HEALTH RESEARCH DATA FOR THE REAL WORLD: THE MARKETSCAN DATABASES

LEIGH G. HANSEN, MS, MBA
STELLA CHANG, MPH
JULY 2011
FILLING A DATA NEED: ORIGINS OF THE MARKETSCAN RESEARCH DATABASES

In the late 1980s, fundamental changes transformed the U.S. healthcare system. In response to rising costs, healthcare delivery shifted 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.

At the time, 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 an estimated 51 percent of the total population.
To address the need for better data on privately insured Americans, Truven Health Analytics created the MarketScan data warehouse. Since its creation, this warehouse has evolved into a suite of proprietary databases containing the largest collection of employer-based patient data in the United States. MarketScan claims 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.

Over the years, the original claims-centric databases have been clinically enriched, fully integrating health and productivity (workplace absence, short-term disability, workers’ compensation), lab test results, and health risk assessment data. In 2006, Truven acquired the Solucient® hospital databases. These datasets have become part of the MarketScan family.

**FIGURE 1 – U.S. Population Distribution by Insurance Status – 2009**

![Pie chart showing U.S. population distribution by insurance status with percentages: Employer 49%, Medicare 12%, Medicaid/Other Public 17%, Individual 5%, Uninsured 17%](source: The Kaiser Family Foundation state-healthfacts.org. Accessed December 2010. Note: Percentages do not add to 100% due to rounding.)

**UNIQUE FEATURES OF MARKETSCAN CLAIMS DATA**

MarketScan claims databases offer several distinct advantages over other types of data sources.

**Large sample size**
The MarketScan claims databases offer the largest convenience sample available in proprietary databases with 143 million unique patients since 1996. In the most recent full data year, MarketScan claims databases contain data on 50 million covered lives. Its sample size is large enough to allow creation of a nationally representative data sample of Americans with employer-provided health insurance and Medicaid. In addition, hospital discharge data contribute over 53 million inpatient records available for research.

**Complete episodes of care**
MarketScan claims 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. Linking hospital discharge records with claims data at the patient level has significantly increased the capability of MarketScan data to capture the continuity of a patient’s drug therapy between the inpatient and outpatient setting.

**Strong longitudinal tracking at the patient level**
The stability of MarketScan data sources allows superior continuity of patients over multiple years, generally longer than other claims databases. This is due to the majority of MarketScan data sourced from large employers. 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 percent of patients in those databases.
Detailed prescription drug information

The MarketScan claims databases contain complete information on outpatient prescriptions. (The linked claims-hospital data file described below contains inpatient drug information linked to outpatient prescriptions.) These databases afford distinct advantages over others that track only prescription fills. MarketScan data allow identification of the type of disease (from medical claims) and can be used to determine whether clinical, demographic, and provider characteristics influence prescribing patterns. Individual patients' prescription fills are recorded so 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.

The MarketScan Hospital Drug Database provides researchers with inpatient drug utilization data derived from hospital discharge records. These data and a proprietary projection methodology allow researchers to understand drug use in the inpatient and outpatient environment, including hospital use patterns, switching behavior, combination therapy, and patient characteristics. This information is used to help determine if introduction or earlier use of a product would improve clinical and overall cost outcomes and to analyze diagnosis volumes.

When researchers need to understand the impact of a hospitalization on prescription drug use, the MarketScan Inpatient Drug Link data file links outpatient data from claims with inpatient drug data from the Hospital Drug Database.

High-quality coding

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

- 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 claims 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:

- The MarketScan claims databases are 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 datasets or be used as benchmarks against them.
- The data come mostly from large employers; medium and small firms are not represented.
- Accessing the data requires data management software. DataProbe®, MarketScan Online Tools — Sample Select, Sample Select Prevalence, Inpatient View, Outpatient View, Disease Profiler, Treatment Pathways — and programmer support can facilitate access.

Numerous research applications

In combination, the features of MarketScan claims databases enable analysts to conduct a broad range of health services studies, including:

- Comparative effectiveness research
- 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 data.
OVERVIEW OF MARKETSCAN CLAIMS DATA

How the MarketScan datasets are constructed
Truven constructs the MarketScan claims databases by collecting data from employers, health plans, and state Medicaid agencies. 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 a MarketScan claims database. 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. We then combine the standardized fields of the individual databases and create links between years of data and across all data types. The data are collected for the MarketScan annual 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 Truven, so 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 dataset and contain none of the data elements prohibited by HIPAA for limited-use datasets. In addition, Truven has taken steps to go beyond these HIPAA requirements. For example, instead of providing five-digit ZIP codes for employees and providers in the databases, we 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 datasets. While meeting these requirements is optional given the current MarketScan licensing process, this additional step further demonstrates the Truven commitment to HIPAA compliance and to protecting the confidentiality of patient-level and provider-level data.

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 Disease Staging
• 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 MARKETSCAN WAREHOUSE: NINE FULLY INTEGRATED CLAIMS 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)
• Over 28 billion patient records
• 122 million covered lives
• 150 contributing employers; 21 contributing health plans
• Representation from over 130 unique carriers
The MarketScan family consists of nine claims databases (see Figure 2) and one hospital discharge database. These are described in further detail below.

1. 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. Just under 40 million individuals are included in the 2009 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, health maintenance organizations (HMOs), and consumer-directed health plans. 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 Truven builds on behalf of large employers and health plans nationwide. Sample data elements are shown in Table 1.

