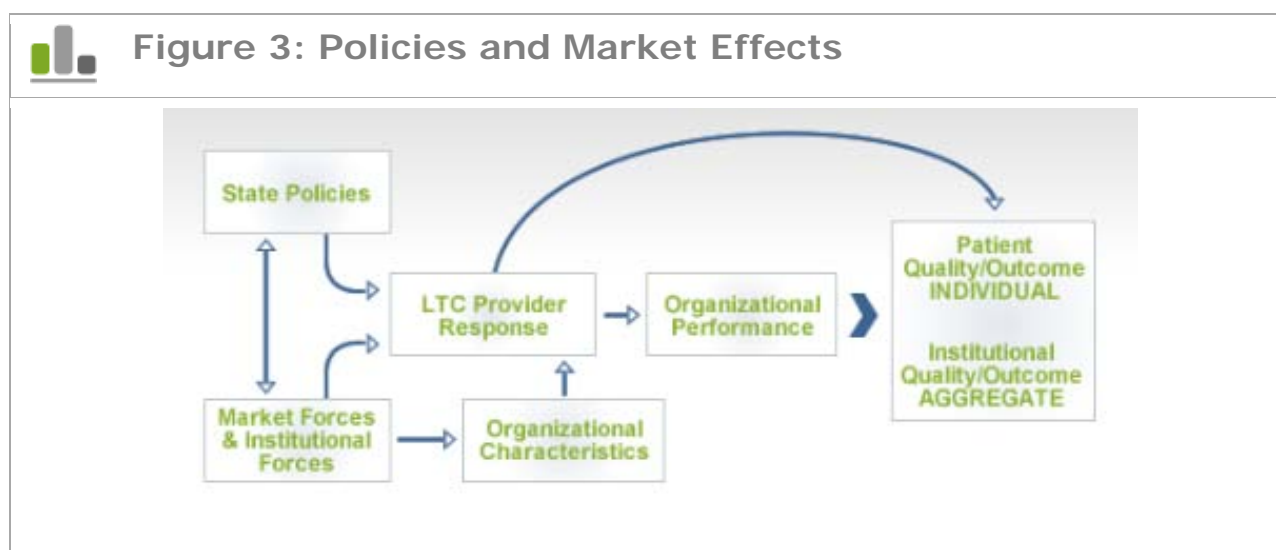
7. Illustrations

Research conducted using this paradigm includes an examination of the impact of states' Medicaid payment levels on different aspects of staffing levels on the assumption that states that pay nursing homes more will end up with higher staffing levels. Both the level and skill mix of nursing home staff appear to be responsive to increases in the payment rate that states give nursing homes (Intrator, Feng et al., 2005; Feng, Grabowski et al., 2008). Additionally, states' adoption of reimbursement based on the acuity mix of the residents served was found to be strongly related to increasing acuity of both patients admitted into nursing homes and those remaining there (Feng, Grabowski et al., 2006). However, there was no evidence that introducing case-mix reimbursement resulted in the insertion of more feeding tubes, something which many had feared would happen in light of the extra acuity points associated with patients requiring tube feeding (Teno, Feng et al., 2008). Other state policies, ranging from paying nursing homes to "hold" the bed of their hospitalized patients to the introduction of provider bed taxes, have been examined using this paradigm with both expected and unexpected results (Intrator, Grabowski et al., 2007; Grabowski, Feng et al., 2008; Intrator, Schleinitz et al., 2009).

Policies of interest have included:
- The introduction of Medicaid nursing home state based reimbursement levels;
- The average level of Medicaid payment provided facilities;
- The presence of policies reimbursing nursing homes for time their Medicaid patients are hospitalized in order to compensate them for holding the bed;
- Focused incremental payments intended for direct care staff;
- Investments in community based long term care; and
- Other reimbursement rules that alter nursing facilities' incentives.

