Health Research Data for the Real World: The MarketScan Databases

David M. Adamson, PhD
Stella Chang, MPH
Leigh G. Hansen, MS, MBA

Research and Pharmaceutical Division
Thomson Medstat
January 2006

THOMSON
™
MEDSTAT

The MarketScan data warehouse, a family of databases, contains individual-level healthcare claims information from large employers, managed care organizations, Medicare, and Medicaid programs. This white paper describes the features and research uses of MarketScan data. Specifically, the paper examines the unique features of individual MarketScan databases, shows how the databases are constructed, describes their uses and data elements, and highlights examples of published studies based on MarketScan data.

Filling a Data Need: Origins of the MarketScan Databases

In the late 1980s, fundamental changes were transforming the U.S. healthcare system. In response to rising costs, healthcare delivery was shifting toward managed care arrangements. At the same time, there was growing interest in greater accountability for care through quality improvement. Many stakeholders sought data on these changes in the system and their impact on costs, quality of care, health outcomes, and the cost-effectiveness of various healthcare arrangements. In particular, employers and health plans, the purchasers and payers for the largest insured segment of the U.S. population (the privately insured), were interested in accurate and timely information on the drivers of cost growth and the return on investments for initiatives to improve employee health and well-being. Healthcare policy makers and practitioners were interested in the prevalence, incidence, and costs of specific diseases as well as the effectiveness and cost implications of interventions, clinical guidelines, and quality improvement initiatives. Providers, healthcare facilities, and pharmaceutical companies were interested in the cost-effectiveness of different therapies under normal care conditions.

Yet data sources to support the requisite analyses were inadequate in various ways. Data about specific conditions that came from randomized clinical trials yielded insight about the efficacy and direct costs of therapies under best-practice conditions, but did not generalize well to broader populations; nor did they provide insight into longer-term outcomes (such as returning to work or avoiding a recurrence of illness). In addition, few clinical trials collected data on costs. Data on U.S. care patterns, health trends, and costs were typically focused on special populations (such as government beneficiaries) or specific care settings (such as hospitals). Most of all, there was a lack of reliable healthcare data on the largest segment of U.S. healthcare users: privately insured patients and their families, who account for over 50% of the insured population.

THOMSON
★
MEDSTAT™

Figure 1: Population Distribution by Insurance Status — 2002
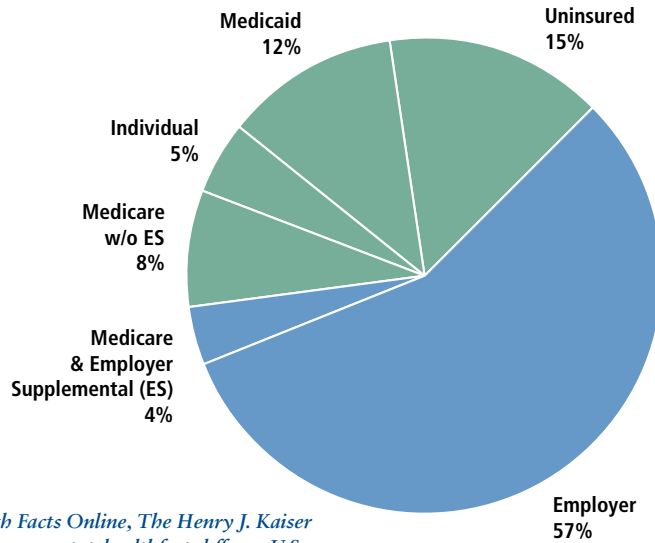
Source: State Health Facts Online, The Henry J. Kaiser Family Foundation, www.statehealthfacts.kff.org; U.S. residents – 285,007,110. Note: Percentages do not add to 100% because of rounding.

To address the need for better data on privately insured Americans, Thomson Medstat created the MarketScan® data collection. Since its creation, MarketScan has been expanded to include data on Medicare and Medicaid populations as well, making it one of the largest collections of claims-based patient data in the nation. MarketScan data reflect the real world of treatment patterns and costs by tracking millions of patients as they travel through the healthcare system, offering detailed information about all aspects of care. Data from individual patients are integrated from all providers of care, maintaining all healthcare utilization and cost record connections at the patient level.

Unique Features of MarketScan Data

MarketScan's claims-based data offer several distinct advantages over other types of data sources:

Large sample size. In its most recent data year, the MarketScan suite of databases contains information on nearly 28 million covered lives. The gender, age, and geographic distribution of the MarketScan population enables the creation of a nationally representative sample of Americans covered by health insurance (when the appropriate weights are applied).

Complete episodes of care. MarketScan databases capture the full continuum of care in all settings, including: physician office visits; hospital stays; retail, mail order, and specialty pharmacies; and carve-out care.

Strong longitudinal tracking at the patient level. More than half of the individuals in MarketScan can be followed for three years or more, a longer period than is typical for other claims databases. Employer-provided data also allow tracking of patients across health plans. This tracking ability is useful because people change health plans more often than they change jobs, and these data are able to capture patients who are "lost" in plan-based data sources — upward of 17% of patients in those databases.

Detailed prescription drug information. MarketScan contains complete information on outpatient prescriptions. These databases afford distinct advantages over others that track only prescription fills: MarketScan allows identification of type of disease (from medical claims) and can be used to determine whether clinical, demographic, and provider characteristics influence prescribing patterns. Because individual patients' prescription fills are recorded, therapies prescribed concurrently (and presumably used in combination) can also be identified. This provides vital information about actual drug use patterns, as opposed to individual drug prescription trends.

High-quality coding. A major advantage of MarketScan data involves their comprehensive and high quality coding. Key examples include:

- Diagnosis coded on 99% of all claims
- Procedure coded on 85% of physician claims
- Fully paid and adjudicated claims
- Complete payment/charge information, including amount of patient responsibility
- Complete outpatient prescription drug information, including patient copayments, mail order, injectables, specialty pharmacies, all carve-outs, manual *and* electronically submitted claims, and plan/formulary summaries.

Limitations of the data. As with any data source, MarketScan data have limitations. Some of these have to do with the nature of claims data, and others with the nature of the MarketScan sample population. Key limitations include the following:

- MarketScan is based on a large convenience sample. Because the sample is not random, it may contain biases or fail to generalize well to other populations. However, these data can complement other data sets or be used as benchmarks against them.
- The data come mostly from large employers; medium and small firms are not represented.
- The encounter data include information on lab tests, but not lab results. Lab results will be available in 2006.
- Accessing the data requires some kind of data management software. Our DataProbe® software and programmer support can facilitate access.