### TABLE 1: Sample Data Elements for Commercial Database and Medicare Supplemental Database

<table>
<thead>
<tr>
<th>DEMOGRAPHIC</th>
<th>MEDICAL INFORMATION (INPATIENT AND OUTPATIENT)</th>
<th>HEALTH PLAN FEATURES</th>
<th>FINANCIAL INFORMATION</th>
<th>DRUG INFORMATION</th>
<th>ENROLLMENT INFORMATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient ID</td>
<td>Admission date and type</td>
<td>Coordination of benefits amount</td>
<td>Total payments</td>
<td>Generic product ID</td>
<td>Date of enrollment</td>
</tr>
<tr>
<td>Age</td>
<td>Principal diagnosis code</td>
<td>Deductible amount</td>
<td>Net payments</td>
<td>Average wholesale price</td>
<td>Member days</td>
</tr>
<tr>
<td>Gender</td>
<td>Discharge status</td>
<td>Copayment amount</td>
<td>Payments to physician</td>
<td>Prescription drug payment</td>
<td>Date of disenrollment</td>
</tr>
<tr>
<td>Employment status and classification (hourly, etc.)</td>
<td>Major diagnostic category</td>
<td>Plan type</td>
<td>Payments to hospital</td>
<td>Therapeutic class</td>
<td></td>
</tr>
<tr>
<td>Relationship of patient to beneficiary</td>
<td>Principal procedure code</td>
<td>Payments – total admission</td>
<td>Days supplied</td>
<td>National drug code</td>
<td></td>
</tr>
<tr>
<td>Geographic location (state, ZIP code)</td>
<td>Secondary diagnosis codes (up to 14)</td>
<td></td>
<td></td>
<td>Refill number</td>
<td></td>
</tr>
<tr>
<td>Industry</td>
<td>Secondary procedure codes (up to 14)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DRG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Length of stay</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Place of service</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Provider ID</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quantity of services</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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? 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 (see Figure 3).

The Commercial Database offers a distinct advantage over other databases for research on medication use. As these data are primarily sourced from employers, claims for mail order prescriptions and specialty pharmacy are included to fully capture prescription fills from all locations. Comprehensiveness of drug data from all sources is particularly important for adherence studies and those looking at injectable drugs.

2. The MarketScan Medicare Supplemental and Coordination of Benefits (Medicare Supplemental) Database is the first in the United States to profile the healthcare experience of retirees with Medicare supplemental insurance paid for by employers. In 2010, 18 percent of the 46.5 million Medicare beneficiaries received their drug benefits through a retiree coverage plan (http://facts.kff.org/chart.aspx?cb=58&scnt=164&ch=1743. Accessed December 2010.). The database includes the Medicare-covered portion of payment (represented as Coordination of Benefits Amount, or COB), the employer-paid portion, and any out-of-pocket patient expenses.

The Medicare Supplemental 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 Medicare Supplemental 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.

3. 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 dataset, 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 Health and Productivity Management Database are presented in Table 2.

| TABLE 2: Sample Data Elements in the Health and Productivity Management Database |
|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| Demographic Information         | Short-Term Disability           | Workers’ Compensation           | Health Plan Features            | Inpatient & Outpatient Medical Information | Drug Information                |
| Age                             | Case diagnosis                  | Indemnity payments              | Coordination of benefits amount | Admission date and type              | Generic product ID              |
| Gender                          | Case diagnosis                  | Copayment amount                | Deductible amount               | Net payments                        | Average wholesale price         |
| Employment status and classification (hourly, etc.) | Total payments                  | Plan type                        | Payments to hospital            | Discharge status                    | Prescription drug payment       |
| Relationship of patient to beneficiary | Cause of injury                | Payments total admission         | Principal procedure code        | Major diagnostic category           | Therapeutic class               |
| Geographic location (state, ZIP Code) | Medical payments               | Principal diagnosis code         | Days supplied                   | National drug code                  |                                |
| Industry                        |                                | Secondary diagnosis codes (up to 14) |                                | Refill number                       |                                |
|                                 |                                | Secondary procedure codes (up to 14) |                                |                                 |                                |
|                                 |                                | DRG                              |                                 |                                 |                                |
|                                 |                                | Length of stay                   |                                 |                                 |                                |
|                                 |                                | Place of service                 |                                 |                                 |                                |
|                                 |                                | Provider ID                      |                                 |                                 |                                |
|                                 |                                | Quantity of services             |                                 |                                 |                                |

4. The Marketscan Medicaid Multi-State Database contains the medical, surgical, and prescription drug experience of nearly 30 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.

Using this database alone or in conjunction with other Marketscan databases, 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
5. **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 Medicare Supplemental 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>
<thead>
<tr>
<th>FINANCIAL PROVISIONS</th>
<th>HEALTH SERVICE BENEFITS</th>
<th>MANAGED CARE FEATURES</th>
<th>HEALTH COVERAGE TYPES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum out-of-pocket (family and individual)</td>
<td>Home healthcare</td>
<td>Pre-certification requirements/ Penalties</td>
<td>Point of Service (capitated and non-capitated)</td>
</tr>
<tr>
<td>Annual and lifetime limits (family and individual)</td>
<td>Extended care/Skilled nursing facility</td>
<td>Utilization review requirements</td>
<td>Health Maintenance Organization</td>
</tr>
<tr>
<td>Coinsurance levels</td>
<td>Hospice</td>
<td>Second surgical opinion requirements</td>
<td>Basic/Major Medical</td>
</tr>
<tr>
<td>Copayments</td>
<td>Prescription drug</td>
<td>Case management requirements</td>
<td>Comprehensive</td>
</tr>
<tr>
<td>Deductibles (family and individual)</td>
<td>Mental health/Substance abuse</td>
<td>Formulary utilization</td>
<td>Preferred Provider Organization</td>
</tr>
<tr>
<td></td>
<td>Physical therapy</td>
<td>Mental health and substance abuse carve-out provisions</td>
<td>Consumer-Directed Health Plan</td>
</tr>
</tbody>
</table>

6. **The MarketScan Lab Database** helps researchers understand:
- How well a drug is performing in the real-world clinical setting
- Diagnostic test results administered prior to initiation of drug therapy
- Lab test results as indicators of results of drug therapy
- Frequency of performing safety monitoring lab tests while a patient is on drug therapy
- Differences in treatment patterns between patients whose disease is under control compared to patients whose disease is not

The Lab Database includes inpatient, outpatient, drug, enrollment, and lab test results for patients during the time period from 2004 through 2010.