Research that has built upon this framework includes a study that tested the effect of the introduction of state-based, Medicaid case mix reimbursement introduced to compensate nursing facilities for admitting more severely impaired residents. Since, during the period in

question, some states introduced this payment model while others did not, it was possible to use a "difference in difference" approach to determine whether case mix acuity of both residents and new admissions to nursing homes increased as a result of states' adoption of case mix reimbursement (Feng, Grabowski et al., 2006). It was also observed that with the introduction of case mix reimbursement additional staff resources were devoted to administrative functions, something that is consistent with the experience of hospitals following the introduction of hospital case mix reimbursement (Hsia, Krushat et al., 1988; Feng, Grabowski et al., 2008).

Racial Disparities in the Quality of Nursing Home Care

The availability of population based data on nursing home use in all regions of the country has made it possible to examine patterns of nursing home use as a function of race and how segregated nursing homes are. Aggregating resident level data on race to the level of the facility, researchers found that African American nursing home residents were much more likely to be in the poorest quality facilities in any market (Mor, Zinn et al., 2004). By looking at who was entering nursing homes, researchers also found that black patients discharged from hospital to a nursing home were 30% more likely to enter a home in the bottom quality quartile in the market than were whites, controlling for clinical needs and educational attainment (Angelelli, Grabowski et al., 2006). Most recently, using MDS data matched to 2000 census level data, researchers found that blacks had higher rates of nursing home use in many areas and that they were much more likely to be served by the poorest quality homes in an area relative to whites (Smith, Feng et al., 2007). And just as important is the fact that, in examining the relative influence of individuals' race as opposed to the proportion of residents in the home that are minority, members of the Brown research team found that over and above patients' own race, both white and black residents of homes with high proportions of African Americans were significantly more likely to be hospitalized for conditions that were potentially avoidable than was the case for residents in homes with fewer blacks (Gruneir, Miller et al., 2008). The power of these nested data on nursing home residents, linked with census and outcome data like hospitalization creates numerous opportunities to more definitively address the issue of disparities in care – is it who you are or where you get your care that matters?

7. Illustrations

Methodological Challenges

Construction of the longitudinal, hierarchical data base took advantage of the ability to create aggregated data from individual-level assessment information, characterizing either the nursing home itself or the market or state in which it was located. At the provider level this greatly enhanced the richness of the publicly available nursing home level data collected annually during the facility inspection process (the Online Survey and Certification Automated Record – OSCAR). Aggregating data on all residents to the level of a state facilitates inter-state comparisons of the nursing home population served. In an effort to explore the extent to which states' policies concern the availability of community based alternatives to long term care facilities, investigators compared the proportion of residents classified as "low care" and found that states with lower investments in home and community based services and more nursing home beds per thousand elderly had higher proportions of low care residents (Mor, Zinn et al., 2007).

In spite of the obvious advantages of having geographically specific, population based data on all nursing home users, the facilities that serve them, and the county resources and state policies, there are several methodological challenges that complicate the interpretation of analyses based upon these observational data. Included are the endogeneity of states' policies in the nursing home arena: data quality varies by state, potentially undermining state effects; and the data don't cover all long term care users, only those who use nursing homes. Each of these issues and the approaches for circumventing them are discussed below.

Many of the examples of analyses conducted testing the impact of selected state policies assume that policies are implemented for fiscal reasons that are largely independent of the outcomes experienced by nursing home residents. Obviously, another causal line of reasoning is possible; states may introduce policies in response to scandals about the quality of care provided to residents of some homes. This argument would be consistent with the background to many policies governing states' approach to reimbursing and/or sanctioning nursing homes which often emerge in response to scandals. Nonetheless, since no state has been immune to such scandals over the past decade, it is unclear whether scandals are state-specific in terms of their likelihood of stimulating a policy.

Facility inspection data are collected by inspectors following federal guidelines which are, however, demonstrably variable in implementation from state to state (Angelelli, Mor et al., 2003). Because of this, most analyses using inspection data have to adjust for such state differences. Staffing data is another factor that varies by state, not necessarily because of the way the data are assembled, but rather because of differences in the tightness of the labor markets for both skilled nurses as well as unskilled nursing assistants. This means that most analyses have to adjust for the local unemployment rate (Weech-Maldonado, Meret-Hanke et al., 2004).