Numerous research applications. In combination, the features of MarketScan's databases enable analysts to conduct a broad range of health services studies, including:

- Cost-effectiveness and cost-offset studies
- Pharmacoeconomic outcomes evaluations
- Burden of illness analyses
- Surgical and pharmaceutical treatment comparisons
- Forecasting and modeling
- Assessment of best practices and benchmarking against empirical norms or clinical practice guidelines
- Clinical trial planning and support.

In the remainder of this white paper, we discuss MarketScan data and research applications in greater detail. We begin by describing how the MarketScan databases are built and the elements contained in each database. Following that, we provide highlights of studies conducted with MarketScan data. At the conclusion of the paper, we offer more information about how to obtain MarketScan.

Overview of MarketScan Data

How the MarketScan Data Sets Are Constructed

Medstat constructs the MarketScan warehouse by collecting data from employers, health plans, and state Medicaid agencies and placing them into databases. Data comprise service-level claims for inpatient and outpatient services and outpatient prescription drugs. All claims have been paid and adjudicated. We standardize financial, clinical, and demographic fields, and then add contributor-specific fields. Drug detail (e.g., therapeutic class, therapeutic group, manufacturer's average wholesale price, and a generic product identifier) and clinical detail (on disease episode grouper) are also added.

A unique enrollee identifier is assigned to each individual in MarketScan. This identifier is created by encrypting information provided by data contributors. This information includes the employee identifier, the relationship of the enrollee to the contract holder, the gender of the enrollee, and the enrollee's date of birth. Medstat then combines the standardized fields of the individual databases and creates links between years of data and across all data types. The data are collected for the MarketScan databases when nearly 100 percent of claims have been paid, eliminating the need for completion factors and improving the reliability and accuracy of the data.

Protecting the privacy of patient data as well as the privacy of our customers is a core principle of Medstat. The MarketScan research databases fully comply with the Health Insurance Portability and Accountability Act of 1996 (HIPAA). The MarketScan databases meet the criteria for a limited-use data set and contain none of the data elements prohibited by HIPAA for limited-use data sets. In addition, Medstat has taken steps to go beyond these HIPAA requirements. For example, instead of providing five-

THOMSON
MEDSTAT

digit ZIP Codes for employees and providers in the databases, we now provide three-digit ZIP Codes. Furthermore, the MarketScan databases underwent a statistical analysis by a third party to verify that they met HIPAA requirements for fully de-identified data sets. While meeting these requirements is optional given the current MarketScan licensing process, this additional step further demonstrates the Medstat commitment to HIPAA compliance and to protecting the confidentiality of patient-level and provider-level data. All patient-level and provider-level data within the MarketScan research databases contain synthetic identifiers to protect the privacy of individuals and data contributors.

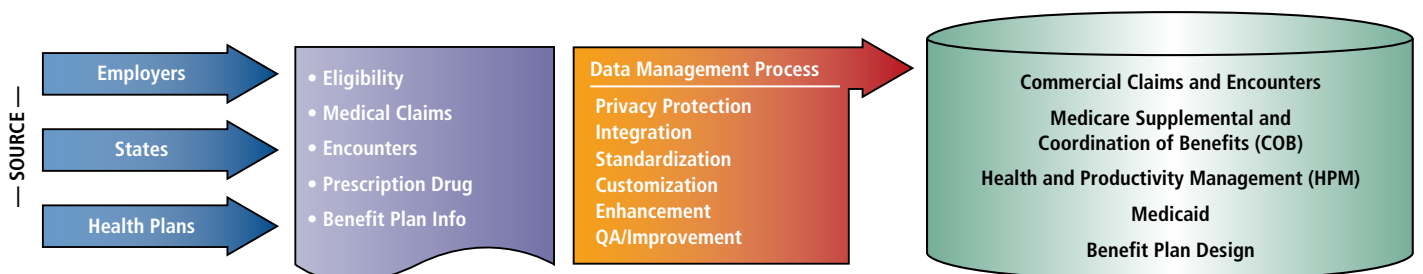Additional enhancements to the data during database creation include:

- Comparing diagnosis and procedure codes to codes that were in effect when the raw data were collected; and editing of the diagnosis and procedure codes, if necessary
- Adding the care provider's Metropolitan Statistical Area (MSA) to claims integration of benefit plan characteristics, enrollment, outpatient prescription drug, and medical/surgical data
- Adding Major Diagnostic Categories (MDCs) and Diagnosis Related Groups (DRGs) to claims, plus application of other classification systems, such as Outpatient Treatment Groups and Medstat Disease Staging Software®
- Identifying the type of plan, such as health maintenance organizations (HMOs), preferred provider organizations (PPOs), and point-of-service (POS) or comprehensive plans
- Verifying that both the experience (claims) and the denominator populations (eligible enrollees) exist for all sets of data contributed to the database.

The Market Scan Warehouse — Five Fully Integrated Databases

The end product is one of the nation's largest collections of patient data, featuring:

- An opportunity sample from multiple sources (employers, states, health plans)
- 28 million covered lives
- More than 50 contributing employers (mostly U.S. Fortune 500 firms)
- Representation from 80 health plan
- Representation from eight states.

Figure 2: MarketScan Databases: Fully Integrated at the Patient Level

— SOURCE —

| Employers | States | Health Plans |

- Eligibility
- Medical Claims
- Encounters
- Prescription Drug
- Benefit Plan Info

Data Management Process
Privacy Protection
Integration
Standardization
Customization
Enhancement
QA/Improvement

Commercial Claims and Encounters
Medicare Supplemental and Coordination of Benefits (COB)
Health and Productivity Management (HPM)
Medicaid
Benefit Plan Design

THOMSON
MEDSTAT™

The MarketScan warehouse consists of five databases (see Figure 2). These are described in detail below.

