IBM MARKETSCAN RESEARCH DATABASES
USER GUIDE

MULTI-STATE MEDICAID DATABASE
DATA YEAR 2017 EDITION

The authorized recipient of these materials shall treat the information contained therein as confidential proprietary information owned by IBM® Watson Health™. The recipient shall not disclose or permit to be disclosed, in full or in part, to third parties any information contained therein. No part of these materials may be reproduced or transmitted in any form or by any means electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from Watson Health.

Requests for permission to make copies of any part of this report should be mailed to:

IBM Watson Health
100 Phoenix Drive
Ann Arbor, Michigan 48108
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>IBM MarketScan Database Overview</td>
<td>1</td>
</tr>
<tr>
<td>OVERVIEW OF TABLES</td>
<td>4</td>
</tr>
<tr>
<td>Medical/Surgical Tables</td>
<td>4</td>
</tr>
<tr>
<td>Enrollment Tables (A, T)</td>
<td>5</td>
</tr>
<tr>
<td>Member Days (MEMDAYS)</td>
<td>5</td>
</tr>
<tr>
<td>OVERVIEW OF ENCOUNTER RECORDS</td>
<td>6</td>
</tr>
<tr>
<td>FINANCIAL VARIABLES</td>
<td>7</td>
</tr>
<tr>
<td>Medical/Surgical Financial Variables</td>
<td>8</td>
</tr>
<tr>
<td>Prescription Drug Financial Variables</td>
<td>10</td>
</tr>
<tr>
<td>Encounter Record Financial Variables</td>
<td>11</td>
</tr>
<tr>
<td>Adjustment Records</td>
<td>11</td>
</tr>
<tr>
<td>Unresolved Adjustments</td>
<td>12</td>
</tr>
<tr>
<td>PERSON-LEVEL IDENTIFIERS</td>
<td>13</td>
</tr>
<tr>
<td>Enrollee Identifiers</td>
<td>13</td>
</tr>
<tr>
<td>CLINICAL VARIABLES</td>
<td>14</td>
</tr>
<tr>
<td>MARKETSCAN DATABASE CONSTRUCTION</td>
<td>15</td>
</tr>
<tr>
<td>Data Quality</td>
<td>15</td>
</tr>
<tr>
<td>PLAN TYPE DEFINITIONS</td>
<td>18</td>
</tr>
<tr>
<td>Plan Type</td>
<td>18</td>
</tr>
<tr>
<td>GLOSSARY OF ACRONYMS, ABBREVIATIONS, AND TERMS</td>
<td>20</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>27</td>
</tr>
<tr>
<td>APPENDIX A: FREQUENTLY ASKED QUESTIONS</td>
<td>61</td>
</tr>
</tbody>
</table>
INTRODUCTION

IBM MARKETSCAN DATABASE OVERVIEW

The IBM® MarketScan® Research Databases capture person-specific clinical utilization, expenditures, and enrollment across inpatient, outpatient, prescription drug, and carve-out services. The data come from a selection of large employers, health plans, and government and public organizations. The MarketScan Research Databases link paid claims and encounter data to detailed patient information across sites and types of providers and over time. The annual medical databases include private-sector health data from approximately 350 payers. Historically, more than 20 billion service records are available in the MarketScan Databases. These data represent the medical experience of insured employees and their dependents for active employees, early retirees, Consolidated Omnibus Budget Reconciliation Act (COBRA) continuees, and Medicare-eligible retirees with employer-provided Medicare Supplemental plans.

The IBM MarketScan Research Databases are composed of seven individual databases, which are described below and summarized in Exhibit 1.

Commercial Claims and Encounters Database

The IBM MarketScan Commercial Claims and Encounters (CCAE) Database contains data from active employees, early retirees, COBRA continuees, and dependents insured by employer-sponsored plans (i.e., individuals not eligible for Medicare).

The database has the following table structure:

- Inpatient Admissions Table (I)
- Facility Header Table (F)
- Inpatient Services Table (S)
- Outpatient Services Table (O)
- Outpatient Pharmaceutical Claims Table (D)
- Long Term Care (L)
- Annual Enrollment Summary Table (A)
- Enrollment Detail Table (T)

Medicare Supplemental and Coordination of Benefits Database

The IBM MarketScan Medicare Supplemental and Coordination of Benefits (COB) Database (MDCR) is created for Medicare-eligible retirees with employer-sponsored Medicare Supplemental plans. This database contains predominantly fee-for-service plan data.

The Medicare Supplemental and COB Database table structure is identical to the CCAE Database table structure.

Both the Medicare-paid amounts and the employer-paid supplemental insurance amounts are included in this database. Only plans in which both the Medicare-paid amounts and the employer-paid amounts were available and evident on the claims were selected for this database.

Health and Productivity Management Database

The IBM MarketScan Health and Productivity Management (HPM) Database is an integrated database that contains absence, short-term disability, long-term disability, and worker’s compensation experiences. This information is linkable to the medical, pharmacy, and enrollment data in the MarketScan CCAE Database for these employees, making the resulting database a unique and valuable resource for examining health and productivity issues for an employed, privately insured population.

A separate User Guide is provided to customers licensing the HPM database.

Health Risk Assessment Database

The MarketScan Health Risk Assessment Database comprises data derived from surveys completed by employees. The Health Risk Assessment Database links with the Commercial Database by enrollee identifier and contains variables reflective of biometric information, self-assessed health, mental health, productivity, health behaviors, and intent to
change behavior. All of the medical and drug claims, as well as enrollment records, are linked with the health risk responses. In addition, for a subset of large employer contributors, enrollment, claims, and health risk data are linkable to the absence, short-term disability, and worker’s compensation records of the Health and Productivity Management Database.

Benefit Plan Design Database
The IBM MarketScan Benefit Plan Design (BPD) Database consists of data for selected benefit plans represented in the MarketScan Research Databases from 1995 forward. A separate User Guide is provided to customers licensing the BPD Database. Benefit plan design information is available for the CCAE and Medicare Supplemental Databases.

Multi-State Medicaid Database
The MarketScan Multi-State Medicaid Database reflects the healthcare service use of individuals covered by Medicaid programs in numerous geographically dispersed states. The database contains the pooled healthcare experience of Medicaid enrollees, covered under fee-for-service and managed care plans. It includes records of inpatient services, inpatient admissions, outpatient services, and prescription drug claims, as well as information on long-term care. Data on eligibility 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 federal aid category (income based, disability, Temporary Assistance for Needy Families) and race.

MarketScan Lab Database
The IBM MarketScan Lab Database contains the pooled health care experience of more than one million covered lives, gleaned from sources that include both commercial and Medicare Supplemental coverage. It captures lab tests for a subset of the covered lives and mainly represents lab tests ordered in office-based practices. Linkage of lab results to claims supports analyses that are not feasible with claims alone, such as determining effectiveness of treatment, measuring severity of illness, identifying patients for whom treatment may be indicated, and verifying diagnoses recorded on claims.

Note: This User Guide is intended to cover the Multi-State Medicaid Database. The data you receive may contain some or all MarketScan data described herein.

Exhibit 1. Overview of the IBM MarketScan Research Databases

<table>
<thead>
<tr>
<th>Database</th>
<th>Content</th>
<th>Covered Lives</th>
<th>Tables</th>
</tr>
</thead>
</table>
| Commercial Claims and Encounters (CCAE) | Health care coverage eligibility and service use of individuals in plans or product lines with fee-for-service plans and fully capitated or partially capitated plans | Active employees and dependents, early (non-Medicare) retirees and dependents, COBRA continuees | Medical/Surgical  
  ○ Inpatient Admissions (I)  
  ○ Facility Header (F)  
  ○ Inpatient Services (S)  
  ○ Outpatient Services (O)  
  • Prescription Drug (D)  
  • Enrollment (A,T) |
| Medicare Supplemental and Coordination of Benefits (COB) (MDCR) | Health care coverage eligibility and service use of individuals in plans or product lines with fee-for-service plans and fully capitated or partially capitated plans | Medicare-eligible active and retired employees and their Medicare-eligible dependents from employer-sponsored supplemental plans | Medical/Surgical  
  ○ Inpatient Admissions (I)  
  ○ Facility Header (F)  
  ○ Inpatient Services (S)  
  ○ Outpatient Services (O)  
  • Prescription Drug (D)  
  • Enrollment (A,T) |
<table>
<thead>
<tr>
<th>Database</th>
<th>Content</th>
<th>Covered Lives</th>
<th>Tables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benefit Plan Design (BPD)</td>
<td>Plan characteristics derived from the medical claims submitted by each plan. Additional information specific to each plan is available in the BPD User Guide.</td>
<td>Not applicable</td>
<td>Links to CCAE and Medicare Supplemental and COB Databases for a subset of plans included in those databases</td>
</tr>
</tbody>
</table>
| Health and Productivity Management (HPM) | Absence, short-term disability, long-term disability, and worker’s compensation experiences for a subset of the covered lives represented in the CCAE Database | Active employees            | • Absenteeism (ABS)  
• Short-Term Disability (STD)  
• Long-Term Disability (LTD)  
• Worker’s Compensation (WC)  
• Eligibility (E)  
Linkable to the medical and prescription drug claims information appearing in the CCAE Database |
| Medicaid                     | Health care coverage eligibility and service use of individuals enrolled in state Medicaid programs for several states and/or Medicaid managed care programs | Medicaid recipients for several states | • Medical/Surgical  
  o Inpatient Admissions (I)  
  o Facility Header (F)  
  o Inpatient Services (S)  
  o Outpatient Services (O)  
  o Long-Term Care (L)  
• Prescription Drug (D)  
• Enrollment (A,T) |
| Lab                          | Health care service use and eligibility for individuals enrolled in commercial and Medicare Supplemental programs, along with lab test records and results | Individuals enrolled in commercial and Medicare Supplemental programs | • Medical/Surgical  
  o Inpatient Admissions (I)  
  o Facility Header (F)  
  o Inpatient Services (S)  
  o Outpatient Services (O)  
• Prescription Drug (D)  
• Enrollment (A,T)  
• Lab Test Results (R) |
| HRA                          | Self-reported biometric and health-related behavioral data               | Active employees            | Links to CCAE and Medicare Supplemental and COB Databases for a subset of enrollees. |

Abbreviation: COBRA, Consolidated Omnibus Budget Reconciliation Act.
OVERVIEW OF TABLES

Note: All tables and databases described below are available in the MarketScan Multi-State Medicaid Database.