Finally, while the LTCFoCUS data base has great breadth and depth, it only includes the "numerator" of nursing home users. The population of individuals who "might" or "could" use a nursing home is much larger than the population of individuals entering nursing home care. To fully understand inter-state differences in nursing home admissions or patterns of care, it is important to know how different are the populations of persons requiring long term care from state to state, including that segment of the population that is not using nursing home care. While there are data on Medicare beneficiaries who use home health care, this covers only a minority of home health care provided in the country and certainly doesn't include all the people using Assisted Living and other non-nursing home residential care arrangements. While it is possible to add information about the supply of such alternative long term care services in each market, not having individual level data on this population necessarily limits the scope of analyses that can be legitimately undertaken.

7. Illustrations

Brown investigators are in the process of building a web site which will permit policy analysts and researchers to download data, thereby extending the availability of the information to a broader community of researchers. The goal of this database, like that of the publicly available Dartmouth Atlas, is to permit analysts and researchers to directly query the data at a variety of different levels to address a policy-relevant issue or to make comparisons of groups of providers. Researchers will be able to obtain longitudinal, historical data at three different levels:

State level data: policies, aggregated provider data, selected aggregated characteristics of the resident population and of the universe of nursing home admissions.

Market level data: selected indicators from the Area Resource File, the Dartmouth Atlas and aggregated nursing home provider and resident level data.

Provider level data: staffing, service availability, aggregated resident and admission population characteristics, and selected "outcomes" such as hospitalization, etc. No data on any provider level parameter that has fewer than 11 observations in the numerator will be made available to limit identifiably.

🖥 Example 1: Use of Administrative Data

The Healthcare Cost and Utilization Project (HCUP) is an on-line query system that provides access to health statistics and information on hospital inpatient and emergency department utilization. Click on the following link, to access (HCUP) and create a set of data tables. As you do so, consider such questions as:

Do these data give you the rate of hospitalization in the population?

Would this database help you understand what is causing those admissions via the ED?

Additional information

To further explore the use of administrative data you can also view Dartmouth Atlas (level of data is hospital or region) to apply concepts presented in this chapter and determine how best to use administrative data.

8. Summary

There is growing use of administrative data in social science and epidemiological research, particularly on topics related to health and aging. Critics of the uses of these data rightly point out that excessive reliance on administrative data lead to studies asking questions that are limited to where the data are. They bemoan the fact that investigators are only looking underneath the street lights when the more informative studies must, of necessity, force us to examine issues that have been, until recently, in the dark. While this is clearly true and by no means should primary data collection efforts be abandoned, it is the case that studies using administrative data have unique advantages and can be powerful complements to primary data collection efforts.

> **Qualitative studies can be greatly enhanced by placing them into context with the use of population-based administrative data, where administrative data are used to identify sources of variation that beg further investigation using surveys and case studies. Insights from the latter generate hypotheses that can then be tested using administrative data.**

Investigators' consideration of the best approach to address a research question inevitably requires balancing design issues such as the representativeness of the sample, the ability to recruit, and the cost of data collection given the size of the sample needed to test the study hypotheses. Because both response rates to household surveys and the representativeness of telephone surveys are declining, the trade-offs between the advantages and disadvantages of surveys versus administrative data may be changing. Notwithstanding the complexities of securing permission from survey respondents to link their responses to their administrative records, increasingly the advantages of administrative data are being added to the advantages of surveys by linking these two powerful sources of data.   The power of these combined data sets to address a broad array of issues is evident in the large number of publications that have emanated from such studies.