The **MarketScan Commercial Claims and Encounters Database** consists of employer- and health plan-sourced data containing medical and drug data for several million individuals annually. More than 13 million individuals are included in the 2004 database, encompassing employees, their spouses, and dependents who are covered by employer-sponsored private health insurance. Healthcare for these individuals is provided under a variety of fee-for-service (FFS), fully capitated, and partially capitated health plans, including preferred and exclusive provider organizations (PPOs and EPOs), point of service plans, indemnity plans, and health maintenance organizations (HMOs). Medical claims are linked to outpatient prescription drug claims and person-level enrollment information.

The Commercial database is constructed by combining, standardizing, and enhancing the databases Medstat builds on behalf of large employers and health plans nationwide. Sample data elements are shown in Table 1.

Understanding the total health cost of a particular medication — as opposed to the direct cost of medication alone — is critical to assessing cost-effectiveness. Why?

Table 1: Sample Data Elements for Commercial and Medicare Databases

Demographic	Medical Information (Inpatient and Outpatient)	Health Plan Features	Financial Information	Drug Information	Enrollment Information
Patient ID	Admission date and type	Coordination of benefits amount	Total payments	Generic product ID	Date of enrollment
Age	Principal diagnosis code	Deductible amount	Net payments	Average wholesale price	Member days
Gender	Discharge status	Copayment amount	Payments to physician	Prescription drug payment	Date of disenrollment
Employment status and classification (hourly, etc.)	Major diagnostic category	Plan type	Payment to hospital	Therapeutic class	
Relationship of patient to beneficiary	Principal procedure code		Payments—total admission	Days supplied	
Geographic location (state, ZIP Code)	Secondary diagnosis codes (up to 14)			National drug code	
Industry	Secondary procedure codes (up to 14)			Refill number	
	DRG			Therapeutic group	
	Length of stay				
	Place of service				
	Provider ID				
	Quantity of services				

Because a more expensive drug therapy may also produce better outcomes, ultimately reducing long-term medical costs. For example, a study compared the cost of two different prescription therapies for treating depression: selective serotonin reuptake inhibitors (SSRIs) and tricyclic antidepressants (TCAs). MarketScan data revealed that the two-year average prescription cost of TCAs was lower, but the overall cost of treatment using TCAs was higher.

Figure 3: Managing populations — Cost-Offset for Treating Depression with SSRIs vs. TCAs
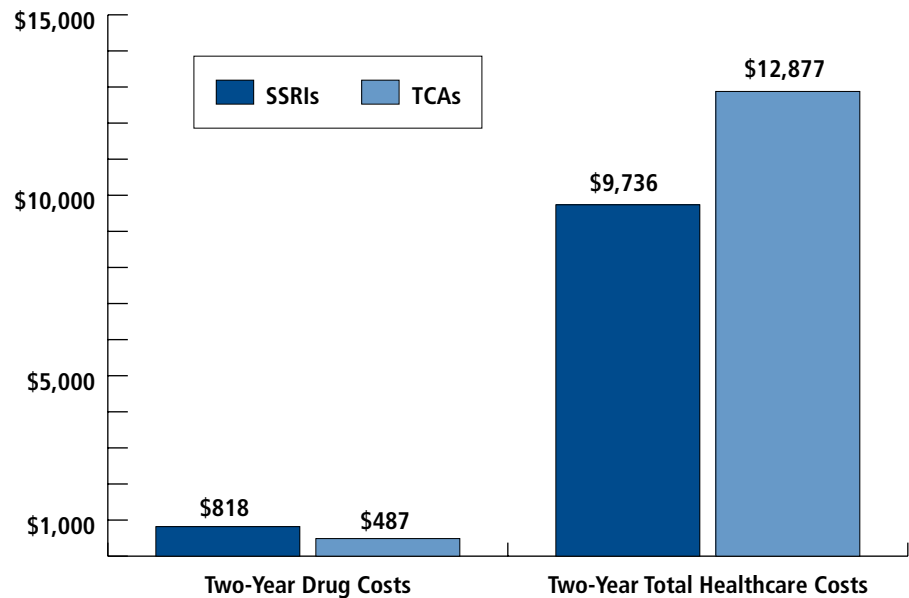
The **MarketScan Medicare Supplemental and Coordination of Benefits** (**Medicare**) **Database** is the first in the U.S. to profile the healthcare experience of retirees with Medicare supplemental insurance paid for by employers. The database includes the Medicare-covered portion of payment (represented as Coordination of Benefits Amount, or COB), the employer-paid portion, and any patient out-of-pocket expenses. The Medicare database provides detailed cost, use, and outcomes data for healthcare services performed in both inpatient and outpatient settings. For most of the population, the medical claims are linked to outpatient prescription drug claims and person-level enrollment data through the use of unique patient or enrollee identifiers.

Beneficiaries in the MarketScan Medicare database have drug coverage; therefore, drug data are available and provide additional valuable information. This feature makes the database a powerful tool for pharmacoeconomic and outcomes research and provides valuable insight into the drug use and spending patterns of older Americans. In addition, this drug data feature can address the same set of questions as the Commercial database.

The data elements in this database are the same as those appearing in the Commercial database, but pertain to patients with Medicare supplemental insurance.

With the implementation of Medicare Part D, Medstat asked its customers who contribute to the Medicare database if they planned to continue to offer Medicare Supplemental coverage. Of the 53 customers responding in the spring of 2005, for Plan Year 2006:

- 62% of responding employers indicated that they would take the 28% subsidy and continue to offer cover to retirees
- 8% will discontinue current coverage
- 6% have applied to become a Prescription Drug Plan
- 6% will provide wrap-around coverage
- 18% were still considering their options at the time information was collected

The **MarketScan Health and Productivity Management (HPM) Database** provides the opportunity to combine data on workplace absence, short-term disability, and workers' compensation with medical/surgical claims and outpatient drug data. The database allows researchers to assess both the direct and indirect costs associated with a particular condition or treatment.

Using this data set, researchers can:

- Assess the direct and indirect costs associated with a clinical condition
- Measure the impact of diseases on absenteeism, short-term disability, and workers' compensation
- Track total healthcare costs across both medical and workers' compensation systems
- Estimate the potential return on investments in wellness or disease management programs
- Assess the impact of a child's or spouse's illness on employee absence
- Determine the relative costs of alternative pharmaceutical and medical device interventions, considering both group medical costs and absenteeism costs
- Develop predictive models that help define relationships between demographic factors and HPM outcomes.

Sample data elements for the MarketScan Health and Productivity Management database are presented in Table 2.