MEDICAL/SURGICAL TABLES

The MarketScan Databases contain inpatient and outpatient medical/surgical data stored in four tables: Inpatient Admissions, Inpatient Services, Facility Header, and Outpatient Services.

Inpatient Admissions Table (I)
The Inpatient Admissions Table contains records that summarize information about a hospital admission. IBM Watson Health constructs this table after identifying all encounters or claims (service records) associated with an admission (e.g., hospital claims, physician claims, surgeon claims, and claims from independent laboratories). Facility and professional payment information then is summarized for all services. The summarized information is stored in an admission record in the Inpatient Admissions Table. Please refer to the section titled Financial Variables for definitions of key financial variables.

The admission record also includes data that can be identified only after all claims for an admission have been identified. These additional data include the principal procedure, principal diagnosis, major diagnostic category (MDC), and diagnosis-related group (DRG). Watson Health uses the Centers for Medicare & Medicaid Services (CMS) DRG Grouper to assign an MDC and DRG to the admission record.

In addition to the principal procedure and diagnosis codes, the admission record includes all diagnoses and procedures (up to 14 each) found on the service records that make up the admission. These additional codes (Diagnosis 2 through Diagnosis 15 and Procedure 2 through Procedure 15) are assigned chronologically on the basis of service dates and do not duplicate the principal code.

To be considered an admission, the grouping of these service records must meet certain criteria (e.g., a room and board claim must be present). If these criteria are not met, the records are stored in the Outpatient Services Table (O) and no admission record is created.

Facility Header Table (F)
The Facility Header Table contains complete header information from facility claims. A Facility Header Record identifier (FACHDID) exists on both the Facility Header Table and the Inpatient Services and Outpatient Services Tables to identify the individual service records that each header record comprises.

Facility inpatient service records are derived from the Uniform Billing (UB04) form. This form does not link financial information to specific procedures or diagnoses.

Inpatient Services Table (S)
The Inpatient Services Table contains the individual facility and professional encounters and services that the inpatient admission record comprises. A Cases and Services Link identifier (CASEID) exists on both the Inpatient Admissions and the Inpatient Services Tables to identify the individual service records that each admission record comprises.

Facility inpatient service records are derived from the UB04 form. This form does not link financial information to specific procedures or diagnoses. Physician services are derived from the CMS 1500 form.

Note: The Inpatient Services Table contains both facility and physician services associated with an inpatient admission. The Inpatient Admissions Table differs from UB04 discharge data in that Watson Health combines the facility charges with the physician services associated with an inpatient admission. UB04 revenue codes are retained in the MarketScan data when available; however, not all
data contributors provide the codes on adjudicated claims.

**Outpatient Services Table (O)**
The Outpatient Services Table contains encounters and claims for services that were rendered in a doctor’s office, hospital outpatient facility, emergency department, or other outpatient facility. A small percentage of claims in this table may represent inpatient services, because the claim was not incorporated into an inpatient admission (e.g., no room and board charge was found). These claims generally have an “inpatient” Place of Service (STDPLAC) code.

**Outpatient Pharmaceutical Claims Table (D)**
Outpatient pharmaceutical claims data are available for a large portion of the individuals represented in the medical/surgical and enrollment tables. The outpatient pharmaceutical data are linked by ENROLID to the medical/surgical data. Each record represents either a mail-order or card program prescription drug claim.

Note: Before you begin your analysis, carefully determine which data sources (e.g., medical/surgical, outpatient pharmaceutical, enrollment) will be necessary to support your analytic plan. If you require more than one of these data sources, it first may be necessary to use the various cohort flags to determine which data contributors or plans have the required data. These are found through the Cohort Drug indicator (DRUGCOVG), Mental Health and Substance Abuse Coverage (MHSACOVG), and/or Enrollee ID Derivation Flag (EIDFLAG) and Enrollment Flag (ENRFLAG) variables.

**LONG TERM CARE TABLE (L)**
The Long Term Care Table contains claims for services that were rendered in a long term care setting as well as room and board claims from long-term care facilities.

**ENROLLMENT TABLES (A, T)**
The Enrollment tables contain person-level enrollment records with demographic and plan information on users and nonusers of services contained in the MarketScan Medicaid Multi-State Database.

The Annual Enrollment Summary Table contains a single record per person per year. The annual summary contains monthly arrays of certain variables such as indicators of enrollment (yes/no), days enrolled, data type, and plan type in each month during the year. There also are variables indicating the number of months during the year with enrollment and the total annual enrollment days.

The Enrollment Detail Table contains one record per person per month of enrollment for an individual enrollee regardless of whether any demographic values have changed from the previous month.

If you need to track changes in variables such as DRUGCOVG indicator or Geographic Location of Employee (E GEOLOC), use the Enrollment Detail Table.

Beginning with the 2001 data, all data contributors submit person-level enrollment information. When using MarketScan Database releases prior to 2001, the ENRFLAG variable allows the user to select only claims supported by person-level enrollment. When ENRFLAG=1, it indicates that person-level enrollment information is available for that data contributor.

**Member Days (MEMDAYS)**
When obtaining an underlying population or covered life count, evaluate the Date Enrollment Start (DTSTART) and Date Enrollment End (DTEND) data before summing Member Days (MEMDAYS). If a time-based subset or study period is required, the DTSTART and DTEND may be outside the beginning and ending dates of the subset criteria. If so, adjust the DTSTART and DTEND to match the study period and recalculate the member days before calculating an enrollee count.

For example, a record may have DTSTART and DTEND of 1/1/2013 and 1/31/2013, respectively. The MEMDAYS variable on this record is 31 days. If the study period of data begins on 1/15/2013, the DTSTART should be reset to reflect the 1/15/2013 beginning date and MEMDAYS should be recalculated to 16 days (MEMDAYS=DTEND–DTSTART+1).
OVERVIEW OF ENCOUNTER RECORDS

Encounter records represent the service use and cost of individuals in partially and fully capitated plans and allow for the empirical investigation of health care under a variety of managed care arrangements.

Historically, not all fully or partially capitated health plans have maintained rigorous cost and utilization data collection systems. Many managed care services are prepaid in fixed sums for each member, which minimizes the need for administrative systems to collect financial encounter information at the time of service delivery. Therefore, unlike indemnity plans (which adjudicate claims for reimbursement), certain types of managed care plans do not process claims for the purpose of financial reporting. For these plans, service delivery information is disconnected from charge and payment information. Instead of generating a claim for reimbursement of prepaid capitated services, a managed care plan generates an encounter record.

An encounter record provides demographic information about the patient, provider characteristics, and diagnosis and procedure codes; however, in many instances it provides only limited financial information. This presents a certain challenge when using encounter records to analyze health care costs.

The challenge involves the correct measurement of reimbursement for capitated managed care plans. Many encounter records contain a Payment (PAY) amount of $1 or $0 for capitated services. The prepaid capitation amounts, whether in the form of per member per month fees or bulk capitation payments, were not contributed by the managed care plans represented in this database. However, managed care plans are beginning to enhance encounter records with fee-for-service-equivalent financial amounts. These amounts are intended to be approximate values for reasonable and customary charges or payments for medical services or procedures. See the Financial Variables section of this User Guide for other important information. The implementation of fee-for-service-equivalent financial amounts is in its early stages; as a result, financial variables are potentially understated. Financial measures derived from encounter records should be interpreted with caution, with the exception of Copayment (COPAY), Deductible (DEDUCT), and Coordination of Benefits and Other Savings (COB) amounts—all of which are recorded with reasonable accuracy.

In constructing the MarketScan Research Databases, encounter records are rigorously tested by overall plan-by-plan utilization rates to ensure that plans appearing to submit incomplete data are excluded.
FINANCIAL VARIABLES

IBM Watson Health receives paid claims from approximately 350 data sources. Financial variables are defined consistently across all data contributors. Exhibit 2 contains an example of a financial variable calculation.

The definitions in Exhibit 3 apply to all MarketScan Research Databases. The definitions apply to the capitated encounter data, even though some of the financial variables are set to zero (0) or one (1) because encounter records may not contain fee-for-service charge and payment equivalents.

To protect business-confidential discount arrangements between our data contributors and their providers, information on submitted charges and allowed amounts are never licensed simultaneously on the same MarketScan dataset.

Exhibit 2. Example of an IBM Watson Health Financial Variable Calculation

<table>
<thead>
<tr>
<th>Charge Type</th>
<th>Amount, $</th>
</tr>
</thead>
<tbody>
<tr>
<td>Submitted charges</td>
<td>1,200.00</td>
</tr>
<tr>
<td>Charges not covered</td>
<td>−100.00</td>
</tr>
<tr>
<td>Eligible charges</td>
<td>1,100.00</td>
</tr>
<tr>
<td>Price reductions</td>
<td>−100.00</td>
</tr>
</tbody>
</table>

*Charge types are not standard MarketScan variables.*

<table>
<thead>
<tr>
<th>Description</th>
<th>Data Element</th>
<th>Amount, $</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gross covered payments</td>
<td>Gross Covered Payments (PAY)</td>
<td>1,000.00</td>
</tr>
<tr>
<td>Remaining deductible</td>
<td>Deductible (DEDUCT)</td>
<td>−100.00</td>
</tr>
<tr>
<td>Coinsurance at 20%</td>
<td>Coinsurance (COINS)</td>
<td>−180.00</td>
</tr>
<tr>
<td>Penalty for no precertification</td>
<td>Coordination of Benefits and Other Savings (COB)</td>
<td>−270.00</td>
</tr>
<tr>
<td>Net payments</td>
<td>Net Payments (NETPAY)</td>
<td>450.00</td>
</tr>
</tbody>
</table>
**MEDICAL/SURGICAL FINANCIAL VARIABLES**

The following abbreviations indicate the tables on which the variable resides:

- **I** – Inpatient Admissions
- **F** – Facility Header
- **S** – Inpatient Services
- **O** – Outpatient Services
- **D** – Outpatient Pharmaceutical Claims
- **L** – Long-Term Care