What is unique about the data systems being assembled by investigators at Dartmouth and Brown is their longitudinal and hierarchical character. Since providers are located in space and

Medicare beneficiaries' zip codes are known, studies linking surveys to Medicare claims could also be used to examine individuals' choice of provider and whether the forces that appear to shape that choice vary from market to market or from state to state. This ability to marry the "demand" side of health care utilization (individuals' wealth, education, insurance and even preferences derived from surveys) with "supply" side information on provider quality, the medical care "culture" prevalent in the market, and relevant state policies will be the new frontier in social science research in health and aging.

9. Resources

Organizations

National Hospital Discharge Survey, National Center for Health Statistics

National Hospital Ambulatory Medical Care Survey (NHAMCS)

Inter-University Consortium for Political and Social Research (ICPSR)

Annual summaries of NHAMCS and NAMCS

Surveillance, Epidemiology, End Stage Results (SEER)

National Health Interview Survey Linked Mortality Public-use Data Files

"Provider of Service" files

"Current Procedural Technology" (CPT) codes

Created and maintained by the American Medical Association

HIPAA (the Health Insurance Portability and Accountability Act, passed in 1996)

Office of Management and Budget (OMB) in OMB Circular No. A-130, Appendix III

Security of Federal Automated Information Systems

Research Data Assistance Center (ResDAC), University of Minnesota

Hospital Cost and Utilization Project

Maintained by the Agency for Health Research and Quality (AHRQ)

Recent report prepared under contract from AHRQ

Bibliography of research studies at the AHRQ web site

National HealthCare Quality and Disparities Report prepared on behalf of AHRQ

Area Resource File (ARF), The Health Resources Service Administration (HRSA)

Inpatient Quality Indicators

10. References

Adamson, D. M. C., Stella, Hansen, Leigh G. (2008). *Health research data for the real world: The MarketScan databases.* T. H. Care. New York, Thompson Health Care.

Angelelli, J., D. C. Grabowski, et al. (2006). "Effect of educational level and minority status on nursing home choice after hospital discharge." *Am J Public Health* 96(7): 1249-53.

Angelelli, J., V. Mor, et al. (2003). "Oversight of nursing homes: pruning the tree or just spotting bad apples?" *Gerontologist* 43 Spec No 2: 67-75.

Baicker, K., A. Chandra, et al. (2004). "Who you are and where you live: How race and geography affect the treatment of medicare beneficiaries." *Health Aff (Millwood)* Suppl Web Exclusives: VAR33-44.

Bernabei, R., G. Gambassi, et al. (1998). "Management of pain in elderly patients with cancer. SAGE Study Group. Systematic Assessment of Geriatric Drug Use via Epidemiology." *Jama* 279(23): 1877-82.

Bradley, E. H., J. Herrin, et al. (2006). "Hospital quality for acute myocardial infarction: correlation among process measures and relationship with short-term mortality." *Jama* 296(1): 72-8.

Bynum, J. P., E. Bernal-Delgado, et al. (2007). "Assigning ambulatory patients and their physicians to hospitals: a method for obtaining population-based provider performance measurements." *Health Serv Res* 42(1 Pt 1): 45-62.

Chulis, G. S., F. J. Eppig, et al. (1993). "Health insurance and the elderly: Data from MCBS (Medicare Current Beneficiary Survey)." *Health Care Financ Rev* 14(3): 163-81.

Cooper, G. S., Z. Yuan, et al. (2001). "Determination of county-level prostate carcinoma incidence and detection rates with Medicare claims data." *Cancer* 92(1): 102-9.

Cooper, G. S., Z. Yuan, et al. (1999). "The utility of Medicare claims data for measuring cancer stage." *Med Care* 37(7): 706-11.

Densen, P., Shapiro, S, Einhorn, M (1959). "Concerning high and low utilizers of service in a medical care plan, and the persistence of utilization levels over a three-year period." *Milbank Q* 37: 217-250.

Dickson, M. and C. A. Plauschinat (2008). "Compliance with antihypertensive therapy in the elderly : A comparison of fixed-dose combination Amlodipine/Benazepril versus component-based free-combination therapy." *Am J Cardiovasc Drugs* 8(1): 45-50.