Table 2: Sample Data Elements in MarketScan Health and Productivity Management Database

Demographic Information	Short-Term Disability	Workers' Compensation	Workplace Absence	Health Plan Features	Financial Information	Inpatient & Outpatient Medical Information	Drug Information
Patient ID	Case days	Body part injured	Dates and hours of absence	Coordination of benefits amount	Total payments	Admission date and type	Generic product ID
Age	Disability type	Case diagnosis	Absence type (sick, disability, vacation, etc.)	Deductible amount	Net payments	Principal diagnosis code	Average wholesale price
Gender	Case diagnosis	Indemnity payments		Copayment amount	Payments to physician	Discharge status	Prescription drug payment
Employment status and classification (hourly, etc.)	Total payments	Case days		Plan type	Payments to hospital	Major diagnostic category	Therapeutic class
Relationship of patient to beneficiary		Cause of injury			Payments total admission	Principal procedure code	Days supplied
Geographic location (state, ZIP Code)		Medical payments				Secondary diagnosis codes (up to 14)	National drug code
Industry						Secondary procedure codes (up to 14)	Refill number
						DRG	Therapeutic group
						Length of stay	
						Place of service	
						Provider ID	
						Quantity of services	

The **MarketScan Medicaid Database** contains the medical, surgical, and prescription drug experience of 13 million Medicaid enrollees from multiple states. It includes records of inpatient services, inpatient admissions, outpatient services, and prescription drug claims, as well as information on long-term care and other medical care. Data on eligibility (by month) and service and provider type are also included. In addition to standard demographic variables such as age and gender, the database includes variables of particular value to researchers investigating Medicaid populations, such as aid category (blind/disabled, Medicare eligible) and race.

THOMSON
MEDSTAT™

Using this database alone or in conjunction with other MarketScan research databases, Medstat researchers can:

- Analyze disease conditions that are especially prevalent among Medicaid populations, such as HIV/AIDS, schizophrenia, and diseases of the elderly
- Assess trends in healthcare cost, utilization, and outcomes for diseases that strike broadly across all populations, such as asthma, cancer, and cardiovascular conditions
- Incorporate variables not available in other claims databases, such as race and aid category
- Determine the cost burden of particular diseases or conditions in Medicaid populations.

The **MarketScan Benefit Plan Design (BPD) Database** contains detailed information about benefit plan characteristics for a subset of the health plans represented in the Commercial database and the Medicare database. This information, which is abstracted from summary plan description booklets, includes financial provisions, health service benefits, managed care features, and health coverage types.

The Benefit Plan Design database allows researchers to:

- Evaluate the impact of health plan features on healthcare utilization
- Assess the relative performance of plan types with varying managed care features
- Include detailed plan provisions — such as copayments, deductibles, and coverage options — in analysis of healthcare cost and use
- Measure changes in plan design and benefit characteristics from 1995 onward.

Sample data elements for the Benefit Plan Design database are presented in Table 3.

Table 3: Sample Data Elements for MarketScan Benefit Plan Design Database

Financial Provisions	Health Service Benefits	Managed Care Features	Health Coverage Types
Maximum out-of-pocket (family and individual)	Home healthcare	Precertification requirements/Penalties	Point of Service (capitated and non-capitated)
Annual and lifetime limits (family and individual)	Extended care/Skilled nursing facility	Utilization review requirements	Health Maintenance Organization
Coinsurance levels	Hospice	Second surgical opinion requirements	Basic/Major Medical
Copayments	Prescription drug	Case management requirements	Comprehensive
Deductibles (family and individual)	Mental health/Substance abuse	Formulary utilization	Preferred Provider Organization
	Physical therapy	Mental health and substance abuse carve-out provisions	
		In-network incentives	
		Out-of-network penalties	
		Preventive care	

Additional Data Tools

Medstat has also created software tools that can be used in conjunction with MarketScan data to increase analytic power. These include Medstat Episode Grouper®, Medstat Disease Staging Software, Sample Select, Disease Profiler, and DataProbe, described in more detail below.

Disease Staging
Disease Staging is a clinically validated risk-adjustment methodology developed by Medstat, for studying the impact of illness severity on care complications, treatment patterns, resource consumption, and costs.

Episode Grouper
Episode Grouper is a methodology that enables researchers to analyze a patient's inpatient, outpatient, and pharmaceutical costs by episode of care.

Sample Select
Sample Select provides Internet access to the most recent five years of MarketScan Commercial and Medicare supplemental data for querying counts of patient cohorts based on disease, diagnosis, and/or procedures. This desktop tool enables researchers to quickly access patient population counts to access research protocols and to gather quick facts. Summary reports provide demographic, clinical, and utilization details on the selected population.

Disease Profiler
Disease Profiler provides Internet-based reports that present summary statistics on more than 600 disease categories. The reports are based on the Commercial and Medicare databases described above. Overall prevalence rates (per 1,000 lives) and mean annual payments (per patient) are provided. Disease Profiler also includes the metrics indicated below, by database and year. Currently, reports are available for 1998 through 2003 on a range of topics, for individual diseases including:

- Prevalence of condition by age group and gender
- Healthcare utilization and payments by place of service
- Mean annual healthcare payment percentages
- Top 5 therapeutic classes
- Top 10 drugs prescribed
- Top 10 comorbidities.

DataProbe
DataProbe is a PC-based software package designed for data analysis to support research and decision making. DataProbe facilitates analyses of multiple data sources, including the MarketScan databases. DataProbe offers users access to healthcare data through applications that guide users in creating customized queries and reports. DataProbe is not data-specific nor does it require particular variables or data formats. It imports any flat file data set and provides several tools for analyzing, combining, and aggregating

data from multiple sources. Examples of such sources include public use files, Medicare, hospital discharge files, patient survey data, vital statistics, and the MarketScan databases.

MarketScan Data in Action: Highlights of Studies

With more than 90 publications appearing in major peer-reviewed journals in the last five years, MarketScan databases are among the most published proprietary databases in the United States, used by government, university, and private-sector researchers. MarketScan data have supported a range of health services research and analysis, including studies of:

- Economic burden of illness: costs and resource utilization
- Evaluation of treatment and interventions
- The effects of healthcare coverage and managed care
- Benchmarking and performance measurement
- Health and workforce productivity
- Pharmacoeconomic analysis
- Methodological studies.