**Exhibit 3. Definitions of Medical/Surgical Financial Variables**

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
<th>MarketScan Variable</th>
<th>Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Payment</td>
<td>Total gross payment to all providers associated with the admission</td>
<td>Payments, Total Case (TOTPAY)</td>
<td>I</td>
</tr>
<tr>
<td>Payment</td>
<td>Total gross payment to a provider for a specific service; that is, the amount eligible for payment after applying pricing guidelines such as fee schedules and discounts and before applying deductibles, copayments, and coordination of benefits</td>
<td>Payment (PAY)</td>
<td>S,O,D,L</td>
</tr>
<tr>
<td>Deductible</td>
<td>Amount of gross covered payments applied toward the deductible</td>
<td>Deductible, Total Case (TOTDED)</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deductible (DEDUCT)</td>
<td>F,S,O,D,L</td>
</tr>
<tr>
<td>Coinsurance/Copayment</td>
<td>Amount of coinsurance applied toward the stop loss and/or amount of copayment</td>
<td>Copayment, Total Case (TOTCOPAY)</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Copayment (COPAY)</td>
<td>F,S,O,D,L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Coinsurance, Total Case (TOTCOINS)</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Coinsurance (COINS)</td>
<td>F,S,O,D,L</td>
</tr>
<tr>
<td>Net Payment</td>
<td>Payment received by the provider excluding patient out-of-pocket and coordination of benefits (i.e., employer or plan liability)</td>
<td>Payments, Net (NETPAY)</td>
<td>F,S,O,D,L</td>
</tr>
<tr>
<td>Total Net Payment</td>
<td>Total net payment to all providers associated with the admission (i.e., sum of service-level net) payments</td>
<td>Payments, Net Case (TOTNET)</td>
<td>I</td>
</tr>
<tr>
<td>Hospital Payments</td>
<td>Total gross payments to the hospital for an admission</td>
<td>Payments, Hospital (HOSPPAY)</td>
<td>I</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
<td>MarketScan Variable</td>
<td>Table</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Physician Payment</td>
<td>Total gross payments to the principal physician (i.e., the professional who charges the most during the admission)(^b)</td>
<td>Payments, Physician (PHYSPAY)</td>
<td>I</td>
</tr>
<tr>
<td>Hospital Net Payment</td>
<td>Payment received by the hospital for an admission excluding patient out-of-pocket and coordination of benefits (i.e., employer or plan liability)</td>
<td>Net Payment, Hospital (HOSPNET)</td>
<td>I</td>
</tr>
<tr>
<td>Physician Net Payment</td>
<td>Payment received by the principal physician (i.e., the professional who charges the most during the admission), excluding patient out-of-pocket and coordination of benefits (i.e., employer or plan liability)</td>
<td>Net Payment Physician, (PHYSNET)</td>
<td>I</td>
</tr>
<tr>
<td>Third-Party Payment</td>
<td>Payment received by the provider from a source other than the patient or the submitting plan</td>
<td>Coordination of Benefits and Other Savings, Total Case (TOTCOB)</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td></td>
<td>COB and Other Savings (COB)</td>
<td>F,S,O,D,L</td>
</tr>
</tbody>
</table>

\(^a\) These variables are formatted in dollars and cents.

\(^b\) Payments to physicians other than the principal physician are included in Payments Total Case (TOTPAY).
The Outpatient Pharmaceutical Claims Table contains the Payment (PAY), Copayment (COPAY), Coinsurance (COINS), Deductible (DEDUCT), and Coordination of Benefits and Other Savings (COB) variables, as previously described. Financial variables specific to prescription drug claims are provided in Exhibit 4.

Exhibit 4. Definitions of Outpatient Pharmaceutical Financial Variables in Table D

<table>
<thead>
<tr>
<th>Term</th>
<th>Definitiona</th>
<th>MarketScan Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Wholesale Priceb</td>
<td>The average wholesale price charged by wholesalers for the specific drug</td>
<td>Average Wholesale Price (AWP)</td>
</tr>
<tr>
<td>Administrative Dispensing Fee</td>
<td>Administrative fee charged by the pharmacy for dispensing the prescription</td>
<td>Dispensing Fee (DISPFEE)</td>
</tr>
<tr>
<td>Ingredient Cost</td>
<td>The cost or charge associated with the pharmaceutical productc</td>
<td>Ingredient Cost (INGCOST)</td>
</tr>
<tr>
<td>Sales Tax</td>
<td>The amount of sales tax applied to the cost of the prescriptiond</td>
<td>Sales Tax (SALETAX)</td>
</tr>
</tbody>
</table>

a These variables are formatted in dollars and cents.

b The IBM RED BOOK™ Systems Licensed Content may be used only as a referential look-up tool and not for an automated claims processing system; use is for RED BOOK System Licensed Content only. The prices contained in the RED BOOK are based on data reported by manufacturers. IBM Micromedex® has not performed an independent analysis of the actual prices paid by wholesalers and providers in the marketplace. Thus, actual prices may vary from the prices contained in this database, and all prices are subject to change without notice. Further, IBM Micromedex does not warrant the accuracy of the database contents or the pricing information. Please refer to the Average Wholesale Price Policy in the RED BOOK product for more information.

c The Ingredient Cost plus the Dispensing Fee and Sales Tax, if applicable, usually represents the entire cost of a prescription.

d Calculation of the sales tax, if applicable, usually is based on the Ingredient Cost plus the Dispensing Fee.
ENCOUNTER RECORD FINANCIAL VARIABLES

Financial information is captured in a variety of ways for encounter claims. A capitated claim may have financial variables with amounts of zero because there is no associated paid claim. At other times, the copayment amount may be the only financial information on the claim. The CAP flag will indicate whether the patient was covered under a capitated arrangement at the time of service. If CAP=1, then the record could be an encounter. If a capitated claim does not include financial information, the financial variables are set to “0” or “1.”

ADJUSTMENT RECORDS

Some claims have negative amounts in payment or other financial variables. These are adjustment records entered by claims processors to correct a payment error or any type of coding error.

Using strict criteria, adjustment records in MarketScan data have been resolved using one of two types of adjustment methods. Resolution of adjustments combines the financial variables on the original record with the financial variables on the adjustment. No information is lost when resolving adjustment records. The sum of the financial variables remains the same. However, instead of reading across multiple records to understand the services rendered, resolution of adjustments creates a single service-level record. Adjustment records are resolved on both the Outpatient Services Table and the Outpatient Pharmaceutical Claims Table. Adjustment records are not resolved on the Inpatient Services Table.

There are two methods that claims processors typically use for entering adjustment records: the Adjustment Method and the Void and Replace Method.

*The Adjustment Method* allows the entry of a new claim that exactly duplicates all the correct variables on the erroneous claim, including the date of service. In the case of financial information being incorrect, an adjusted dollar amount is entered in the appropriate financial variable(s) (e.g., PAY), and all other financial variables are $0. In the case of a nonfinancial variable being incorrect, the data in the appropriate variable (e.g., DX1) are corrected, and all financial variables are $0 on the adjustment record. This way, the sum of the financial variables of the erroneous claim and the adjustment claim equals the correct financial amounts. Under this method, therefore, two records represent a single transaction.

To resolve the adjustment, the MarketScan Database build process matches the adjustment with the original record, with the requirement that certain nonfinancial variables are exactly the same on both records. The financial information on the two records is summed, creating one resulting record. Exhibit 5 presents an example of the adjustment method.

Exhibit 5. Example of the Adjustment Method

<table>
<thead>
<tr>
<th>Record</th>
<th>ENROLID</th>
<th>SVCDATE</th>
<th>DX1</th>
<th>PAY</th>
<th>NETPAY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>9876501</td>
<td>19970630</td>
<td>12345</td>
<td>100</td>
<td>70</td>
</tr>
<tr>
<td>Adjustment</td>
<td>9876501</td>
<td>19970630</td>
<td>12345</td>
<td>−20</td>
<td>0</td>
</tr>
<tr>
<td>Resulting</td>
<td>9876501</td>
<td>19970630</td>
<td>12345</td>
<td>80</td>
<td>70</td>
</tr>
</tbody>
</table>

*The Void and Replace Method* allows entry of a new claim that exactly duplicates all variables of the erroneous claim except that the financial variables are entered as negatives. In this way, the original erroneous claim is fully voided, and the claim is re-entered with complete correct data in each variable.

Under this method, three records are present to represent a single transaction.

To resolve the adjustment, the MarketScan Database build process matches the void record with the original record, with the requirement that certain nonfinancial information is exactly the same on both records and the financial information on the
void record is the exact negative of the original record. The void and original record are dropped from the database because the combined record is a record in which all financial information is zero.

Only the replacement record remains. Exhibit 6 presents an example of the Void and Replace Method.

### Exhibit 6. Example of the Void and Replace Method

<table>
<thead>
<tr>
<th>Record Type</th>
<th>ENROLID</th>
<th>SVCDATE</th>
<th>DX1</th>
<th>PAY</th>
<th>NETPAY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>9876501</td>
<td>19970630</td>
<td>12345</td>
<td>100</td>
<td>70</td>
</tr>
<tr>
<td>Void</td>
<td>9876501</td>
<td>19970630</td>
<td>12345</td>
<td>-100</td>
<td>-70</td>
</tr>
<tr>
<td>Replacement</td>
<td>9876501</td>
<td>19970630</td>
<td>12345</td>
<td>80</td>
<td>70</td>
</tr>
<tr>
<td>Resulting</td>
<td>9876501</td>
<td>19970630</td>
<td>12345</td>
<td>80</td>
<td>70</td>
</tr>
</tbody>
</table>

### UNRESOLVED ADJUSTMENTS

Because strict matching criteria are required to resolve adjustments, some adjustment records remain unresolved; these account for less than 1 percent of the records in the MarketScan Outpatient Services Table. These records generally contain changes to a variable that normally would be used to match the original and adjustment records. For example, if the original Provider ID (PROV_ID) was incorrect and the adjustment record adjusted for that ID, the two records would not match because PROV_ID is a key variable. Both records would remain. When performing person-level analysis, or higher levels of analysis such as geographic region, all claims must be included.
PERSON-LEVEL IDENTIFIERS

ENROLLEE IDENTIFIERS

One of the major strengths of the MarketScan Research Databases is their ability to track patients and families longitudinally. The MarketScan Databases maintain person-level identifiers, consisting of family and member identifiers, across all years of data and across all tables, including medical/surgical and outpatient pharmaceutical claims.