Eggers, P. W. (1984). "Trends in Medicare reimbursement for end-stage renal disease: 1974-1979." *Health Care Financ Rev* 6(1): 31-8.

Eppig, F. J. and G. S. Chulis (1997). "Matching MCBS (Medicare Current Beneficiary Survey) and Medicare data: the best of both worlds." *Health Care Financ Rev* 18(3): 211-29.

Etzioni, R., G. F. Riley, et al. (2002). "Measuring costs: Administrative claims data, clinical trials, and beyond." *Med Care* 40(6 Suppl): III63-72.

Feng, Z., D. C. Grabowski, et al. (2006). "The effect of state Medicaid case-mix payment on nursing home resident acuity." *Health Serv Res* 41(4 Pt 1): 1317-36.

Feng, Z., D. C. Grabowski, et al. (2008). "Medicaid payment rates, case-mix reimbursement, and nursing home staffing-1996-2004." *Med Care* 46(1): 33-40.

Fisher, E. S. and H. G. Welch (1999). "Avoiding the unintended consequences of growth in medical care: How might more be worse?" *Jama* 281(5): 446-53.

Fisher, E. S., H. G. Welch, et al. (1992). "Prioritizing Oregon's hospital resources. An example based on variations in discretionary medical utilization." *Jama* 267(14): 1925-31.

Fisher, E. S., J. E. Wennberg, et al. (1994). "Hospital readmission rates for cohorts of Medicare beneficiaries in Boston and New Haven." *N Engl J Med* 331(15): 989-95.

Fisher, E. S., J. E. Wennberg, et al. (2000). "Associations among hospital capacity, utilization, and mortality of US Medicare beneficiaries, controlling for sociodemographic factors." *Health Serv Res* 34(6): 1351-62.

Foster, R. W. (1985). "Cost-shifting under cost reimbursement and prospective payment." *J Health Econ* 4(3): 261-71.

Gambassi, G., D. E. Forman, et al. (2000). "Management of heart failure among very old persons living in long-term care: has the voice of trials spread? The SAGE Study Group." *Am Heart J* 139(1 Pt 1): 85-93.

Gambassi, G., F. Landi, et al. (1998). "Validity of diagnostic and drug data in standardized nursing home resident assessments: potential for geriatric pharmacoepidemiology. SAGE Study Group. Systematic Assessment of Geriatric drug use via Epidemiology." *Med Care* 36(2): 167-79.

Gambassi, G., K. L. Lapane, et al. (2000). "Effects of angiotensin-converting enzyme inhibitors and digoxin on health outcomes of very old patients with heart failure. SAGE Study Group. Systematic Assessment of Geriatric drug use via Epidemiology." *Arch Intern Med* 160(1): 53-60.

Gibbs, P., S. McLaughlin, et al. (2006). "Re: Completion of therapy by Medicare patients with stage III colon cancer." *J Natl Cancer Inst* 98(21): 1582.

Gibson, T. B., T. L. Mark, et al. (2006). "Impact of statin copayments on adherence and medical care utilization and expenditures." *Am J Manag Care* 12 Spec no.: SP11-9.

Gilman, B. H. and J. Kautter (2008). "Impact of multitiered copayments on the use and cost of prescription drugs among Medicare beneficiaries." *Health Serv Res* 43(2): 478-95.

Goodman, D. C., S. S. Mick, et al. (2003). "Primary care service areas: A new tool for the evaluation of primary care services." *Health Serv Res* 38(1 Pt 1): 287-309.

Grabowski, D. C., Z. Feng, et al. (2008). "Medicaid Nursing Home Payment and the Role of Provider Taxes." *Med Care Res Rev*. 65(4):514-27. Epub 2008 Mar 27.

Greenfield, C. A., P. M. Densen, et al. (1978). "Use of out-of-plan services by Medicare members of HIP." *Health Serv Res* 13(3): 243-60.

Greenlick, M. R. and B. J. Darsky (1968). "A comparison of general drug utilization in a metropolitan community with utilization under a drug prepayment plan." *Am J Public Health Nations Health* 58(11): 2121-36.