Each of these areas is discussed in more detail and exemplified below:

Economic burden of illness: costs and resource utilization

Using the MarketScan databases, researchers can track complete episodes of care for patients and their families. Therefore, the data enable analysis of the direct and indirect costs of specific diseases — such as cancer, depression, or diabetes — as well as other health conditions, such as heroin addiction (discussed below). Two recent examples of research drawing on MarketScan data have studied the economic burden of particular conditions.

Analysts used MarketScan data from the Commercial, Medicare, and HPM databases to estimate the cost burden of cancer. Prior to the study, most estimates for cancer costs had been based on data sources with significant limitations for generalizing to the broader U.S. population, such as Medicare data or survey results that included only inpatient spending. This study focused on newly diagnosed patients with one of seven kinds of tumors in 1999 and 2000 and compared their direct and indirect costs with those of a matched control group. Results revealed that the total cost burden of cancer during the study years was substantial — ranging from $2,187 for prostate cancer to $7,616 for pancreatic cancer. By comparison, costs for the control group averaged $329 per month. The study also demonstrated the utility of using administrative data for deriving tumor-specific estimates of cancer costs, suggesting that policy makers charged with creating a national estimate of cancer costs should broaden the customary range of data sources to obtain a more comprehensive picture.

Reference: Chang, S., S.R. Long, L. Kutikova, L. Bowman, D. Finley, W.H. Crown, and C.L. Bennett. "Estimating the Cost of Cancer: Results on the Basis of Claims Data Analyses for Cancer Patients Diagnosed with Seven Types of Cancer During 1999 to 2000." *Journal of Clinical Oncology* 22(17) (2004): 3524–3530.

Another study examined the economic costs of heroin addiction in the United States. Although heroin use represents only a small proportion of total illicit drug use, its highly addictive nature suggests that its economic impact could surpass a less intensely addicting substance like marijuana. Using an array of data sources, including MarketScan data, the analysis estimated the total cost of heroin addiction in the U.S. in 1996. The results showed that estimated total costs of heroin addiction during the study period were $21.9 billion. Of these costs, productivity losses accounted for the greatest share — 53% — followed by criminal activities (24%), medical care (23%), and social welfare spending (0.5%). The large economic burden resulting from heroin addiction suggests the importance of investing in prevention and treatment.

Reference: Mark, T.L., G.E. Woody, T. Juday, and H. Kleber. "The Economic Costs of Heroin Addiction in the United States." *Drug and Alcohol Dependence* 61(2) (2001): 195–206.

Evaluation of treatment and outcomes

MarketScan data enable analysts to examine shifts in treatment patterns and assess their impact on a range of outcomes. These data also allow comparative assessment of different interventions.

For example, there are concerns that providers are shifting mental health patients from inpatient to outpatient care as a strategy to contain costs. If this is occurring, the shift could have implications for the quality of care such patients receive. A study used mental health claims data from 1993 to 1995 to examine whether this shift has occurred, the magnitude of any cost savings, and what components of care might account for savings. The study found that both inpatient and outpatient costs and utilization had fallen during the study period. Inpatient mental health costs fell by 30%, driven primarily by decreases in hospital days. Outpatient costs also fell (by 15%), as did costs for patients using both kinds of care (by 14%). The study concluded that there has not been a significant shift in the pattern of mental healthcare; rather, an overall reduction of care for mental health patients has occurred.

Reference: Leslie, D.L. and R. Rosenheck. "Shifting to Outpatient Care? Mental Health Use and Cost Under Private Insurance." *The American Journal of Psychiatry* 156 (1999): 8.

THOMSON
MEDSTAT

Another study examined the impact of new treatments on patients with Type 2 diabetes. In the 1990s, two new medications for treating hyperglycemia (high blood sugar associated with Type 2 diabetes) appeared on the market. The two new medications could be prescribed simultaneously for patients. A study used data from MarketScan sources to examine whether prescribing patterns for hyperglycemia changed concurrently with the introduction of these new medications and whether these changes were associated with changes in the underlying patient population treated for this condition. The results suggested that from 1997 through 2000, prescription patterns for patients with hyperglycemia shifted in a manner parallel with introduction of the new treatments.

Reference: Cohen, F.J., C.A. Neslusan, J.E. Conklin, and X. Song. "Recent Antihyperglycemic Prescribing Trends for U.S. Privately Insured Patients with Type 2 Diabetes." *Diabetes Care* 26(6) (2003): 1847–1851.

The effects of health coverage and managed care

MarketScan data contain information on health plan types and features. These details allow researchers to examine utilization patterns in different types of plans, including health maintenance organizations (HMOs), preferred provider organizations (PPOs), and fee-for-service (FFS). The data also contain benefit information, which can be analyzed for their impact on utilization and cost.

For example, managed care strategies that seek to contain costs by reducing hospital lengths of stay (LOS) have come under criticism. Are managed care patients experiencing shorter hospital stays and does this raise concerns about quality of care under such plans? To examine this issue, researchers used MarketScan data to analyze the impact of different health insurance plans on lengths of stay for hospitalizations related to one specific condition: cervical cancer. Because cervical cancer is more likely to affect younger women, the MarketScan data were particularly useful because they reflect a population slightly younger than that of the U.S. overall. The study found that after controlling for other variables such as severity of illness, no significant differences emerged in LOS for patients covered by comprehensive FFS plans versus those covered by other types of plans, including managed care.

Reference: Berg, G.D. and S.K. Chattopadhyay. "Determinants of Hospital Length of Stay for Cervical Dysplasia and Cervical Cancer: Does Managed Care Matter?" *American Journal of Managed Care* 10(1) (2004): 33–38.