7. **The MarketScan Health Risk Assessment (HRA) Database** provides specialized data allowing researchers to understand the contribution of patient behaviors to health outcomes. Health risk assessments can also be invaluable for researchers as they provide self-reported data on clinical variables that may be otherwise unavailable. Like other MarketScan databases, the HRA Database standardizes and links HRA data with the claims experience of patients presenting an opportunity for innovative research. In addition to medical and drug claims, absence, short-term disability, and workers’ compensation data, HRAs can provide key data inputs for analyzing the health and productivity of patient cohorts. There is significant overlap between the HRA Database and HPM Database, enriching health and productivity management studies. Researchers looking at diabetes, cardiovascular disease, insomnia, and smoking cessation find these data uniquely valuable.
8. The **MarketScan Dental Database** is the only integrated medical, drug, and dental database of its kind. Linking dental claims with medical claims, this database allows researchers to conduct research on the relationship between dental care, use of pharmaceuticals for oral health, and patients’ medical conditions such as:

- Respiratory tract infections
- Chronic sinus infections
- Diabetes
- Chronic acid reflux
- Liver or kidney problems
- Infective endocarditis
- Cardiovascular disease
- Pre-term birth

The Dental Database provides data covering the entire continuum of healthcare including services received in the dental office. The current release covers 2005–2009. It includes 6.7 million patients with dental and medical claims.

Sample data elements for the Dental Database are presented in Table 5.

---

### TABLE 5: Dental Database Sample Data Elements

<table>
<thead>
<tr>
<th>DEMOGRAPHIC</th>
<th>CLINICAL</th>
<th>PROVIDER</th>
<th>FINANCIAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Dental procedure (ADA coding)</td>
<td>Specialty</td>
<td>Total payment</td>
</tr>
<tr>
<td>Gender</td>
<td>Procedure group</td>
<td>Setting of care</td>
<td>Deductible</td>
</tr>
<tr>
<td>Geography</td>
<td>Service type</td>
<td></td>
<td>Copay</td>
</tr>
<tr>
<td>Type of dental coverage</td>
<td></td>
<td></td>
<td>Coinsurance</td>
</tr>
<tr>
<td>Type of medical coverage</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug coverage</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

9. The **MarketScan Hospital Drug Database**, is derived primarily from hospital billing systems from 593 hospitals. This database provides the most detailed and comprehensive data available for understanding hospital care, including drug utilization in the inpatient setting. The database includes more than 53 million hospital discharges between January 2002 and March 2011. Data are updated monthly with a 45-day lag after the close of the month.

A unique feature of this database is the proprietary projection methodology allowing data users to forecast market potential, monitor product uptake and share, determine if earlier use of a product would improve clinical and overall cost outcomes, and analyze diagnosis volume and drug ordering by physician specialty. A full spectrum of semi-custom data reports are supported by the Hospital Drug Database including those used to understand hospital market share, combination therapy, switching behavior, and for profiling patients.
Sample data elements for the Hospital Drug Database are presented in Table 6.

<table>
<thead>
<tr>
<th>TABLE 6: MarketScan Hospital Drug Database Sample Data Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PATIENT</strong></td>
</tr>
<tr>
<td>Projected number of patients</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Principal/Secondary diagnosis codes</td>
</tr>
<tr>
<td>Principal/Secondary procedure codes</td>
</tr>
<tr>
<td>Admission type</td>
</tr>
<tr>
<td>Admission source</td>
</tr>
<tr>
<td>Length of stay (calculated)</td>
</tr>
</tbody>
</table>

10. The MarketScan Inpatient Drug Link File, nicknamed the Hospital Spillover File, helps to answer research questions regarding the effect of an inpatient stay on drug utilization. The file matches patients from MarketScan claims data (Commercial, Medicare Supplemental, and Medicaid) and hospital discharge records. These data help researchers understand:

- Drug use (spillover), switching, and adherence between settings of care
- Pre- and post-hospitalization treatment
- Repeated hospitalizations
- Health outcomes
- Drug-specific and/or total healthcare costs

Claims data frame the picture of the continuum of care pre-, during, and post-hospitalization providing rich cross-sectional and longitudinal details about patient treatment patterns. Hospital discharge data provide the inpatient drug component resulting in enriched insights into the transition between inpatient and outpatient treatment.

**FIGURE 4: Tracking Prescription Drug Treatment Through a Hospitalization**

Inpatient Drug Link Data File

<table>
<thead>
<tr>
<th>Pre-admission treatment</th>
<th>Hospital setting</th>
<th>Post-discharge treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outpatient care</td>
<td>Inpatient care</td>
<td>Outpatient care</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>Diagnosis</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>Procedure</td>
<td>Procedure</td>
<td>Procedure</td>
</tr>
<tr>
<td>Drug</td>
<td>Drug</td>
<td>Drug</td>
</tr>
</tbody>
</table>
MARKETSCAN DATA TO MATCH YOUR RESEARCH POINT OF VIEW

MarketScan data releases
The MarketScan Commercial Database and Medicare Supplemental Database are updated on a quarterly and annual basis. Quarterly releases include Early View and Standard Updates. The Annual File is a calendar-year release. The Health Risk Assessment, Health and Productivity Management, Benefit Plan, Dental, and Lab databases are released annually, while the Multi-State Medicaid Database is released semi-annually. The Hospital Drug Database is updated monthly.