Gruneir, A., S. C. Miller, et al. (2008). "Relationship between State Medicaid Policies, Nursing Home Racial Composition, and the Risk of Hospitalization for Black and White Residents." *Health Serv Res* 43(3): 869-81.

Holzer, S. S., L. Engelhart, et al. (1997). "Asthma treatment costs using inhaled corticosteroids." *Am J Manag Care* 3(6): 891-7.

Hsia, D. C., C. A. Ahern, et al. (1992). "Medicare reimbursement accuracy under the prospective payment system, 1985 to 1988." *Jama* 268(7): 896-9.

Hsia, D. C., W. M. Krushat, et al. (1988). "Accuracy of diagnostic coding for Medicare patients under the prospective-payment system." *N Engl J Med* 318(6): 352-5.

Hurtado, A. V. and M. R. Greenlick (1971). "A disease classification system for analysis of medical care utilization, with a note on symptom classification." *Health Serv Res* 6(3): 235-50.

Iezzoni, L. I. (2002). "Using administrative data to study persons with disabilities." *Milbank Q* 80(2): 347-79.

Intrator, O., Z. Feng, et al. (2005). "The employment of nurse practitioners and physician assistants in U.S. nursing homes." *Gerontologist* 45(4): 486-95.

Intrator, O., D. C. Grabowski, et al. (2007). "Hospitalization of nursing home residents: The effects of states' Medicaid payment and bed-hold policies." *Health Serv Res* 42(4): 1651-71.

Intrator, O., M. Schleinitz, et al. (2009). "Maintaining Continuity of Care for Nursing Home

Residents: Effect of States' Medicaid Bed-Hold Policies and Reimbursement Rates." *Health Serv Res* 44(1):33-55. Epub 2008 Sep 8.

Jencks, S. F., J. Daley, et al. (1988). "Interpreting hospital mortality data. The role of clinical risk adjustment." *Jama* 260(24): 3611-6.

Klar, R. M. (1979). "ESRD program projections and experience." *Jama* 241(3): 239-40.

Liperoti, R., G. Gambassi, et al. (2005). "Conventional and atypical antipsychotics and the risk of hospitalization for ventricular arrhythmias or cardiac arrest." *Arch Intern Med* 165(6): 696-701.

Liperoti, R., V. Mor, et al. (2003). "The use of atypical antipsychotics in nursing homes." *J Clin Psychiatry* 64(9): 1106-12.

Liperoti, R., G. Onder, et al. (2007). "Conventional or atypical antipsychotics and the risk of femur fracture among elderly patients: results of a case-control study." *J Clin Psychiatry* 68(6): 929-34.

Manton, K. G., J. C. Vertrees, et al. (1990). "Changes in health service use and mortality among U.S. elderly in 1980-1986." *J Aging Health* 2(2): 131-56.

Manton, K. G., M. A. Woodbury, et al. (1993). "Use of Medicare services before and after introduction of the prospective payment system." *Health Serv Res* 28(3): 269-92.

Marks, S., Greenlick, MR, Hurtado, AV, Johnson, JD, Henderson, BA (1980). "Ambulatory surgery in an HMO: A study of costs, quality of care and satisfaction." *Med Care* 18(2): 127-146.

McCarthy, E. (1981). "Inpatient utilization of short-stay hospitals by diagnosis United States, 1978." *Vital Health Stat* 55: 1-79.

McPherson, K., J. E. Wennberg, et al. (1982). "Small-area variations in the use of common surgical procedures: an international comparison of New England, England, and Norway." *N Engl J Med* 307(21): 1310-4.

Mor, V. (2004). "A comprehensive clinical assessment tool to inform policy and practice: applications of the minimum data set." *Med Care* 42(4 Suppl): III50-9.

Mor, V., V. Wilcox, et al. (1994). "Functional transitions among the elderly: patterns, predictors, and related hospital use." *Am J Public Health* 84(8): 1274-80.