The enrollee identifier (ENROLID) is the unique identifier assigned during MarketScan data build and cannot be linked to recipient ID, Social Security number, or any other external identifier. Enrollee identifiers are derived from all data contributors, not only those submitting person-level enrollment data. The methodology used to assign ENROLID differs depending on the level of information available from a particular data contributor.
CLINICAL VARIABLES

**Diagnosis codes** in MarketScan data use the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) classification system for service dates on or before September 30, 2015. For service dates starting October 1, 2015, the International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) classification system is used. A Diagnosis Version (DXVER) field is included in the data to indicate which coding system is in use. Note that it is possible for one string to be valid in both systems.

ICD-9-CM diagnosis codes are three to five digits. The first character can be alphanumeric (0–9, E or V); characters two through five are numeric or blank. There are approximately 15,800 valid ICD-9-CM codes. In MarketScan data, the decimal point is implied between the third and fourth digit of the code. The data are left justified. Examples are provided in Exhibit 7a.

**Exhibit 7a. Examples of ICD-9-CM Diagnosis Codes**

<table>
<thead>
<tr>
<th>ICD-9-CM</th>
<th>MarketScan Data Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>390</td>
<td>390 (followed by 2 spaces)</td>
</tr>
<tr>
<td>012.1</td>
<td>0121 (followed by 1 space)</td>
</tr>
<tr>
<td>223.89</td>
<td>22389</td>
</tr>
</tbody>
</table>

ICD-10-CM diagnosis codes are three to seven digits in length. The first character can be alphanumeric, the second character is numeric, the third character is alphanumeric, and the fourth through seventh characters are alphanumeric or blank. There are approximately 70,000 valid ICD-10-CM codes. In MarketScan data, the decimal point is implied between the third and fourth digit of the code. The data are left justified. Examples are provided in Exhibit 7b.

**Exhibit 7b. Examples of ICD-10-CM Diagnosis Codes**

<table>
<thead>
<tr>
<th>ICD-10-CM</th>
<th>MarketScan Data Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>E02</td>
<td>E02 (followed by 4 spaces)</td>
</tr>
<tr>
<td>M86.9</td>
<td>M869 (followed by 3 spaces)</td>
</tr>
<tr>
<td>C72.20</td>
<td>C7220 (followed by 2 spaces)</td>
</tr>
<tr>
<td>B08.010</td>
<td>B08010 (followed by 1 space)</td>
</tr>
<tr>
<td>W00.9XXA</td>
<td>W009XXA</td>
</tr>
</tbody>
</table>

Up to four diagnosis codes (DX1, DX2, DX3, DX4) are recorded on every Inpatient Service record. The principal diagnosis on the Inpatient Admissions Table generally is identified as the discharge diagnosis on a hospital claim. Up to 14 secondary diagnosis codes (DX2 through DX15) from individual Inpatient Service records are included on the corresponding Inpatient Admission record. Up to four diagnosis codes (DX1, DX2, DX3, DX4) are recorded on each Outpatient Service record. Up to nine diagnosis codes (DX1 through DX9) are recorded on each Facility Header record.

**Procedure codes** in MarketScan data are three to seven digits, depending on the classification system used. The Current Procedural Terminology, 4th Edition (CPT-4) coding system is most prevalent. CPT-4 procedure codes appear on physician claims and many outpatient facility claims. CPT-4 codes are five-digit numeric codes.

ICD-9-CM procedure codes or International Classification of Diseases, Tenth Revision, Procedure Coding System (ICD-10-PCS) procedure codes are found on facility claims. These codes are three to four digits and are all numeric. There is an

---

1 CPT copyright 2016 American Medical Association (AMA). All rights reserved. Applicable Federal Acquisition Regulation (FAR) and Defense Federal Acquisition Regulation Supplement (DFARS) restrictions apply to government use.

2 Fee schedules, relative value units, conversion factors, and related components are not assigned by the AMA and are not part of CPT; the AMA is not recommending their use. The AMA does not directly or indirectly practice medicine or dispense medical services. The AMA assumes no liability for data contained or not contained herein.
implied decimal point between the second and third digits for ICD-9-CM procedure codes; there is no decimal point in ICD-10-PCS procedure codes. Examples are provided in Exhibit 8.

Exhibit 8. Examples of ICD-9-CM and ICD-10-PCS Procedure Codes

<table>
<thead>
<tr>
<th>ICD-9-CM, ICD-10-PCS</th>
<th>MarketScan Data Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.9</td>
<td>139 (followed by 4 spaces)</td>
</tr>
<tr>
<td>13.19</td>
<td>1319 (followed by 3 spaces)</td>
</tr>
<tr>
<td>001U3J7</td>
<td>001U3J7</td>
</tr>
</tbody>
</table>

Effective with the 2000 data year, the MarketScan Databases contain CPT-4 procedure code modifiers for some data contributors (see related references in footnotes on previous page).

The CMS Healthcare Common Procedural Coding System (HCPCS) procedure codes are found in MarketScan data less often than CPT and ICD procedure codes. These codes are five digits. The first character is alpha; all other characters are numeric. HCPCS codes beginning with “J” are included in the MarketScan Databases and represent injectable drugs.

One procedure code (PROC1) is stored on each Inpatient Services record. From the individual Inpatient Services constituting one Inpatient Admissions record, one procedure code is identified and assigned as the principal procedure (PPROC). Up to 14 secondary procedure codes (PROC2 through PROC15) from individual Inpatient Service records are included on the corresponding Inpatient Admissions record. One procedure code (PROC1) is included on each Outpatient Service record. Up to six procedure codes (PROC1 through PROC6) are included on each Facility Header record. Most procedure codes on the Facility Header Table use the ICD-9-CM or ICD-10-PCS systems.

The variable Procedure Code Type (PROCTYP) identifies the type of procedure code (e.g., HCPCS, CPT-4). Use this variable in conjunction with the Procedure Code 1 (PROC1) variables on the Inpatient Services and Outpatient Services records to designate the coding system of interest.

The quality of diagnosis and procedure coding varies among the approximately 350 payers or administrators represented in the MarketScan Databases. Every effort is made to select the data contributors with the best coding. The diagnosis and procedure codes are validated and edited, if necessary. The IBM MarketScan Multi-Medicaid Database contains the pooled health care experience of approximately seven million Medicaid enrollees from multiple states.

Each contributor database is constructed by collecting raw data from the appropriate payer(s). These raw data are service-level adjudicated paid claims and capitated encounters containing both inpatient and outpatient services. Financial, clinical, and demographic variables are standardized to common definitions, and variables specific to employers also are added. Clinical detail is added to the Outpatient Pharmaceutical Claims Table. Examples of such detail include therapeutic class, therapeutic group, manufacturer’s average wholesale price, and a generic product identifier.

IBM Watson Health then applies an admission construction methodology to assemble the inpatient paid services into one record per inpatient admission. During the admission creation process, variables such as Primary Diagnosis (PDX) are created and included on both the inpatient admission record and the inpatient service record.

**DATA QUALITY**

Edits on the reasonableness of data check the distribution of categorical fields to ensure that they
are reasonable against norms. Validity checks are conducted for selected fields, including diagnosis codes, procedure codes, date(s) of service, sex, and age, to compare recorded values with lists of possible valid values for those fields. Improper coding is flagged to recommend data quality improvement actions to the carrier or data processor.

The MarketScan Databases are created by combining the standard variables of the individual databases (data contributors) and by creating links between years of data and across all data types. The MarketScan Databases are created as a snapshot in time and are based on a calendar-year incurred period. The MarketScan data flow is depicted in Exhibit 9.

Claims lag periods (the amount of time between the date of service on the claim and the date payment is made) vary considerably across the insurance carriers in the MarketScan Databases. Because of this, the data are collected when close to 100 percent of claims have been paid, which takes about 6 months after year end.

Additional enhancements to the data during the MarketScan Database creation process include the following:

- Comparing and validating diagnosis and procedure codes to codes that were in effect at that time
- Adding the Metropolitan Statistical Area (MSA) of the primary beneficiary to claims
- Integrating benefit plan characteristics, enrollment, outpatient pharmaceutical claims, and medical/surgical data
- Adding MDCs and DRGs to claims
- Creating a common synthetic patient identifier, which enables a patient to be tracked over years across medical/surgical, outpatient, pharmaceutical, enrollment, and benefit plan files and across databases (e.g., CCAE and MDCR Databases) while ensuring patient confidentiality
- Identifying the type of plan for the patient, such as preferred provider organization (PPO), point-of-service (POS) plan, or comprehensive plan
- Verifying that both the experience and the denominator populations exist for all subsets of the data
- Standardizing place, service type, and provider type values and industry classifications.

Note: Data are not edited for concordance between diagnosis or procedure codes and demographic variables such as sex.
Exhibit 9. MarketScan Data Flow Chart

**MarketScan® Data Flow**

**Abbreviation:** DX/Proc, diagnosis/procedure.
PLANT TYPE DEFINITIONS

The plan types in the MarketScan Databases are based on the definitions provided below. The summary grid identifies the basic differences between plan types.

Exhibit 10. Type of Plan (PLANTYP)

<table>
<thead>
<tr>
<th>Definition Number and Plan Type</th>
<th>Patient Incentive to Use Certain Providers?</th>
<th>PCP Assigned?</th>
<th>Referrals From PCP to Specialists Required?</th>
<th>Out-of-Network Services Covered?</th>
<th>Partially or Fully Capitated?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. B/MM</td>
<td>No</td>
<td>No</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>2. COMP</td>
<td>No</td>
<td>No</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>3. EPO</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>4. HMO</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>5. Non-Cap POS</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>6. PPO</td>
<td>Yes</td>
<td>No</td>
<td>n/a</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>7. Cap or Part Cap POS</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>8. CDHP</td>
<td>Varies</td>
<td>No</td>
<td>n/a</td>
<td>Varies</td>
<td>No</td>
</tr>
<tr>
<td>9. HDHP</td>
<td>Varies</td>
<td>No</td>
<td>n/a</td>
<td>Varies</td>
<td>No</td>
</tr>
</tbody>
</table>

Abbreviations: n/a, not applicable; PCP, primary care physician. Plan type abbreviations are defined below.