The release schedule in Figure 5 illustrates how the Early View, Standard Updates, and Annual File work together to provide the most current and complete data available for researchers. Early View always provides the most timely data. Both Early View and Standard Updates link back to the MarketScan Annual File, each representing a subsequent time period. By bridging the time period, Standard Updates are required to fill in patient history between Early View and the Annual File.

| FIGURE 5: MarketScan Data Coverage (effective July 2011) |
|---------------------------------|-----------------|----------------|
| 1995-2009                       | 2010            | 2011           |
| Q1 – Q4                         | Q1              | Q1             |
| Annual Releases                 | Standard Update | Early View     |

Minimum completion rates vary by data release
Researchers using healthcare claims data have to balance considerations of data completeness with the need for timely information. The completeness of claims databases is determined by three important dates. These include the:

- Service date: The date the patient received the service.
- Paid date: The date the claim was paid by the health plan.
- Pull date: The date the claim was pulled to construct the research database.

To understand how complete a claims database is, one must understand the impact of these three dates on completion rates and run-off time requirements for each data release.

Healthcare services for prescription drugs, inpatient admissions, facilities, and professional services are submitted to and paid by health plans at different rates. Drug claims are paid quite quickly, whereas inpatient admissions, facility, and professional claims take longer to process.

Completion rates refer to the cumulative percent of claims that have been paid by the health plan at a given point in time after the service date. The time period is measured in months. As Table 7 illustrates, the time period between the service date and the paid date varies by the type of claim. One month after the service date, 97.2 percent of drugs claims, 64.5 percent of inpatient admissions, 65.3 percent of facility claims, and 74.1 percent of professional claims have been paid.
### Table 7: MarketScan Estimated Completion Rates

<table>
<thead>
<tr>
<th>Months from Service Date</th>
<th>Drug Claims (%)</th>
<th>Inpatient Admissions (%)</th>
<th>Facility Claims (%)</th>
<th>Professional Claims (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>63.6</td>
<td>15.1</td>
<td>21.4</td>
<td>34.9</td>
</tr>
<tr>
<td>1</td>
<td>97.2</td>
<td>64.5</td>
<td>65.3</td>
<td>74.1</td>
</tr>
<tr>
<td>2</td>
<td>97.7</td>
<td>81.6</td>
<td>79.9</td>
<td>84.3</td>
</tr>
<tr>
<td>3</td>
<td>97.8</td>
<td>87.0</td>
<td>86.1</td>
<td>88.5</td>
</tr>
<tr>
<td>4</td>
<td>97.9</td>
<td>89.8</td>
<td>89.5</td>
<td>91.2</td>
</tr>
<tr>
<td>5</td>
<td>98.0</td>
<td>91.9</td>
<td>91.8</td>
<td>93.1</td>
</tr>
<tr>
<td>6</td>
<td>98.1</td>
<td>93.7</td>
<td>93.6</td>
<td>94.5</td>
</tr>
<tr>
<td>7</td>
<td>98.3</td>
<td>94.9</td>
<td>94.9</td>
<td>95.6</td>
</tr>
<tr>
<td>8</td>
<td>98.5</td>
<td>95.7</td>
<td>96.0</td>
<td>96.5</td>
</tr>
<tr>
<td>9</td>
<td>98.7</td>
<td>96.6</td>
<td>96.8</td>
<td>97.3</td>
</tr>
<tr>
<td>10</td>
<td>98.9</td>
<td>97.3</td>
<td>97.5</td>
<td>97.9</td>
</tr>
<tr>
<td>11</td>
<td>99.2</td>
<td>97.9</td>
<td>98.2</td>
<td>98.5</td>
</tr>
<tr>
<td>12</td>
<td>99.5</td>
<td>98.4</td>
<td>98.7</td>
<td>99.0</td>
</tr>
<tr>
<td>13</td>
<td>99.7</td>
<td>99.0</td>
<td>99.2</td>
<td>99.4</td>
</tr>
<tr>
<td>14</td>
<td>99.8</td>
<td>99.5</td>
<td>99.6</td>
<td>99.7</td>
</tr>
<tr>
<td>15</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

When creating the MarketScan claims databases, the elapsed time between the service date and the paid date captured when data are pulled to create a release file affects the completeness of the database. The longer the period between the service date and last paid date, the more complete the database will be. Early View, Standard Updates, and the Annual File pull data with progressively longer time periods between the last service date and the last paid date, thus the Annual File is more complete than an Early View file. Each of our data releases has a specific run-off requirement. Run-off is the time between the last service month of the time period and the latest paid month available at the time data are pulled.

- Early View data has no minimum run-off requirement.
- Standard Updates require three months of run-off for the last month of data.
- The MarketScan Annual File requires a minimum of six months of run-off for the last month of data.

To match data completeness with your research requirements, you will want to select the right data release.

**Choose Early View when you need timely drug claims**

The Early View release provides the most recent claims data available. Early View may include zero to seven months of run-off.

Early View provides the most timely view of drug claims. One month after the service date, over 97 percent of drug claims are complete, as these claims are usually submitted electronically and paid quickly. The drug claim provides the ordering physician’s identification number (DEA number), National Drug Code (NDC), patient identification number, basic patient demographics, new/refill status, and payment. Early View, as its name implies, also provides a preview of all other claims before all claim types are fully paid.