Mor, V., J. Zinn, et al. (2004). "Driven to tiers: socioeconomic and racial disparities in the quality of nursing home care." *Milbank Q* 82(2): 227-56.

Mor, V., J. Zinn, et al. (2007). "Prospects for transferring nursing home residents to the community." *Health Aff (Millwood)* 26(6): 1762-71.

Morris, J. N., C. Hawes, et al. (1990). "Designing the national resident assessment instrument for nursing homes." *Gerontologist* 30(3): 293-307.

Morris, J. N., S. Nonemaker, et al. (1997). "A commitment to change: Revision of HCFA's RAI." *J Am Geriatr Soc* 45(8): 1011-6.

Nattinger, A. B., M. S. Gottlieb, et al. (1992). "Geographic variation in the use of breast-conserving treatment for breast cancer." *N Engl J Med* 326(17): 1102-7.

Pedone, C., K. L. Lapane, et al. (2004). "Donepezil use in US nursing homes." *Aging Clin Exp Res* 16(1): 60-7.

Pine, P. L., M. Gornick, et al. (1981). "Analysis of services received under Medicare by specialty of physician." *Health Care Financ Rev* 3(1): 89-116.

Polsky, D., K. A. Armstrong, et al. (2006). "Variation in chemotherapy utilization in ovarian cancer: The relative contribution of geography." *Health Serv Res* 41(6): 2201-18.

Roos, L. L., C. A. Mustard, et al. (1993). "Registries and administrative data: Organization and accuracy." *Med Care* 31(3): 201-12.

Sangl, J., D. Saliba, et al. (2005). "Challenges in measuring nursing home and home health quality: Lessons from the First National Healthcare Quality Report." *Med Care* 43(3 Suppl): I24-32.

Sgadari, A., K. L. Lapane, et al. (2000). "Oxidative and nonoxidative benzodiazepines and the risk of femur fracture. The Systematic Assessment of Geriatric Drug Use Via Epidemiology Study Group." *J Clin Psychopharmacol* 20(2): 234-9.

Shapiro, S., W. Venet, et al. (1985). "Selection, follow-up, and analysis in the Health Insurance Plan Study: a randomized trial with breast cancer screening." *Natl Cancer Inst Monogr* 67: 65-74.

Sirovich, B. E., D. J. Gottlieb, et al. (2005). "Variation in the tendency of primary care physicians to intervene." *Arch Intern Med* 165(19): 2252-6.

Sirovich, B. E., D. J. Gottlieb, et al. (2006). "Regional variations in health care intensity and physician perceptions of quality of care." *Ann Intern Med* 144(9): 641-9.

Skinner, J., J. N. Weinstein, et al. (2003). "Racial, ethnic, and geographic disparities in rates of knee arthroplasty among Medicare patients." *N Engl J Med* 349(14): 1350-9.

Smith, D. B., Z. Feng, et al. (2007). "Separate and unequal: racial segregation and disparities in quality across U.S. nursing homes." *Health Aff (Millwood)* 26(5): 1448-58.

Sorrentino, E. A. (1989). "Hospital mission and cost differences." *Hosp Top* 67(3): 22-5.

Teno, J. M., Z. Feng, et al. (2008). "Do financial incentives of introducing case mix reimbursement increase feeding tube use in nursing home residents?" *J Am Geriatr Soc* 56(5): 887-90.

Weech-Maldonado, R., L. Meret-Hanke, et al. (2004). "Nurse staffing patterns and quality of care in nursing homes." *Health Care Manage Rev* 29(2): 107-16.

Weinstein, J. N., J. D. Lurie, et al. (2006). "United States' trends and regional variations in lumbar spine surgery: 1992-2003." *Spine* 31(23): 2707-14.

Welch, W. P., M. E. Miller, et al. (1993). "Geographic variation in expenditures for physicians' services in the United States." *N Engl J Med* 328(9): 621-7.

Wennberg, J. E. (1985). "Variations in medical practice and hospital costs." *Conn Med* 49(7): 444-53.