PLAN TYPE

1. Basic/Major Medical Plan

There is no incentive for the patient to use a specific list of providers. Coverage is handled in two phases: a basic policy covers the first set of charges—usually a hospital admission—with no out-of-pocket charge. After the basic policy will no longer pay, a major medical policy assumes coverage, usually with a deductible and coinsurance.

2. Comprehensive Plan

There is no incentive for the patient to use a specific list of providers. Coverage is handled by only one policy with a deductible and coinsurance.

3. Exclusive Provider Organization Plan

Patients must choose from an approved list of providers for all nonemergency care. Each patient chooses a primary care physician (PCP) to manage all care. Referral from the PCP is required for treatment by specialists. The plan pays for all services on a capitated basis.

4. Health Maintenance Organization Plan

Patients must choose from an approved list of providers for all nonemergency care. Each patient chooses a PCP to manage all care. Referral from the PCP is required for treatment by specialists. The plan pays for all services on a capitated basis.
5. Noncapitated Point-of-Service Plan
Patients are offered financial incentives through a lower copay or deductible to use an approved list of providers. Each patient chooses a PCP to manage all care. Referral from the PCP is required for treatment by specialists. No services are capitated, and patients may seek treatment outside the network, usually involving a severe financial penalty.

6. Preferred Provider Organization Plan
Patients have financial incentives, often through a lower copay or deductible, to use an approved list of providers. No PCP is required, nor are referrals necessary. No services are capitated. Patients may seek treatment outside the network but usually at a higher cost to the patient. The financial incentives may be offered only through discounted rates within the network.

7. Capitated (Cap) or Partially Capitated (Part Cap) Point-of-Service Plan
Patients are offered financial incentives through a lower copay or deductible to use an approved list of providers. Each patient chooses a PCP to manage all care. Referral from the PCP is required for treatment by specialists. All or some services are paid on a capitated basis. Patients may seek treatment outside the network, usually with a financial penalty.

Note: For Medicaid, this plan type also includes Primary Care Case Management.

8. Consumer-Directed Health Plan
A consumer-directed health plan (CDHP) is a PPO plan coupled with a health reimbursement arrangement (HRA). The PPO plan typically has a relatively high deductible but may carve drugs in or out of the HRA and plan deductible. The HRA is a notional account that is 100 percent paid from employer funds; an HRA is not prefunded with employer monies.

9. High-Deductible Health Plan
A high-deductible health plan (HDHP) is a statutory HDHP (as defined in the Medicare Modernization Act of 2003) that is coupled with a health savings account (HSA). An employee is vesting 100 percent in HSA funds, and either the employer or employee can contribute to the HSA. The HSA is a tax-advantaged, portable savings account owned by the employee. HDHP plan design features such as deductibles and contribution limits are indexed each year by the Treasury Department. An HDHP must conform to the statutory plan design requirements in order to use an HSA to defray HDHP costs.
GLOSSARY OF ACRONYMS, ABBREVIATIONS, AND TERMS

**Acute care.** (1) Services within a hospital setting intended to provide patients with medical and surgical care over a relatively short period of time. (2) A hospital that provides short-term medical and surgical care.

**Adjudication.** The process of claims review by the carrier to determine whether the claims should be paid and, if so, how much money should be paid for each claim.

**Adjustment records.** Claims in some databases that represent financial adjustments to original claims. The dollar amounts of these adjustments may be negative, or the record may include an adjustment indicator that shows whether the adjustment is positive or negative. There also are specific terms that refer to adjustments as we receive them from carriers. A bulk adjustment is a single quarterly or annual adjustment for a hospital discount (not typically loaded on the database). A void adjustment is a record that simply cancels an earlier claim record. A replacement claim record usually follows it. A void and replace adjustment is a single record that stores both the cancellations of the earlier claim and the new claim. An adjustment to net pay just shows the difference between the original net pay amount and what the carrier actually paid.

**Administrator.** Person or firm that pays claims under an Administrative Services Only (ASO) contract—also known as a third-party administrator.

**Admission.** An acute inpatient hospital stay covered by the patient’s benefit plan. To the extent that such care is covered, admissions may include hospital stays, psychiatric stays, psychiatric night care, and stays for alcoholism, substance abuse, and rehabilitative care. An admission also may be called a case or a stay.

**Admission date.** The date a patient begins a stay in a hospital or other overnight health care facility.

**Ambulatory care.** Medical services provided on an outpatient (nonhospitalized) basis. Services may include diagnosis, treatment, surgery, and rehabilitation.

**Ambulatory surgery.** Surgery for which there is no overnight stay in a hospital. The patient comes into and out of the hospital on the same day.

**Annualization.** A statistical technique for estimating a yearly rate using data collected over a shorter time period (e.g., a quarter or month) or over a longer time period (e.g., 30 months).

**Average length of stay (ALOS).** The average number of days per hospital admission for a group of admissions. Analysts typically examine the ALOS for a single MDC or DRG at a given beneficiary’s location or other variable and compare it with a norm, another location, or other measure. See length of stay.

**Benefit.** Conventionally defined as the amount payable for a loss under a specific type of insurance coverage (indemnity benefits) or as the guarantee that certain services will be paid.

**Business coalitions.** Groups of employers, which may or may not include health plans, that seek to control health care costs and ensure quality by aggressively regulating prices, assuming administrative tasks related to health care, and/or asking health plans to develop and provide data on measures of quality and outcomes.

**Capitation.** (1) A predetermined amount prepaid to a provider for a specific group of services that are defined in the contract, usually in a health maintenance organization (HMO) arrangement. The provider is paid on the basis of the number of members who have selected him or her as their primary care physician (PCP). (2) A fixed, predetermined amount paid to a provider for each member who has elected to seek care from that provider. Total payment to the provider (sum of per person enrolled payment amount) is based on the actual number or nature of services provided to members. This is the characteristic payment method for primary care in HMOs.
Carrier. The party to the group contract that agrees to underwrite and provide certain types of coverage and services. Examples are commercial insurers (e.g., Aetna®, Metropolitan Insurance Services, Prudential) and Blue Cross Blue Shield.

Carve-out. A program that is separate from the primary group health plan and designed to provide a specialized type of care, such as mental health services. Carve-out also may describe a method of integrating Medicare with an employer’s retiree health plan (making the employer plan excess or secondary), which tends to produce the lowest employer cost.

Case level. A variable that is found in the Inpatient Admissions Table. Case-level variables may be demographic variables that are the same for the entire case (e.g., patient age, sex, enrollee ID number), clinical variables that refer to the case as a whole (e.g., MDC, DRG), or financial variables that summarize all services for a case (e.g., total payments). See service level for comparison.

Centers for Medicare & Medicaid Services (CMS). (1) A division within the U.S. Department of Health and Human Services that oversees all regulatory and financing activities for Medicare and Medicaid. (2) The portion of the federal government responsible for payment of Medicare. Prior to June 2001, CMS was named the Health Care Financing Administration.

Charges. The amount patients or third-party payers are billed for care.

Claims data. Information that comes from provider claims to third-party payers. Claims data usually include personal patient-identification information, the services performed, and the amount paid by the patient. Claim forms generally are used by enrollees of standard indemnity plans (i.e., fee-for-service plans).

Claims lag. (1) This lag generally refers to the period between the date a health care service is incurred and the date the claim for that service is submitted to the administrator for payment. (2) The IBM Watson Health definition is the period between the service date and the paid date on a claim. See runoff.

Coding. The handling process for the carrier’s claims data. A coding problem indicates that the carrier entered inaccurate or imprecise data into the claims record, failed to fill in one or more data variables, or failed to include one or more variables in the record extract.

Coinsurance. (1) The percentage of a covered medical expense that a health plan or beneficiary must pay after a deductible is met. (2) A policy provision by which both the insured and the insurer share hospital and medical expenses in a specified ratio (commonly 20 percent to 80 percent), after the deductible is met. Coinsurance amounts are stored in the IBM Watson Health variable COINS.

Completion factors. (1) Factors that allow a quantitative measure of data completeness. These factors range in value between 0 (no data) and 100 (a full month of data) for services in any month. Completion factors are used to derive the number of months of data and an annualization factor for rate calculations. They also are used to derive weighted population averages. (2) A percentage that estimates how many of the cases that occurred in a given month are online in a client database. Completion factors of less than 100 percent are due to runoff or runup. The percentage of data missing for each month is used to annualize the cost and use rates for that month on clinical reports.

Comprehensive Omnibus Budget Reconciliation Act (COBRA). (1) A congressional act passed in 1985 that requires continuation of benefits to plan participants who previously would have been ineligible because of a qualifying event. (2) A program that gives employees who leave an employer the option of continuing their health coverage with that employer for a period of time. The employee pays the premium.

Coordination of benefits (COB). (1) After one insurance carrier has paid a claim, the second carrier pays an amount that covers the patient up to the benefit level of the second policy only. (2) COB coverage between carriers so that the insured does not receive double payment for services when a subscriber has coverage from two or more sources. An example is when a Medicaid beneficiary is also covered by Medicare. COB policies also establish primary and secondary payment responsibilities. (For older databases in the IBM Watson Health system, the COB variable may represent dollars.
saved for reasons other than COB, such as penalties for noncompliance.)

Copay or copayments. (1) Copayments are generally a preset amount per covered visit or service (e.g., $10) paid by the patient. (2) A fixed payment, paid by the patient, for a given service or procedure. This payment customarily is made at the time of service. Copayment amounts are stored in the MarketScan variable COPAY.

Cost sharing. Arrangements whereby consumers pay a portion of the cost of the health services, sharing costs with employers or, for Medicare and Medicaid beneficiaries, with the government. Deductibles, copayments, coinsurance, and payroll deductions (premium contributions) are forms of cost sharing.