For research on acute conditions treated in an office setting, Early View data provide high completion rates and high patient counts for the seven months immediately post-Standard Update. Early View can support research that is closer to real time.
Choose Standard Updates and the MarketScan Annual File when you need longitudinal integrity. The Standard Updates, released quarterly, include data with a minimum three months of run-off. This means that service dates in the last month of the release have the following completion rates: drug claims are 98 percent complete; inpatient admissions are 87 percent complete; facility claims are 86 percent complete; and professional claims are 89 percent complete. All other months in the release are more complete as additional time has elapsed since the service date.

Illustrated in Figure 5, Standard Updates bridge the gap between the calendar year Annual File and Early View. There are four Standard Update releases each year. The first release covers Q1. The second covers Q1 and Q2, up to the fourth quarter release which provides a preview of the full calendar year. The Q4 Standard Update, which is the first full look at a calendar year of data, has 15 months of run-off for January service dates and three months of run-off for December service dates.

The MarketScan Annual File is released with the most complete data available. The Annual File version one (v1) release is built when the last month, December, has six months of claims run-off (in contrast to the fourth quarter Standard Update, which has three months run-off). The six-month run-off period guarantees very high completeness of all claims even in the last month of the database.

The MarketScan Standard Updates plus the Annual File provide the best data for chronic disease and longitudinal studies for several reasons.
• The MarketScan Commercial Database and Medicare Supplemental Database are exceptionally large, supporting studies requiring smaller patient populations.
• These databases offer some of the best patient retention (i.e., continuous enrollment) percentages in the industry allowing researchers to follow patients from several years to more than a decade.
• For trend analyses, particularly pre- and post-a new intervention or therapy (e.g., a new immunization), MarketScan data go back to 1992 with Annual Files available for licensing from 1995–2008.
• The required run-off periods for these data releases provide for high completion rates for all types of claims.

ADDITIONAL DATA TOOLS

Software tools and methodologies can be used in conjunction with MarketScan data to increase analytic power and gain quick access to summarized information. These include Disease Staging, Medical Episode Grouper, MarketScan Sample Select, MarketScan Sample Select Prevalence, MarketScan Disease Profiler, DataProbe, MarketScan Inpatient View, MarketScan Outpatient View, and MarketScan Treatment Pathways, described in more detail below.

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

Medical Episode Grouper
Medical Episode Grouper is a methodology that enables researchers to analyze a patient’s inpatient, outpatient, and pharmaceutical costs by episode of care.
MarketScan Sample Select
Sample Select provides Internet access to the most recent five years of MarketScan Commercial and Medicare Supplemental, HPM, and HRA 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 assess research protocols and to gather quick facts. Summary reports provide demographic, clinical, and utilization details on the selected population.

MarketScan Sample Select Prevalence
Sample Select Prevalence uses the online Sample Select platform to quickly project the estimated prevalence of a treated condition or diagnosis among patients actively engaged in the U.S. healthcare system and covered by employer-sponsored insurance — an estimated 51 percent of Americans. Importantly, these estimates represent those patients seeking care and covered by an important payer in the healthcare arena. Data come from the MarketScan Commercial Database and Medicare Supplemental Database, and projections are based on the proprietary MarketScan Weights.

MarketScan Disease Profiler
Disease Profiler provides online reports that present summary statistics on more than 600 disease categories. Overall prevalence rates per 1,000 lives and mean annual payments per patient are provided based on the Commercial and Medicare databases. Disease Profiler includes the metrics indicated below by selected disease category.
• Prevalence of condition by age group and gender
• Healthcare utilization and payments by place of service
• Mean annual healthcare payment percentages
• Top five 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 them 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 dataset 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 Inpatient View
Inpatient View, another online tool, uses a comprehensive catalog of all U.S. hospital-based inpatient care to provide users with total diagnosis and procedure volumes and key statistics for the most recent year with trending information. For each diagnosis code, statistics are provided including detailed patient demographics, admissions data, length-of-stay distributions, cost, regional and facility distributions, patient disposition, and payer mix information.

Inpatient View includes a feature called Link View. Link View is comprised of a set of reports linking principal diagnoses, procedures, and diagnostic-related groups (DRGs) with the 50 most common corresponding or secondary diagnoses and procedures that appeared on patient discharge records from U.S. hospitals. Link View allows users to determine the frequency of a principal event occurring in conjunction with another event.

Data are derived from all-payer data gathered from 20 million actual inpatient records, representing approximately 55 percent of discharges from U.S. hospitals per year. This detail-rich database is the Projected Inpatient Database (PIDB), a proprietary Truven database, the largest all-payer inpatient database available. The PIDB supports publications, products, and custom studies, the results of which are applicable to all short-term, general, nonfederal hospitals in the United States. This exclusive database combines data from both public and proprietary state data as well as individual and group hospital contracts.
MarketScan Outpatient View
Outpatient View details the total annual volume for diagnoses and procedures by outpatient setting: ambulatory surgery, outpatient hospital, and physician offices. Volumes are detailed by region, age, gender, outpatient setting, and payer for the most recent year. Like Inpatient View, Outpatient View also provides a five-year forecast.

MarketScan Treatment Pathways
Treatment Pathways is a visual interface into MarketScan data allowing users to follow patients through their course of treatment over time – forward and backward – very rapidly and without programmer support. Treatment pathways include the medical, surgical, and drug services found in MarketScan data. The software allows users to graphically see the patient journey as a series of events sequenced into easily-drawn treatment tree maps. Treatment Pathways may be used to understand events leading up to a diagnosis, time to treatment, switching patterns, and outcome events.