Wennberg, J. E., E. S. Fisher, et al. (2005). "Evaluating the efficiency of California providers in caring for patients with chronic illnesses." *Health Aff (Millwood)* Suppl Web Exclusives: W5-526-43.

Wennberg, J. E., E. S. Fisher, et al. (2004a). "Use of Medicare claims data to monitor provider-specific performance among patients with severe chronic illness." *Health Aff (Millwood)* Suppl Web Exclusives: VAR5-18.

Wennberg, J. E., E. S. Fisher, et al. (2004b). "Use of hospitals, physician visits, and hospice care during last six months of life among cohorts loyal to highly respected hospitals in the United States." *Bmj* 328(7440): 607.

Wennberg, J. E., J. L. Freeman, et al. (1989). "Hospital use and mortality among Medicare beneficiaries in Boston and New Haven." *N Engl J Med* 321(17): 1168-73.

Wennberg, J. E., N. Roos, et al. (1987). "Use of claims data systems to evaluate health care outcomes. Mortality and reoperation following prostatectomy." *Jama* 257(7): 933-6.

Werner, R. M. and E. T. Bradlow (2006). "Relationship between Medicare's hospital compare performance measures and mortality rates." *Jama* 296(22): 2694-702.

Wolinsky, F. D., J. Krygiel, et al. (2002). "Hospitalization for prostate cancer among the older men in the longitudinal study on aging, 1984-1991." *J Gerontol A Biol Sci Med Sci* 57(2): M115-21.

Wolinsky, F. D., T. E. Stump, et al. (1995). "Hospital utilization profiles among older adults over time: Consistency and volume among survivors and decedents." *J Gerontol B Psychol Sci Soc Sci* 50(2): S88-100.

Won, A., K. Lapane, et al. (1999). "Correlates and management of nonmalignant pain in the nursing home. SAGE Study Group. Systematic Assessment of Geriatric drug use via Epidemiology." *J Am Geriatr Soc* 47(8): 936-42.

Wright, G. E., W. E. Barlow, et al. (2007). "Differences among the elderly in the treatment costs of colorectal cancer: How important is race?" *Med Care* 45(5): 420-30.

Zechnich, A. D., M. Greenlick, et al. (1998). "Elimination of over-the-counter medication coverage in the Oregon Medicaid population: The impact on program costs and drug use." *Med Care* 36(8): 1283-94.

Zinn, J., Z. Feng, et al. (2008). "Restructuring in response to case mix reimbursement in nursing homes: a contingency approach." *Health Care Manage Rev* 33(2): 113-23.

11. Author Biography

Vincent Mor, PhD is Professor and Chair of the Department of Community Health at the Brown University School of Medicine and formerly served as the Director of the Brown University Center for Gerontology and Health Care Research. Dr. Mor has been on the faculty of the Department of Community Health since 1981, becoming tenured in 1987. Dr. Mor was one of the founders of the Department's graduate program in 1986 and directed it until becoming chair in 1996.

Dr. Mor has been Principal Investigator of over 20 NIH funded grants focusing on the organizational and health care delivery system factors associated with variation in use of health services and the outcomes experienced by frail and chronically ill persons. He has had multiple grants from the Robert Wood Johnson Foundation, the Pew Memorial Trust and the Retirement Research Foundation, as well as contracts from the Health Care Financing Administration and the Office of the Assistant Secretary for Planning and Evaluation to evaluate the impact of programs and policies in aging and long term care including Medicare funding of hospice, the costs and benefits of day-hospital treatment of cancer, patient outcomes in nursing homes, the impact of short term case management for cancer patients, several studies documenting age discrimination in cancer treatment and use of home care services, and a national study of residential care facilities. Over the past 20 years Dr. Mor's research has frequently integrated quantitative and qualitative data, particularly in program evaluations examining the approaches communities, organizations and specific providers use to adjust to health policy changes such as financing and reimbursement or to the emergence of integrated delivery systems.