Cost shifting. Occurs when a provider inflates charges for a given procedure or patient to cover losses associated with charges (payments received) for other patients or procedures.

CPT or CPT-4 codes. Physicians’ Current Procedural Terminology codes. (1) Physicians’ most commonly used coding scheme (five-digit codes) used to identify the medical or surgical procedure that occurred for a patient; most frequently used for billing by professionals. (It often is referred to as CPT-4, with 4 representing the fourth edition). (2) A system developed by the American Medical Association to classify procedures and services rendered by physicians. Physicians use the CMS 1500 form to describe services rendered to a patient and to request payment for those services. See ICD-9-CM, ICD-10-CM/PCS, HCPCS.

Deductible. The portion of a subscriber’s health care expenses that must be paid out of pocket before any insurance coverage applies—commonly $100 to $300. It is not allowed in federally qualified HMOs. The deductible usually must be met again each benefit year before the insurer will begin paying for benefits. The deductible amount is stored in the IBM MarketScan variable DEDUCT.

Dependent. An insured individual’s spouse or (in many policies) domestic partner and unmarried children who meet certain eligibility requirements and who are not otherwise insured under the same group policy. The precise definition of a dependent varies by insurer or employer.

Diagnosis (Dx). The determination of the nature of a disease based on the medical symptoms of a patient; a concise technical classification of a health situation. The diagnosis helps determine necessary procedures.

Discount. Arrangement whereby a payer has negotiated a reduced payment with a provider in return for a patient incentive.

Eligible. A contract holder and his or her spouse and dependents who are enrolled in a benefit plan.

Encounter. (1) A unit of measure denoting one patient-provider contact or appointment. Multiple services may be delivered during one encounter. Encounters can take place on an inpatient or outpatient basis. (2) A patient visit to a capitated provider; no fee-for-service payment.

Encounter record. A record of a patient encounter reflecting who visited a given provider and which services were provided. The form used to capture encounter data applies to non-fee-for-service arrangements (capitated).

Enrollees. Employees, beneficiaries, contract holders, spouses, and dependents who are enrolled in a benefit plan (also known as covered lives).

Exclusions. Services or procedures that are not covered according to the plan provisions.

Exclusive provider organization (EPO). A preferred provider organization (PPO) in which patients are required to use the PPO network providers.

Fee for service (FFS). A method of payment based on reimbursing providers for each unit of service or treatment provided.

Fee-for-service equivalent (FFSE). An amount specified on claims records representing what would have been charged for a service if the service had not been covered by a capitation arrangement.

Gatekeeper. (1) The PCP responsible for managing medical treatment rendered to an enrollee of a health plan. (2) A designated health care practitioner who provides primary care services and coordinates specialist and other care for health plan members. Members typically are charged extra
costs for care that is not provided or coordinated by the gatekeeper.

**Grouper.** Software that assigns claims to a common clinical grouping. In the MarketScan Databases, groupers are used to assign a DRG and MDC to each inpatient admission. The assignment is based on diagnosis and procedure coding received from the carrier (provided the diagnosis and procedure coding from the carrier is adequate).

**Healthcare Common Procedure Coding System (HCPCS).** (1) A procedure coding system that includes all CPT-4 codes plus supplemental codes not included in CPT-4 (e.g., ambulance, chiropractic services). (2) One of several schemes used to classify health care activity. HCPCS was based on CPT-4 coding and expanded to include nonphysician provider procedures. The acronym is pronounced “hick-picks.” See CPT-4, ICD-9-CM, ICD-10-CM/PCS.

**Health maintenance organization (HMO).** (1) An entity that accepts responsibility and financial risk for providing specified health care services to a defined population during a defined period at a fixed price. There generally is no coverage for non-emergency-department care panels of practitioners and providers. (2) The Health Maintenance Act of 1973 (PL93-222) defines an HMO as a legal entity or organized system of health care that provides an agreed-upon set of comprehensive health services to a voluntarily enrolled population in exchange for a predetermined, fixed, and periodic payment. See *open-ended HMO*.

**Hospital payments.** Facility payments only.

**Incurred but not reported (IBNR).** Claims for services that have been incurred but not yet paid by the carrier. See *claims lag*.

**International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM).** A nationally uniform system for coding clinical conditions (diagnoses) that was used prior to October 1, 2015, by nearly all providers and claims payers. It also includes procedure coding used by hospitals. ICD-9-CM includes both diagnostic and procedure coding required by the Grouper to assign DRGs and MDCs. It is also known as I9. See *CPT-4, HCPCS, ICD-10-CM/PCS*.

**International Classification of Diseases, Tenth Revision, Clinical Modification/Procedure Coding System (ICD-10-CM/PCS).** A nationally uniform system for coding clinical conditions (diagnoses), used effective October 1, 2015, by nearly all providers and claims payers. It also includes procedure coding used by hospitals. ICD-10-CM/PCS includes both diagnostic and procedure coding required by the Grouper to assign DRGs and MDCs. It is also known as I10. See *CPT-4, HCPCS, ICD-9-CM*.

**Incurred date.** The date on which the activity or service took place. See *paid date, claims lag, IBNR*.

**Indemnity (traditional) insurance.** (1) A health care insurance plan designed to reimburse patients for losses due to health care costs; typically used to characterize fee-for-service payment plans. (2) The most common form of health insurance coverage in recent decades. The indemnity insurer usually administers claims and does not provide health care services. A typical coverage arrangement is 80 percent of a claim covered by the insurer and 20 percent covered by the patient or enrollee (also referred to as coinsurance). Indemnity plans typically also require that the covered person meet an annual deductible (e.g., $200) before the insurer will begin to pay a percentage of claims incurred.

**Individual practice association (IPA).** A type of HMO. A group of physicians who practice independently but also provide services for an HMO under a contract agreement. An IPA physician also can and does provide “traditional” fee-for-service health care to patients not covered by an HMO.

**Inpatient.** (1) Pertaining to the medical care of an individual admitted to the hospital for at least 1 night. (2) That portion of the base relating to hospital admissions. Length of stay (DAYS) is at least 1 day.

**Inpatient payments.** All facility, professional, and other payments related to a hospital admission.

**Length of stay (LOS).** The number of days (DAYS) the patient was confined (spent in the hospital) during the inpatient admission. Also see *average length of stay*. 
Long-term disability (LTD). (1) A significant period of disability generally ranging from 6 months to life. (2) Wage replacement insurance for individuals who are (partially or totally) permanently disabled.

Mail-order pharmacy. A company that receives prescriptions from physicians or patients via fax or mail and then mails the medication to patients. Meanwhile, the physician provides the patient with enough of the medication to last until the prescription arrives. Generally, the cost per prescription from mail-order pharmacies is lower than the cost at other pharmacies because of higher volume and lower overhead.

Major diagnostic category (MDC). (1) A classification system for grouping medical conditions into 1 of 25 categories. The first 16 categories refer to major body systems; the remaining categories encompass more than one body system. (2) A widely recognized classification system that groups medical conditions into broad classifications, mostly by body system. Each DRG is assigned to one MDC.

Managed care. (1) Employing incentives at both the provider and patient levels that encourage the efficient provision of health care services. Common elements of managed care include capitation, a primary care physician acting as a gatekeeper, and patient copayments. (2) An organized system of health care services in contrast to the fee-for-service system.

Medical. Clinical in nature, as opposed to surgical.

Medicare. (1) A system of medical insurance provided by the federal government for all Americans aged 65 years and older and for Americans who are permanently disabled or have renal failure. (2) A federal program under Title XIX of the Social Security Act that provides health insurance for individuals aged 65 years and older and for other specified groups. Part A of Medicare covers hospitalization and is compulsory (i.e., automatically provided to any beneficiary who has qualified for participation in Social Security). Part B of the program covers outpatient services and is voluntary.

National Drug Code (NDC). A standard 12-digit coding system used to identify drugs on drug claims.

Not elsewhere classified (NEC). An abbreviation used to indicate the most generic category. There may be insufficient information to assign a more specific code.

Net pay. The portion of the charge for a health care service that the carrier paid to the beneficiary or assigned provider. NETPAY is calculated as PAY minus DEDUCT minus COPAY minus COINS minus COB.

Network providers. Providers who have contracted to be part of a plan’s network; they may be paid through a capitated or a discounted fee-for-service arrangement. Patients who visit out-of-network providers generally pay greater out-of-pocket amounts.

Open-ended HMO. An HMO that allows the patient to receive services from a nonnetwork provider. Although such services will be covered, the patient must pay higher-than-normal copayments and deductibles.

Out-of-pocket (OOP) costs. The portion of the claim that the patient or enrollee is obligated to pay (e.g., copayments, coinsurance, deductible). There typically is an annual OOP maximum. If the maximum is met, the insurer pays 100 percent of the costs incurred by the enrollee for the remainder of the plan year.

Paid date. The date on which a claim is paid (PDDATE). Claims data usually are received from carriers on the basis of paid date. For example, a submitted data file may contain all claims that were paid during the fourth quarter of 2013, regardless of when the claims were incurred. See incurred date, claims lag.

Point-of-service (POS) plan. Replacement of an indemnity plan. (1) A managed care plan that pays (reduced) benefits when patients receive health care services either from non-managed-care network providers or without proper referral by their primary care physician. (2) A benefit plan design in which enrollees must access the health care system through a gatekeeper. In addition to differential coinsurance and copayment levels described under
PPO, POS plans may include a differential deductible for in- and out-of-network services used (e.g., in-network deductible may be $250 and out-of-network deductible may be $500).

Precertification or preauthorization. Permission from the administrator for the hospital admission to occur or the services to be performed. This is a form of utilization review based on the patient’s health status and treatment needs.

Preferred provider arrangement or prudent purchaser arrangement (PPA). Same as a preferred provider organization.

Preferred provider organization (PPO). (1) A health plan that gives patients lower rates if they use the physicians in the preferred group of providers. Patients may use doctors outside that list, but they usually pay more to do so. Participating physicians normally are under a contract and keep an independent practice in the community. They also typically enroll in other preferred provider programs. Physicians receive reduced rates in return for a larger patient flow—lower price for the promise of higher volume. (2) Providers (e.g., hospitals, physicians) offering discounts or other reduced rates to a health care purchaser. Patients usually are “channeled” by receiving improved benefits (e.g., lower/no deductibles or copayments). See EPO, point-of-service PPO.