MarketScan data also allow analysis of the impact of copayments. One increasingly common feature of drug benefit packages is a three-tiered copayment: the first tier includes generic medications, the second includes specific branded medicines designated as preferred-formulary, and the third includes other non-preferred brands. Patients have an incentive to choose a generic drug, which costs less than the brand-name products. Drugs in the third tier are non-preferred and therefore more expensive. Evidence has shown that

this arrangement can reduce drug spending. Less well understood, however, is whether three-tiered copayments influence the choice of medication or restrict the access of high-risk populations to recommended therapies. A recent study examined the impact of a three-tiered copay arrangement on the use of non-steroidal anti-inflammatory drugs (NSAIDs) among patients with arthritis. In particular, the study focused on the use of recently introduced COX-2 selective inhibitors, which several consensus guidelines recommend for arthritic patients at greatest risk of gastrointestinal complications, but which typically cost more than the older drug treatments. The study employed data from the 2000 MarketScan databases (specifically, Commercial Claims and Encounters, Medicare, and Benefit Plan Design). The findings suggested that the three-tiered arrangement was influencing the selection of medications, and that patients may have been unaware that the more expensive alternative had therapeutic or safety advantages. The analysts concluded that more information about the relative health and safety benefits of different drugs should be made available both to doctors and patients in the context of tiered copayment schemes.

Reference: Briesacher, B., S. Kamal-Bahl, M. Hochberg, D. Orwig, and K.H. Kahler. "Three-Tiered-Copayment Drug Coverage and Use of Nonsteroidal Anti-inflammatory Drugs." *Archives of Internal Medicine* 164(15) (2004):1679–1684.

Benchmarking and performance measurement

MarketScan data resources allow benchmarking analysis to compare the relative health of different populations or the quality of treatment for specific conditions.

One study used the MarketScan databases along with data from the Department of Veterans Affairs (VA) to compare the quality of drug treatment for patients with a specific mental health condition: schizophrenia. The analysis focused on the use of antipsychotic medications and whether dosing adhered to established treatment recommendations. Results revealed no statistical difference between the two populations in terms of the percentage of patients receiving recommended dosages (60% for the VA; 58% among privately insured patients). However, VA patients were more likely either to be dosed above or below guideline-recommended amounts. The study noted that there was room for improvement in the treatment of both populations in terms of adhering to guidelines.

Reference: Leslie, D.L. and R.A. Rosenheck. "Benchmarking the Quality of Schizophrenia Pharmacotherapy: A Comparison of the Department of Veterans Affairs and the Private Sector." *The Journal of Mental Health Policy and Economics* (Italy) 6(3) (2003): 113–121.

MarketScan data can also support performance measurement analyses. For example, one study measured the short-term cost of adhering to clinical guidelines. The analysis focused on relationships between adherence to seven guidelines for treating diabetes

and medical spending based on 1996 data. The study found that overall expenditures during the year were $713 higher if diabetes patients received all guideline-based treatments. There were two important exceptions: adherence to the suggested frequency of hemoglobin A1c blood sugar tests and of eye exams was associated with significantly lower expenditures. One implication of the study was that further research was needed to assess the potential for longer-term payoffs of guideline adherence.

Reference: Ozminkowski, R., S. Wang, W.D. Marder, and J. Azzolini. "Short-run Associations Between Medical Care Expenditures and Adherence to Clinical Practice Guideline-Based Measures for Diabetes." *Value in Health* 3(s1) (2000): 29–38.

Health and workforce productivity

MarketScan workforce data allow researchers to assess the impact of specific conditions and particular courses of treatment on health maintenance and job productivity.

A team of Medstat researchers examined the health and productivity costs of the "Top 10" physical and mental health conditions affecting six large U. S. employers. The analysis was intended as a first step in identifying high-cost conditions that are most amenable to intervention. By far the most expensive condition for U.S. employers in the study year (1999) was angina pectoris (chronic chest pain). Costs for this condition averaged $235 across all eligible employees. The next closest was hypertension, which costs $160 per employee. The most expensive mental health condition was bipolar disorder, which costs $64 across all eligible employees.

Reference: Goetzel, R.Z., K. Hawkins, R.J. Ozminkowski, and S. Wang. "The Health and Productivity Cost Burden of the 'Top 10' Physical and Mental Health Conditions Affecting Six Large U.S. Employers in 1999." *Journal of Occupational and Environmental Medicine* 45(1) (2003): 5–14.

Another study examined the impact of treatment for depression on workplace productivity. Because one goal of antidepressant therapy is to return patients to normal functioning, an important measure of effectiveness is restoring normal work capability, as measured in terms of days present on the job. The study analyzed a population of employees diagnosed with depression. It recorded absences for the six-month period prior to treatment and the six-month period following for two groups of patients: those treated with tricyclic medication and those treated with selective serotonin reuptake inhibitors (SSRIs). The study found that anti-depressant therapy was effective in reducing absenteeism for both groups.

Reference: Claxton, A.J., A.J. Chawla, and S. Kennedy. "Absenteeism Among Employees Treated for Depression." *Journal of Occupational and Environmental Medicine* 41(7) (1999): 605–611.

Pharmacoeconomics

Pharmacoeconomic analysis focuses on understanding the impact of pharmaceutical use on treatment costs. In most cases, a treatment followed consistently over a long time will cost more than one followed sporadically or for a short period. However, the longer, more consistent treatment may produce better clinical outcomes, thus reducing the long-term utilization of medical resources. Therefore, a complete understanding of how drug use affects treatment costs requires examining the direct costs of spending on drugs as well as indirect cost offsets that may reduce the overall use of resources. MarketScan data combine clinical encounter and prescription drug data to enable this kind of analysis.

One study examined the impact of drug selection and adherence to guidelines in treating depression. The study compared three different SSRI drug treatments for depression. The results showed that patients on one of the three treatments were more likely to be treated according to consensus guidelines; however, adherence to these guidelines did not produce significantly higher costs.

Reference: Crown, W., M. Treglia, L. Meneades, and A. White. "Long-Term Costs of Treatment for Depression: Impact of Drug Selection and Guideline Compliance." *Value in Health* 4(4) (2001): 295–307.

MarketScan data also support analysis of trends in drug costs. In recent years, the price of prescription drugs has increased rapidly as a share of overall medical spending. Researchers sought to determine the drivers of cost growth in drug use. The study employed MarketScan data on prescription drug use in conjunction with other data to study the cost of treatments for seven common health problems. The analysis sought to disentangle the relative contributions of price increases and volume growth in explaining the rising cost of drugs. Results revealed that price increases accounted for only a small part of cost growth and that growth in the volume of prescribing accounted for a much larger share of the increase. The analysis identified several trends driving this increase in volume: increased prevalence of disease as a result of new knowledge and improved clinical practice; demographic shifts toward an older population that requires more treatment; shifts in the mix of existing therapies toward more costly treatments; increasing quantities of medication per patient; and introduction of new therapies for patients who were previously untreated.