MARKETSCAN DATA IN ACTION: HIGHLIGHTS OF STUDIES

The MarketScan Research Databases are exceptionally well published. The first publication appeared in 1990 in the New England Journal of Medicine. Since then, more than 425 articles have appeared in major peer-reviewed journals. MarketScan data have supported a range of health services research conducted by government, academic, and private researchers. These studies have been in the areas of:

• Comparative effectiveness research
• Economic burden of illness
• Clinical results
• Dental research
• The economic costs of health risks
• Health and workforce productivity
• Projected disease prevalence
• Benefit plan design and adherence
• Adverse events rates
• Outcomes of treatment
• Population studies

Comparative Effectiveness Research
The American Recovery and Reinvestment Act of 2009 earmarked substantial funds for comparative effectiveness research (CER). The Federal Coordinating Council, responsible for coordinating the initiative, defines CER as the conduct and synthesis of systematic research comparing different interventions and strategies to prevent, diagnose, treat, and monitor health outcomes.

The MarketScan databases are well suited to observational comparative effectiveness research whether comparing drugs and/or procedures, or comparing the effects of health policy on outcomes, population disparities, or prevention strategies.

In one MarketScan study, complication rates for two colorectal cancer treatments were analyzed by researchers using the Commercial Database and Medicare Supplemental Database. Capecitabine is an oral alternative to 5-fluorouracil, which has equal clinical efficacy and has a favorable safety profile. After adjusting for differences in patient demographics, clinical history, and treatment setting, the analysis found that even though the most frequently observed complications (nausea, vomiting, infection, anemia, neutropenia, and diarrhea) were the same for the two treatments, the probability of developing any of 23 studied complications per episode was greater in patients on 5-flurouracil monotherapy. There were no significant differences in total costs for acquisition, administration, and complications between the two regimens; however, for patients with colorectal cancer, capecitabine therapy was associated with lower complication rates.

Economic Burden of Illness

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, including cancer, depression, and diabetes.

Analysts used MarketScan data from the Commercial Database and Health and Productivity Management Database to quantify direct costs, utilization, and rate of comorbidities in patients with fibromyalgia, a poorly understood illness, and compare them to patients with rheumatoid arthritis. Employer absence and disability data linked at the patient level to the patient’s medical claims was a key benefit of using MarketScan data for the study. Patients with fibromyalgia exhibited higher use of the emergency room and physician and physical therapy visits than rheumatoid arthritis patients. Mean annual expenditures were about the same for patients with fibromyalgia and patients with rheumatoid arthritis — $10,911 and $10,716 respectively; however, costs were almost double for patients with both diseases. Indirect costs (absence and short-term disability) were similar. Researchers concluded that the burden of illness for fibromyalgia is substantial and comparable to that of rheumatoid arthritis.


In another study, researchers analyzed the direct healthcare costs for hip, vertebral, and non-hip, non-vertebral fractures using the Commercial Database and Medicare Supplemental Database to quantify first-year healthcare costs for patients 50 and older. Using data from July 1, 2001, to December 31, 2004, researchers found that the adjusted mean first-year per-patient cost for hip fractures was $26,545 for patients aged 50-64 and $15,196 for those aged 65 and older. For vertebral fractures, first-year per-patient costs were $14,977 and $6,701 for patients aged 50-64 and those aged 65 and older, respectively. For non-hip, non-vertebral fractures, first-year per-patient costs were $9,183 and $6,106, respectively, for the two patient cohorts. Although non-hip, non-vertebral fractures had lower total per-patient costs than either hip or vertebral fractures, overall costs were higher because the ratio of non-hip, non-vertebral fractures to hip, vertebral fractures was 11:1 for patients aged 50-64 and 2:1 for patients aged 65 and older.


Clinical Results

In some diseases, clinical results are surrogate markers for risk reduction and disease management. The MarketScan Lab Database provides researchers with the ability to conduct studies on laboratory outcomes using lab results data linked to patients’ claims.

For diabetes patients to achieve and maintain control, the American Diabetes Association guidelines recommend that A1c tests be repeated within three months following an out-of-range test result or a change in therapy. Truven researchers used the MarketScan Lab Database, which includes both commercial and Medicaid patients, to compare frequency of retesting between groups of patients at varying A1c levels based on an initial test, and to explore factors such as changes in therapy or insurance coverage associated with more frequent retesting.

The research found that retesting within six months is no more frequent after out-of-range tests or changes in therapy. Retesting rates were in fact lower for Medicaid patients than for those with commercial insurance. Thirty-five percent of Medicaid patients with an initial A1c result of greater than nine percent were retested versus 51.7 percent of commercial patients. These results supported a previous finding of clinical inertia in response to poor glycemic control.

This study won an ISPOR Best-Podium Award in 2007 at the 12th Annual International Conference.

**Dental Research**

Observational studies have suggested a relationship between cardiovascular disease and periodontitis due to systemic inflammation that may impair the vascular system. In October 2008, the American Academy of Periodontology reported that patients with periodontal disease were twice as likely to suffer from coronary artery disease as those without the disease. To our knowledge there are no commercially available research databases that link dental claims with medical claims at the patient level to study this relationship except for the MarketScan Dental Database.