Premium. An amount paid periodically to purchase health benefits; for self-insured groups that do not purchase insurance, the term may refer to the per employee or per family cost of health benefits and may be used for planning and analysis purposes, even when no contribution to coverage is collected from the employee.

Primary care physician (PCP). The physician whom a patient in a managed care plan must see first for any health problem; the PCP acts as a gatekeeper and determines whether and when the patient needs to see a specialist. PCPs generally are internists, pediatricians, family physicians, general practitioners, and occasionally obstetricians/gynecologists.

Procedure group. Outpatient procedure groupings based on CPT-4 and HCPCS procedure code values.

Provider. A person or organization that provides health care services, such as a physician or hospital.

Referral. (1) Written authorization from a patient’s PCP for the patient to see a specialist. (2) An arrangement for a patient to be evaluated or treated by another provider.

Reimbursement. The dollar cost of covered products and services for which insurers pay.

Risk sharing. An agreement whereby the risks of providing care under a capitated arrangement are shared by multiple parties. For example, a pharmaceutical manufacturer assumes a portion of the financial risk for the use of a product with the provider. A risk-sharing arrangement may include a capitated payment for the unlimited use of a product, promotion of appropriate usage by the manufacturer, or performance guarantees based on predetermined outcomes.

Runoff period. The period of time representing the number of months between a claim’s service date and paid date. For example, if the runoff month’s variable is equal to 6, it indicates that most claims are paid within 6 months of their service date.

Self-insurance. Funding of medical care expenses in whole or part through internal resources rather than through transfer of risk to an insurer.

Service date. The date that a medical care service is provided (SVCDATE).

Service level. A variable that is found in the Inpatient Services Table. These variables can be different for each service within an admission. Examples are service date, provider ID, diagnosis and procedure codes, and financial variables that contain only the amount for that service (e.g., charge, payment). See case level for comparison.

Short-term disability (STD). (1) Wage replacement insurance for individuals temporarily disabled because of nonoccupational injury or illness. (2) Often considered to be a disability lasting not longer than 6 months.

Stop-loss (out-of-pocket maximum). (1) Usually, this refers to the maximum out-of-pocket amount that an individual or family could pay in a single plan year, including deductibles and copayment amounts. Alternatively, it may refer to the total...
dollar value of covered services after which the plan pays 100 percent. (2) The maximum out-of-pocket liability for a patient each year for deductibles, copayment, and coinsurance.

**Subrogation.** The assumption by a third party (such as an insurance company) of another’s legal right to collect a debt or damages. It is related to COB (e.g., recoveries from auto insurance may reduce an insurer’s health benefit liability).

**Summary Plan Description (SPD).** A legally required document that summarizes a company’s health care benefit plan.

**Surgical.** Pertaining to a service performed by a surgeon or involving surgery.

**Third-party administration or administrator (TPA).** (1) Administration of a group insurance plan by some person or firm other than the insurer or the policyholder. TPAs also may pay claims. (2) The administrator or claims administrator.

**Total charges.** Total eligible charges, prior to reductions for reasonable and customary limits and PPO discounts.

**Total payments.** Total eligible charges less any reasonable and customary amounts and discounts for PPO services, but prior to reductions for deductibles, copayments, and other savings.

**Uniform Billing (UB).** A standardized billing format for hospitals to use when submitting data to third-party payers. The term usually is followed by a year that indicates when the format was last revised (e.g., UB04).

**Unbundling.** Creative or fraudulent billing practices used by providers to increase payment by charging for components of a medical procedure on an item-by-item basis. **Usual, customary, and reasonable (UCR).** A method of payment to physicians based on the usual (U) charge of a particular physician for the procedure, the customary (C) charge for the procedure among physicians in the community, and a determination of what a payer’s reasonable (R) payment should be. This system is highly inflationary, because physicians typically increase their charges substantially to ensure that they attain a certain income. Plans often pay a percentage of UCR or a percentage of R and C. The patient is liable for the remainder, unless the physician is contractually obligated to accept the adjusted payment in full. (Balance billing is the practice of billing the patient for the remainder.)

**Utilization review (UR).** (1) A generic term referring to any program to control hospital runoff and runup admissions, lengths of stay, or both. Examples are second surgical opinion programs, length-of-stay certification, concurrent review, and preadmission certification. (2) A managed care process focused on the point at which care is (or is to be) provided, typically for expensive events; for example, in the case of hospital admission or outpatient surgery, a third party reviews the necessity and appropriateness of the procedure against medical criteria.

**Wellness benefits.** A broad range of employer or union-sponsored facilities and activities designed to promote safety and good health among employees. The purpose is to increase worker morale and reduce the costs of accidents and ill health such as absenteeism, lower productivity, and health care costs. It may include physical fitness programs, smoking cessation, health risk appraisals, diet information and weight loss, stress management, and blood pressure screening.

**Withhold amount/pool process.** The dollar amount retained or withheld from the servicing provider and placed in a risk-sharing pool for future distribution.
In preparing an analytic plan, it may be useful to refer to studies that have used the IBM MarketScan Research Databases. It also may be helpful to examine other references regarding analysis of administrative data from these databases. Since 1988, health care researchers have used MarketScan data to understand disease progression, treatment patterns, health outcomes, and their associated costs to patients, employers, health plans, and the government. The MarketScan Databases are fully compliant with the Health Insurance Portability and Accountability Act (HIPAA) of 1996. They are considered the gold standard in proprietary databases used for health care research in the United States. MarketScan data have been used in more than 300 peer-reviewed articles published in leading journals since the first article by J. B. Hillman and colleagues appeared in the *New England Journal of Medicine* in 1990. Research using MarketScan data has made a substantial contribution to the body of literature used to formulate policy decisions and improve health care for Americans.

**MarketScan Studies: Abbreviated Bibliography 2018:**


Akhuemonkhan E, Parian A, Carson KA, Hutfless S. Adverse reactions after intravenous iron infusion among inflammatory bowel disease patients in the United States, 2010-2014. *Inflamm Bowel Dis* 2018; April 13; [Epub ahead of print].


---


Castillo H, Chintapalli RTV, Boyajian HH, Cruz SA, Morgan VK, Shi LL, Lee MJ. Lumbar discectomy is associated with higher rates of lumbar fusion. Spine J 2018; May 21; [Epub ahead of print].

Cerullo M, Gani F, Chen SY, Canner JK, Herman JM, Laheru D, et al. Assessing the financial burden associated with treatment options for resectable...


Coleman CI, Bunz TJ, Eriksson D, Meinecke AK, Sood NA. Effectiveness and safety of rivaroxaban vs warfarin in people with non-valvular atrial fibrillation and diabetes: An administrative claims database analysis. Diabet Med 2018; April 16; [Epub ahead of print].


Dandolu V, Pathak P. Health resource utilization and costs during the first 90 days following robot-assisted hysterectomy. Int Urogynecol J 2018; 29(6): 865–872.


Day AT, Chang HY, Quon H, Kang H, Kiess AP, Eisele DW, Frick KD, Gourin CG. Treatment, short-term outcomes, and costs associated with


Ellhoff JJ, McHugh KE, Buckley JR, Morris SA, Simpson KN, Scheurer MA. Out-of-pocket medical expenses in severe CHD. Cardiol Young 2018; June 20; [Epub ahead of print].

Ellison JS, Dy GW, Fu BC, Holt SK, Gore JL, Merguerian PA. Neonatal circumcision and urinary tract infections in infants with hydronephrosis. Pediatrics 2018; June 7; [Epub ahead of print].


Gargano JW, Zhou F, Stokley S, Markowitz LE. Human papillomavirus vaccination in


Hepp Z, Lage MJ, Espaillat R, Gossain VV. The association between adherence to levothyroxine and economic and clinical outcomes in patients with hypothyroidism in the US. J Med Econ 2018; June 22; [Epub ahead of print].


Kalb LG, Stuart EA, Vasa RA. Characteristics of psychiatric emergency department use among privately insured adolescents with autism spectrum disorder. Autism 2018; January 1; [Epub ahead of print].


Khan AI, Patil D, Kawwass JF, Zholudev V, Mehta A. Surgical sterilization among US men and women with employer-based insurance: A claims data analysis. Contraception 2018; May 18; [Epub ahead of print].


Kim SC, Solomon DH, Rogers JR, Gale S, Klearman M, Sarsour K, Schneeweiss S. No difference in cardiovascular risk of tocilizumab versus abatacept for rheumatoid arthritis: A multi-
database cohort study. Semin Arthritis Rheum 2018; March 22; [Epub ahead of print].


Lee HH, Dalesio NM, Lo Sasso AT, Van Cleve WC. Impact of clinical guidelines on revisits after ambulatory pediatric adenotonsillectomy. Anesth Analg 2018; June 12; [Epub ahead of print].


Lewis SJ, Olufade T, Anzalone DA, Malangone-Monaco E, Evans KA, Johnston S. LDL cholesterol levels after switch from atorvastatin to rosuvastatin. Curr Med Res Opin 2018; February 21; [Epub ahead of print].

Li K, Moubayed SP, Spataro E, Most SP. Risk Factors for corrective septorhinoplasty associated with initial treatment of isolated nasal fracture. JAMA Facial Plast Surg 2018; June 14; [Epub ahead of print].


MacGeorge CA, Simpson KN, Basco WT Jr, Bundy DG. Constipation-related emergency department use, and associated office visits and payments among commercially insured children. Acad Pediatr 2018; April 16; [Epub ahead of print].


McHorney CA, Crivera C, Laliberté F, Germain G, Wynant W, Lefebvre P. Adherence to rivaroxaban versus apixaban among patients with non-valvular atrial fibrillation: Analysis of overall population and

Mellinger JL, Shedden K, Winder GS, Tapper E, Adams M, Fontana RJ, Volk ML, Blow FC, Lok ASF. The high burden of alcoholic cirrhosis in privately insured persons in the United States. Hepatology 2018; March 26; [Epub ahead of print].