Reference: Dubois, R.W., A.J. Chawla, C.A. Neslusan, M.W. Smith, and S. Wade. "Explaining Drug Spending Trends: Does Perception Match Reality?" *Health Affairs* 19(2) (2000): 231–239.

THOMSON

MEDSTAT

Methodological analysis

MarketScan data can also be used to advance methodological expertise or to support studies that improve analytic capabilities or inform practice.

For example, a team of Medstat analysts measured how well current risk-adjustment systems can predict individuals' future healthcare costs based on current health characteristics. The ideal for such systems is not to overstate future costs, which creates incentives for plans to bar sicker patients, or to understate costs, which may invite overuse of services. It is especially important to understand how well risk-adjustment systems predict the future costs of people with chronic conditions and whether certain kinds of systems work better for certain conditions. The study focused on patients with potentially disabling chronic conditions and compared actual expenditures with those that various risk-adjustment models would have predicted. It found that the use of existing risk-adjustment systems would reduce incentives for health plans to discourage higher-risk enrollees, but not eliminate them. Nevertheless, risk-adjustment would provide a more accurate alternative for predicting future medical costs than simpler formulas based on age, gender, or wages.

Reference: Mark, T.L., R.J. Ozminkowski, A. Kirk, S.L. Ettner, and J. Drabek. "Risk Adjustment for People with Chronic Conditions in Private Sector Health Plans." *Medical Decision Making* 23(5) (2003): 397–405.

Another study set out to define an episode of care. Episode-of-care calculations can be important for informing treatment algorithms or end-point assessments in clinical trials. Using MarketScan claims data from 1993-95, the analysis focused on a specific condition — diabetic foot ulcer, a well-understood and costly complication of diabetes. The study compared two utilization measures: a daily average of patient charges versus a proportion of patients charging. The results showed that the latter measure was a more accurate indicator of an episode of care.

Reference: Mehta, S.S., S. Suzuki, H.A. Glick, and K.A. Schulman. "Determining an Episode of Care Using Claims Data. Diabetic Foot Ulcer." *Diabetes Care* 22(7) (1999): 1110–1115.

THOMSON

MEDSTAT

Using and Obtaining the MarketScan Databases

The MarketScan databases offer a powerful, flexible resource for healthcare research. The databases have several distinctive features:

- Fully integrated, patient-level data are pooled from diverse points-of-care, reflecting the true continuum and cost of healthcare (including the indirect costs).
- The longitudinal tracking of patient data from all sources of care is the strongest in the industry.
- Use in more than 90 studies published in peer-reviewed journal articles during the past five years places MarketScan among the most published healthcare databases in the United States.

For more information on how to obtain the MarketScan research databases for your healthcare research, please contact Medstat at 866-301-5419 or by e-mail at medstat.marketscan@thomson.com. Customized data sets and licensing agreements are available to suit specific data needs.

Selected Bibliography of Studies Using MarketScan Data

Economic burden of illness: costs and resource utilization

Chang, S., S.R. Long, L. Kutikova, L. Bowman, D. Finley, W.H. Crown, and C.L. Bennett. "Estimating the Cost of Cancer: Results on the Basis of Claims Data Analyses for Cancer Patients Diagnosed with Seven Types of Cancer During 1999 to 2000." *Journal of Clinical Oncology* 22(17) (2004): 3524–3530.

Crystal-Peters, J., C.A. Neslusan, M.W. Smith, and A. Togias. "Healthcare Costs of Allergic Rhinitis-Associated Conditions Vary with Allergy Season." *Annals of Allergy, Asthma & Immunology* 89(5) (2002): 457–462.

Hu, X., L. Markson, R. Lipton, W. Stewart, and M. Berger. "Burden of Migraine in the United States: Disability and Economic Costs." *Archives of Internal Medicine* 159(8) (1999): 813–818.

Mark, T.L., G.E. Woody, T. Juday, and H. Kleber. "The Economic Costs of Heroin Addiction in the United States." *Drug and Alcohol Dependence* 61(2) (2001): 195–206.

Stang, P.E., W.H. Crown, R. Bizier, M.L. Chatterton, and R. White. "The Family Impact and Costs of Migraine." *The American Journal of Managed Care* 10(5) (2004): 313–320.

Evaluation of treatment and intervention

Berg, G.D. and S.K. Chattopadhyay. "Determinants of Hospital Length of Stay for Cervical Dysplasia and Cervical Cancer: Does Managed Care Matter?" *American Journal of Managed Care* 10(1) (2004): 33–38.

Berndt, E.R., A. Bir, S.H. Busch, R.G. Frank, and S.L. Normand. "The Medical Treatment of Depression, 1991–1996: Productive Inefficiency, Expected Outcome Variations, and Price Indexes." *Journal of Health Economics* 21(3) (2002): 373–396.

Cohen, F.J., C.A. Neslusan, J.E. Conklin, and X. Song. "Recent Antihyperglycemic Prescribing

Trends for U.S. Privately Insured Patients with Type 2 Diabetes." *Diabetes Care* 26(6) (2003): 1847–1851.

Kong, S.X., H.T. Hatoum, S.Z. Zhao, N.M. Agrawal, and S.G. Geis. "Prevalence and Cost of Hospitalization for Gastrointestinal Complications Related to Peptic Ulcers with Bleeding or Perforation: Comparison of Two National Databases." *American Journal of Managed Care* 4(3) (1998): 399–409.

Mark, T.L. "The Costs of Treating Persons with Depression and Alcoholism Compared with Depression Alone." *Psychiatric Services* 54(8) (2003): 1095–1097.

Effects of health coverage and managed care

Briesacher, B., S. Kamal-Bahl, M. Hochberg, D. Orwig, and K.H. Kahler. "Three-Tiered-Copayment Drug Coverage and Use of Nonsteroidal Anti-inflammatory Drugs." *Archives of Internal Medicine* 164(15) (2004):1679–1684.

Crown, W.H., E.R. Berndt, O. Baser, S.N. Finkelstein, W.P. Witt, J. Maguire, and K.E. Haver. "Benefit Plan Design and Prescription Drug Utilization Among Asthmatics: Do Patient Copayments Matter?" *Frontiers in Health Policy Research* 7 (2004): 95–127.