Researchers from Truven examined the relationship between periodontal disease, statin (HMG-CoA) use, and cardiovascular disease using 2005-2007 data from the MarketScan Dental Database. Using a matched control group, patients identified as having periodontal disease by ICD-9 code or related dental procedures were divided into two cohorts — treated with statins and not treated with statins. These patients were followed for 12 months to observe evidence of cardiovascular events. Ninety-nine percent of patients with periodontal disease did not have a cardiovascular event. Twenty-five percent of patients with the disease were on statin medications. The rate of cardiovascular events for statin-users and non-statin users were 0.42 percent and 1.16 percent respectively. Controlling for differences in demographics and preexisting clinical conditions, researchers observed that patients without statin treatment were at a greater risk for cardiovascular events (OR = 2.77, 95 percent CI, p<0.0001) in the 12 months following the periodontal diagnosis than were treated patients.


**The Economic Cost of Health Risks**

MarketScan data offer unique patient-level linkages to data unavailable from administrative claims. The MarketScan Health Risk Assessment Database allows researchers to tie patient-reported risk data to patient claims data. With the epidemic of obesity in the United States, self-reported data found in this database can help to identify the societal burden of this ever-growing risk.

Researchers used MarketScan health risk assessment (HRA) and health and productivity management data to quantify the direct and indirect costs of obesity to U.S. self-insured employers. Body mass index (BMI) derived from HRA survey results allowed researchers to divide patients into risk groups as defined by the Centers for Disease Control and the World Health Organization. Direct costs from claims and indirect costs from these same patients in the Health and Productivity Management Database presented an analysis consistent with previous research — that patients classified as obese or severely obese had higher overall healthcare costs. Lost time due to obesity was higher in overweight, obese, and severely obese patients. This study provided current payer estimates attributable to BMI categories.


**Health and Workplace Productivity**

Absence, short-term disability, and workers’ compensation data linked to medical claims in the MarketScan Health and Productivity Management Database allow researchers to assess the impact of specific medical conditions and particular courses of treatment on health maintenance and job productivity.

In one study using this database, researchers examined the association between non-adherence to bipolar medications and lost productivity costs by employers. Adult patients with a biopolar diagnosis and at least one prescription claim for a mood stabilizer or atypical antipsychotic were selected for study. In the selected cohort, only 35.3 percent of patients were adherent to their medication as determined by a medication possession ratio of greater or equal to 80. Non-adherent patients had higher adjusted indirect costs of +$771.41 due to absence, +$285.72 in short-term disability, and +$360.62 in workers’ compensation. Extrapolating these findings to a fictional employer with 70,000 employees and an incidence rate of 3.3 percent for bipolar disorder, this employer could potentially save $578,378 in combined absence, short-term disability, and workers’ compensation indirect costs if all employee-patients were adherent to their bipolar treatment.

In another study, researchers sought to calculate productivity losses among depressed employees treated with anti-depressants. Through the examination of absence and short-term disability data for patients diagnosed with depression and a control cohort, researchers found that even when depressed patients were treated with anti-depressants there were still substantial productivity losses. Mean short-term disability costs were $713 higher for depressed patients than for controls and $1,345 higher for a sub-cohort of severely depressed patients than their controls. The marginal impact of treated patients was $377 for absenteeism.


Projected Disease Prevalence
Researchers at Truven have developed the MarketScan Weights that allow MarketScan Commercial and Medicare Supplemental data to provide reliable projections for the number of patients in the employer-insured (ESI) population seeking treatment for a disease or medical procedure of interest. This population represents over 50 percent of Americans and is a large healthcare payer segment. The MarketScan Weights combine the power of a large claims database, a convenience sample, with the representativeness of a probability sample such as the Medical Expenditure Panel Survey (www.meps.ahrq.gov).

Truven researchers looked at “the prevalence of antipsychotic use among privately insured patients ages 0-64” to understand the implications for education and safety studies. In the 1990s, second generation antipsychotics (atypicals) came on the market, presenting a better side effect profile than older typical antipsychotics. Recently, atypical medications have been shown to have other potential liabilities, and the FDA issued warnings regarding the cardiac, metabolic, cerebrovascular, and mortality risks associated with the use of antipsychotic medication in elderly patients with dementia. The study used Sample Select Prevalence to calculate antipsychotic medication use by age and diagnosis. Results estimated that one child in every 185 and one in every 100 adults aged 45-64 in the ESI population filled a prescription for an antipsychotic medication in 2005. The most commonly indicated diagnosis was bipolar disorder.


Benefit Plan Design and Adherence
MarketScan data contain information on health plan types and features such as patient cost-sharing amounts. These details allow researchers to examine utilization patterns in different types of plans. 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 increase patient cost-sharing may have an impact on patient adherence to treatment plans. Some employers have actually reduced copayments for certain maintenance drugs where it has been shown that lower out-of-pocket costs for patients improves drug adherence and reduces overall healthcare costs.

One such study looked at the impact of out-of-pocket expenses (copayments and coinsurance) on adherence for patients newly diagnosed with rheumatoid arthritis and treated with etanercept or adalimumab. Not surprisingly, medication adherence decreased as out-of-pocket expenses increased. Ninety-two percent of patients paid less than $20 per week for their medication.
As patient costs increased in increments of $5.50 per week, one week of therapy adherence was lost. When weekly costs exceeded $50 per week, patients were more likely to discontinue therapy than lower-paying patients.


In another study using patient-quarter data from the MarketScan Commercial Database, researchers used generalized estimating equations to determine the effects of patient cost-sharing on adherence to second generation anti-psychotic medications. As in the previous study, results demonstrated that higher cost-sharing was inversely associated with high adherence, particularly when cost-sharing levels were above $30. Higher cost sharing was associated with shorter time to discontinuation.


**Adverse Event Rates**

Due to their size and longitudinal integrity, the MarketScan databases enable epidemiologists to understand the context around adverse events.