Mosenthal WP, Boyajian HH, Ham SA, Conti Mica MA. Treatment trends, complications, and effects of comorbidities on distal radius fractures. Hand (NY) 2018; January 1; [Epub ahead of print].


Robinson LC, Anderson RCE, Brockmeyer DL, Torok MR, Hankinson TC; Pediatric Craniocervical Society. Comparison of fusion rates based on graft material following occipitocervical and atlantoaxial arthrodesis in adults and children. Oper Neurosurg (Hagerstown) 2018; March 15; [Epub ahead of print].

Rose S. Robust machine learning variable importance analyses of medical conditions for health care spending. Health Serv Res 2018; March 11; [Epub ahead of print].

Rudy SF, Li K, Moubayed SP, Most SP. Risk of venous thromboembolism in patients with keratinocyte carcinoma. JAMA Facial Plast Surg 2018; May 24; [Epub ahead of print].


Tang WY, Grothe D, Keshishian A, Morgenstern D, Haider S. Pharmacoeconomic and associated cost savings among women who were prescribed systemic conjugated estrogens therapy compared with those without menopausal therapy. Menopause 2018; 25(5): 493–499.


Tortella BJ, Alvir J, McDonald M, Spurden D, Fogarty PF, Chhabra A, Pleil AM. Real-world analysis of dispensed ius of coagulation factor IX and resultant expenditures in hemophilia B patients receiving standard half-life versus extended half-life products and those switching from standard half-life to extended half-life products. J Manag Care Spec Pharm 2018; January 24; [Epub ahead of print].


Winterfield JR, Kent AR, Karst E, Dalal N, Mahapatra S, Bunch TJ, Reynolds MR, Wilber DJ.


Zhang X, Song X, Lopez-Gonzalez L, Jariwala-Parikh K, Cong Z. Economic burden associated with adverse events of special interest in patients with relapsed Philadelphia chromosome-negative B-cell acute lymphoblastic leukemia in the united


2017


Lagman C, Ugiliweneza B, Boayke M, Drazin D. Spine surgery outcomes in the elderly versus the general adult patient population in the United


2016


Baggs J, Fridkin SK, Pollack LA, Srinivasan A, Jermigan JA. Estimating national trends in inpatient

IBM Watson Health. MarketScan Medicaid User Guide


Brown JD, Shewale AR, Dherange P, Talbert JC. A comparison of oral anticoagulant use for atrial


Carney PI, Lin J, Xia F, Law A. Temporal trend in the use of hysteroscopic vs laparoscopic sterilization and the characteristics of commercially insured and Medicaid-insured females in the US who have had the procedures. International Journal of Women’s Health 2016; 8: 137–44.


Clemow DB, Nyhuis AW, Robinson RL. Clinical impact of not achieving recommended dose on duration of atomoxetine treatment in adults with attention-deficit/hyperactivity disorder. CNS Neuroscience & Therapeutics 2016; 22(12): 970–8.


Nazarian S, Reynolds MR, Ryan MP, Wolff SD, Mollenkopf SA, Turakhia MP. Utilization and


Neugut AI, Zhong X, Wright JD, Accordino M, Yang J, Hershman DL. Nonadherence to medications for chronic conditions and nonadherence to adjuvant hormonal therapy in women with breast cancer. JAMA Oncology 2016; 2(10): 1326–32.


Nichols CI, Vose JG. Clinical outcomes and costs within 90 days of primary or revision total joint arthroplasty. Journal of Arthroplasty 2016; 31(7): 1400–6.e3.


Obama B. United States Health care reform: Progress to date and next steps. JAMA 2016; 316(5): 525–32.


Pace LE, Dusetzina SB, Keating NL. Early impact of the Affordable Care Act on uptake of long-acting reversible contraceptive methods. Medical Care 2016; 54(9): 811–7.


Piccini JP, Mittal S, Snell J, Prillinger JB, Dalal N, Varma N. Impact of remote monitoring on clinical


Veeravagu A, Connolly ID, Lamsam L, Li A, Swinney C, Azad TD, et al. Surgical outcomes of


APPENDIX A: FREQUENTLY ASKED QUESTIONS

Q1. How do I track individuals longitudinally across years and plans?

IBM Watson Health maintains a unique person-level identifier that is consistent across all tables, plans, and years. Individuals who moved from state to state cannot be tracked, even if the states are all MarketScan Database contributors.

Q2. How do I identify continuously enrolled covered lives?

To determine whether an individual was enrolled for an entire calendar year, the sum of Member Days (MEMDAYS) for each enrollment detail record should equal 365.

Q3. How do I know whether a patient’s lack of utilization data represents a lack of health care use or disenrollment from a plan?

You can match the patient’s utilization data to enrollment information by following these steps:

1. Create a subset of medical and/or outpatient pharmaceutical claims.
2. Use Enrollee ID (ENROLID) from the claims utilization as the subset of criteria for the enrollment data.
3. The resulting subset contains the enrollment records for the patients in the corresponding claims.

You can establish a fixed window of continuous enrollment by following these steps:

1. Use the Enrollment Detail Table and subset records with enrolled months that are within the time window of interest (i.e., all MEMDAYS > 0).
2. Subset the utilization information (e.g., claims) to Date of Service Incurred (SVCDATE) within the time window of interest.
3. Sort the utilization information (e.g., claims) by ENROLID.
4. Merge restricted and sorted enrollment data with sorted utilization information by ENROLID in cases in which records appear in both sets.

Q4. How can I ensure that diagnoses, procedures, and demographic information are in concordance with each other?

Diagnosis and procedure codes are edited for validity. If they are invalid, they are set to missing.

Q5. What variables can I use to calculate a rate (e.g., per capita, per employee)?

Metrics that require a population-based denominator (e.g., procedures per 1,000 covered lives) can be calculated only by subsetting demographic variables that are contained in the Enrollment Detail Table. Typical subsets for such counts include the type of plan (PLANTYP) or the sex of the patient (SEX).
Please refer to the MarketScan Enrollment Detail Tables in the Database Dictionary for a full list of population-supported variables.

Q6. How do I calculate utilization rates and payments by procedure?

When calculating a utilization rate by procedure, using the count of claims as the number of procedures oversates the number of procedures because a single procedure can generate more than one claim (e.g., a surgeon’s claim, an anesthesiologist’s claim, and a facility claim). Follow these steps to construct a day-episode record for the procedure to collapse the related services for each of the procedures of interest:

1. Using the variable PROC1, subset the Inpatient Services Table and/or the Outpatient Services Table for the procedures of interest.
2. To eliminate multiple claims, aggregate the data on ENROLID, PROC1, and SVCDATE to create one record per patient per procedure for a single service date. Sum any other variables of interest (e.g., Payment [PAY], Payments Net [NETPAY]). The number of procedures performed equals the record count in the resulting subset.
3. Divide the procedure count by the number of covered lives to calculate a utilization rate.
4. To calculate the covered life counts, sum POPCNT on the Populations Table and divide the resulting number by the number of calendar quarters.
5. To calculate payments per procedure, sum PAY and divide by the number of procedures.

Q7. Can a diagnosis be linked to drug claims (and vice versa)?

The Outpatient Pharmaceutical Claims Table does not contain diagnosis variables because they are not provided regularly by the physician on a prescription form. Therapeutic class (e.g., corticosteroids) is provided on the pharmaceutical claims, representing the broad classification of the drug. However, to impute the diagnosis, one must access the related medical claims for the individual—usually the claims filed within a specific time window around the prescription:

1. Subset the National Drug Codes of interest on the Outpatient Pharmaceutical Claims Table.
2. Use ENROLID and SVCDATE as the selection criteria to subset all services from the medical tables (I, S, O) that fall within a predefined time window around the SVCDATE. The resulting diagnoses on the medical claims may be associated with the pharmaceutical claim.

These steps may be modified to identify the prescriptions associated with a specific diagnosis. First, subset a diagnosis in the medical claims, and then select all pharmaceutical claims for each person with the diagnosis (using ENROLID as the linkage variable) within a predefined time window around the date of the prescription.

Q8. How do I count emergency department (ED) visits, which can occur in the Inpatient Services Table or in the Outpatient Services Table?

The Service Subcategory (SVSCAT) field can be used to identify most types of service. The field is structured so that the first three digits describe the facility type and the last two digits identify service type. To select ED visits, choose from the S or O table any records with a SVSCAT value that ends in 20.

Because multiple claim records can be generated for a single ED visit, count the number of ED visits by creating day-episode records from the data table produced by aggregating ENROLID/SVCDATE combinations. Accumulate all analytic variables of interest.

Q9. The National Drug Code in the MarketScan Database is 11 digits long, but the codes I have from my Food and Drug Administration search are only 10 digits long. How can I convert them?

The 10-digit codes should be padded with zeros in the appropriate places until the 11 digit, 5-4-2 format is established. See schematic below:

4-4-2: XXXX-XXXX-XX \(\rightarrow\) 0XXXX-XXXX-XX
5-3-2: XXXXXX-XXX-XX \(\rightarrow\) XXXXX-0XXX-XX
5-4-1: XXXXXX-XXXX-X \(\rightarrow\) XXXXX-XXXX-0X
Q10. Why are there no geographic variables?
Under the agreement with the contributors to this database, IBM Watson Health cannot reveal the identity of the states or health plans. Because certain states cover most of a single geographic region, geographic information on any level (from ZIP Code to census region) could not be included in this database.

Q11. Are prescription drugs covered under Medicare Part D in this database?
Prescription drugs covered under Medicare Part D are not included in this database, but some information about drugs for beneficiaries who were dually eligible for Medicaid and Medicare are available for data years before 2006, when Medicare Part D took effect. For individuals with dual Medicaid and Medicare coverage, Medicaid pays for certain services that are not covered by Medicare, but Medicare is considered the primary payer. When a claim is generated, Medicaid processes it only if Medicaid provides part of the payment. Historically, Medicare has not provided drug benefits, and Medicaid has provided drug benefits for individuals who are dually eligible. However, after the implementation of Medicare Part D, Medicare became the primary payer for outpatient prescription drugs for those with dual coverage. As of the 2006 data year, individuals who were dually eligible for Medicaid and Medicare in the MarketScan Multi-State Medicaid Database are noted as not having drug capture by the DRUGCOVG flag.