Goodman, A.C., J.R. Hankin, E. Nishiura, and J.J. Sloan. "Impacts of Insurance on the Demand and Utilization of Drug Abuse Treatment: Implications for Insurance Mandates." *International Journal of the Economics of Business* 6(3) (1999): 331–348.

Kamal-Bahl, S. and B. Briesacher. "How Do Incentive-Based Formularies Influence Drug Selection and Spending for Hypertension?" *Health Affairs* 23(1) (2004): 227–236.

Leslie, D.L. and R. Rosenheck. "Shifting to Outpatient Care? Mental Health Use and Cost Under Private Insurance." *The American Journal of Psychiatry* 156 (1999): 8.

Benchmarking and performance measurement

Busch, S.H., D.L. Leslie, and R.A. Rosenheck. "Comparing the Quality of Antidepressant Pharmacotherapy in the Department of Veterans Affairs and the Private Sector." *Psychiatric Services* 55(12) (2004): 1386–91.

Leslie, D.L. and R.A. Rosenheck. "Comparing Quality of Mental Healthcare for Public-Sector and Privately Insured Populations." *Psychiatric Services* 51(5) (2000): 650–655.

Leslie, D.L. and R.A. Rosenheck. "Benchmarking the Quality of Schizophrenia Pharmacotherapy: A Comparison of the Department of Veterans Affairs and the Private Sector." *The Journal of Mental Health Policy and Economics* (Italy) 6(3) (2003): 113–121.

Ozminkowski, R., S. Wang, W.D. Marder, and J. Azzolini. "Short-run Associations Between Medical Care Expenditures and Adherence to Clinical Practice Guideline-based Measures for Diabetes." *Value in Health* 3(s1) (2000): 29–38.

Health and workforce productivity

Claxton, A.J., A.J. Chawla, and S. Kennedy. "Absenteeism Among Employees Treated for Depression." *Journal of Occupational and Environmental Medicine* 41(7) (1999): 605–611.

Crystal-Peters, J., W.H. Crown, R.Z. Goetzel, and D.C. Schutt. "The Cost of Productivity Losses Associated with Allergic Rhinitis." *The American Journal of Managed Care* 6(3) (2000): 373–378.

Goetzel, R.Z., K. Hawkins, R.J. Ozminkowski, and S. Wang. "The Health and Productivity Cost Burden of the 'Top 10' Physical and Mental Health Conditions Affecting Six Large U.S.

Employers in 1999." *Journal of Occupational and Environmental Medicine* 45(1) (2003): 5–14.

Goetzel, R.Z., S.R. Long, R.J. Ozminkowski, K. Hawkins, S. Wang, and W. Lynch. "Health, Absence, Disability, and Presenteeism Cost Estimates of Certain Physical and Mental Health Conditions Affecting U.S. Employers." *Journal of Occupational and Environmental Medicine* 46(4) (2004): 398–412.

Pharmacoeconomic studies

Croghan, T., C. Melfi, W. Crown, and A. Chawla. "Cost-Effectiveness of Antidepressant Medication." *The Journal of Mental Health Policy and Economics* 1 (1998): 109–117.

Dubois, R.W., A.J. Chawla, C.A. Neslusan, M.W. Smith, and S. Wade. "Explaining Drug Spending Trends: Does Perception Match Reality?" *Health Affairs* 19(2) (2000): 231–239.

Hellinger, F.J. and W.E. Encinosa. "Antiretroviral Therapy and Healthcare Utilization: A Study of Privately Insured Men and Women with HIV Disease." *Health Services Research* 39(4 Pt 1) (2004): 949–967.

Ozminkowski, R., S. Wang, W.D. Marder, J. Azzolini, and D. Schutt. "Cost Implications for the Use of Inhaled Anti-Inflammatory Medications in the Treatment of Asthma." *Pharmacoeconomics* 18(3) (2000): 253–264.

Methodological analysis

Crown, W., R. Obenchain, L. Englehart, T. Lair, D. Buesching, and T. Croghan. "The Application of Sample Selection Models to Outcomes Research: The Case of Evaluating Effects of Antidepressant Therapy on Resource Utilization." *Statistics in Medicine* 17 (1998): 1943–1958.

Lestina, D.C., T.R. Miller, and G.S. Smith. "Creating Injury Episodes Using Medical Claims Data." *The Journal of Trauma* 45(3) (1998): 565–569.

Mark, T.L., R.J. Ozminkowski, A. Kirk, S.L. Ettner, and J. Drabek. "Risk Adjustment for People with Chronic Conditions in Private Sector Health Plans." *Medical Decision Making* 23(5) (2003): 397–405.

Mehta, S.S., S. Suzuki, H.A. Glick, and K.A. Schulman. "Determining an Episode of Care Using Claims Data. Diabetic Foot Ulcer." *Diabetes Care* 22(7) (1999): 1110–1115.

Rosen, A.K., R.L. Houchens, T.B. Gibson, and A. Mayer-Oakes. "Developing Episodes of Care for Adult Asthma Patients: A Cautionary Tale." *American Journal of Medical Quality* 13 (1998): 25–35.

For a more complete listing of the MarketScan bibliography, please e-mail medstat.marketscan@thomson.com and place "Request for MarketScan bibliography" in the subject line.

For more information about MarketScan, please call 866-301-5419 or e-mail medstat.marketscan@thomson.com.

About The Thomson Corporation and Medstat

The Thomson Corporation (www.thomson.com), with 2004 revenues of $8.10 billion, is a global leader in providing integrated information solutions to business and professional customers. Thomson provides value-added information, software tools and applications to more than 20 million users in the fields of healthcare, scientific research, law, tax, accounting, financial services, higher education, reference information, and corporate e-learning and assessment. With operational headquarters in Stamford, Conn., Thomson has approximately 40,000 employees and provides services in approximately 130 countries. The Corporation's common shares are listed on the New York and Toronto stock exchanges (NYSE: TOC; TSX: TOC).

Medstat (www.medstat.com) is a Thomson Healthcare brand (www.thomson.com/healthcare) which provides market intelligence and benchmark databases, decision support solutions, and research services for managing the cost and quality of healthcare. Medstat applies these capabilities to improve policy and management decision making for employers, government agencies, health plans, hospitals and provider networks, and pharmaceutical companies.

THOMSON

MEDSTAT