One study analyzed healthcare claims data from six influenza seasons (2000-2006) to examine the safety of oseltamivir in influenza patients of all ages. Central nervous system (CNS) and neuropsychiatric events, occurring within 14 and 30 days following influenza diagnosis, were compared between those treated with oseltamivir and those who were not given the antiviral treatment. The treated and untreated cohorts were propensity-matched in each flu season. The study found that there was no increase in CNS and neuropsychiatric events in children, adolescents, and adults who were prescribed oseltamivir — in fact, the study found that these events were less likely to occur in the first 14 days.


**Outcomes of Treatment**

Characterized by robust longitudinal integrity and cross-sectional detail, the MarketScan databases are ideal for studying the economic and utilization outcomes of treatment over time. Continuous enrollment is strong in the MarketScan databases because of the predominance of employer-sourced data. Data in their current research format go back to 1995; however, data are available back to 1992.

The U.S. Centers for Disease Control and Prevention have conducted longitudinal studies such as the one below looking at the impact of vaccines on overall healthcare utilization. This study was designed to understand the impact of the varicella (chickenpox) vaccine. While incidence of the disease dropped substantially since varicella was recommended for routine immunization in 1995, incomplete data made it difficult to track medical visits and associated expenses related to the disease. Researchers used the MarketScan databases to conduct a retrospective, population-based study examining trends in rates and costs for varicella-related hospitalizations and ambulatory visits from 1994 to 2002. During this time period, hospitalizations due to varicella declined by 88 percent and ambulatory visits declined by 59 percent. Total estimated direct medical expenditures declined by 74 percent, representing a savings of $62.8 million.

This study was a MarketScan Award winner in 2006 and published in *The Journal of the American Medical Association*.

**Population Studies**

In a world of comparative effectiveness research, understanding differences in treatment response in patient subpopulations is critical. The MarketScan databases offer strength for examining patients within the Medicaid population as well as the commercially insured.

A study on lipid-lowering therapy was conducted using the MarketScan Medicaid Database to look at possible racial differences in switching, augmentation, and titration for Medicare/Medicaid dual-eligible patients. Studies have shown that African Americans have a higher risk of coronary heart disease morbidity/mortality and are less likely to achieve cholesterol level targets while on therapy than their Caucasian counterparts. In the study, estimates were developed using logistic regression models to determine the probability of patients switching to or augmenting with another lipid-lowering therapy, or titrating up their current therapy during a year. Results revealed that African American patients were less likely to switch, augment, or titrate up than Caucasian patients, suggesting a treatment disparity by race.


Vlahiotis A, Devine ST, Eichholz J, Kautzner A. Discontinuation rates and healthcare costs in adult patients starting generic versus brand SSRI or SNRI antidepressants in commercial health plans. *Journal of Managed Care Pharmacy* 2011; 17(2): 123–32.


Asche CV, Singer ME, Jhaveri M, Chung H, Miller A. All-cause healthcare utilization and costs associated with newly diagnosed multiple sclerosis in the United States. *Journal of Managed Care Pharmacy* 2010; 16(9): 703–12.


Barlev A, Song X, Ivanov B, Setty V, Chung K. Payer costs for inpatient treatment of pathologic fracture, surgery to bone, and spinal cord compression among patients with multiple myeloma or bone metastasis secondary to prostate or breast cancer. *Journal of Managed Care Pharmacy* 2010; 16(9): 693–702.


Evans-Lacko SE, Dosreis S, Kastelic E, Riley AW. Mental health service use before and after diagnosis of early-onset bipolar disorder. *Journal of Behavioral Health Services and Research* 2010; September 15; [Epub ahead of print].


Lee DW, Foster DA. The association between hospital outcomes and diagnostic imaging: Early findings. *Journal of the American College of Radiology* 2009; 6; 780–785.


ABOUT TRUVEN HEALTH ANALYTICS

Truven Health Analytics delivers unbiased information, analytic tools, benchmarks, and services to the healthcare industry. Hospitals, government agencies, employers, health plans, clinicians, and life sciences companies have relied on us for more than 30 years. We combine our deep clinical, financial, and healthcare management expertise with innovative technology platforms and information assets to make healthcare better. We collaborate with our customers to uncover and realize opportunities for improving quality, efficiency, and outcomes. With more than 2,000 employees, we have major offices in Ann Arbor, Michigan; Chicago; and Denver. Advantage Suite, Micromedex, ActionOi, Marketscan, and 100 Top Hospitals are registered trademarks of Truven Health Analytics.

truenhealth.com

©2012 Truven Health Analytics All rights reserved.

PH 11238 0612 TEMP

ADDITIONAL PAPERS THAT MAY BE OF INTEREST TO YOU

- MarketScan Bibliography: Lists all known peer-reviewed publications for studies using MarketScan data back to 1990.
- The Value of Understanding Health and Productivity Costs. The MarketScan Health and Productivity Management Database. A white paper.
- Clinical Laboratory Results Provide a More Complete Understanding of Patient Outcomes. MarketScan Lab Database. A white paper.
- MarketScan in Action: Epidemiological Studies. A white paper.

To receive these papers, please email your request to marketscan@truvenhealth.com.

USING AND OBTAINING THE MARKETSCAN DATABASES

The MarketScan databases offer a powerful, flexible resource for healthcare research. These claims 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 425 studies published in peer-reviewed journal articles places the MarketScan Research Databases among the most published in the United States

FOR MORE INFORMATION

For more information on how to obtain the MarketScan Research Databases for your healthcare research, please contact Truven Health Analytics by email at marketscan@truvenhealth.com. Customized datasets and licensing agreements are available to suit specific data needs.