IQVIA PharMetrics® Plus

User Guide

June, 2020

TABLE OF CONTENTS

OVERVIEW

The IQVIA PharMetrics® Plus database is comprised of fully adjudicated medical and pharmacy claims. It contains a longitudinal view of inpatient and outpatient services, prescription and office/outpatient administered drugs, costs and enrollment information. With PharMetrics Plus, an enrolled patient can be tracked across all sites of care: hospital, specialist, emergency room, pharmacy, primary care, and more.

The enrollee population in the database is generally representative of the under 65 commercially insured population in the U.S. with respect to both age and gender.

Aside from being able to follow patients across sites of care, key aspects of the adjudicated claims data are that patient enrollment is explicitly captured on a month by month basis, and there is robust information around costs.

DATABASE STRUCTURE

The standard data deliverable consists of three types of files: health plan enrollment details, claims and reference files for clinical code sets.

Health Plan Enrollment Detail (ENROLL and ENROLL2)

There are two different files that will contain information relative to the patient's enrollment status, ENROLL, and ENROLL2.

ENROLL

The ENROLL file is a patient-level file of key demographics and enrollment standard variables for the selected population. The file consists of one record per enrollee. The records are typically sorted according to patient ID and can be linked to the CLAIMS file using the patient ID variable. The enrollment information contained in this file pertains to the patient's overall enrollment in the health plan.

ENROLL2

The ENROLL2 file is also a patient-level file with six string-based records per person. These string-based records serve as a month-by-month indicator of coverage at a more granular level. This file also provides product and payer type information by month, with six records per enrollee.

Claims Files (CLAIMS)

These files contain transactional records sourced from the claims submitted by providers to health plans for reimbursement. These claims may have been submitted a variety of ways, for example:

- HCFA-1500
- UB04
- via a pharmacy claims system
- via a PBM

Users will see all available medical and pharmacy claims for the selected population or cohort. For large cohorts, users may see multiple claims files.

Reference Files

Diagnosis Reference [pp_dx_lookup]

The diagnosis reference file contains reference information for the ICD-9 and ICD-10 diagnosis codes found in the database. This will include diagnosis names and descriptions.

Medication Reference [pp_rx_lookup]

The medication reference file contains reference information for the NDC codes found in the database. This will include product names, dosage form, route of administration, strength, and generic product identifier codes (GPI).

Procedure Reference [pp_pr_lookup]

The procedure reference file contains reference for the ICD-9, ICD-10, CPT and HCPCS procedure codes found in the database, including procedure names and descriptions.

Place of Service [pp_pos_lookup]

The place of service codes are codes used to specify where the service was rendered.

Revenue Codes [pp_rev_lookup]

The revenue codes represent a high-level description of services performed by a hospital/other facility.

ENROLLMENT

Enrollment is captured in the database in two different files, ENROLL and ENROLL2.

ENROLL File

The ENROLL file contains one record per individual enrolled in the plan. This record includes the demographics for that patient along with information relative to their enrollment in the health plan.

The file contains separate fields that represent the first and last dates of enrollment, along with a string variable 'ESTRING' which provides the month by month indication of enrollment in a health plan.

Patient Demographics

The ENROLL file contains the demographics for the patients. These include:

- Year of Birth
- Gender
- Geography (State, and/or Zip 3 if purchased)

The following modifications are made to the data to comply with the Health Insurance Portability and Accountability Act (HIPAA):

- For enrollees who are greater than 85 years of age in the current year, their year of birth (DER_YOB) is set to '0000'.

- For enrollees who are in a 3-digit zip with less than 20,000 people, their 3-digit zip is reset to '000'.

ESTRING

ESTRING is within the ENROLL file. It is a 360-character string that can contain only "X" or "-" characters. Each position in the 360-character string corresponds to a month and year, beginning in January 2001. An "X" indicates the individual was enrolled in the plan during that specific month.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2001	1	2	3	4	5	6	7	8	9	10	11	12
2002	13	14	15	16	17	18	19	20	21	22	23	24
2003	25	26	27	28	29	30	31	32	33	34	35	36
2004	37	38	39	40	41	42	43	44	45	46	47	48
2005	49	50	51	52	53	54	55	56	57	58	59	60
2006	61	62	63	64	65	66	67	68	69	70	71	72
2007	73	74	75	76	77	78	79	80	81	82	83	84
2008	85	86	87	88	89	90	91	92	93	94	95	96
2009	97	98	99	100	101	102	103	104	105	106	107	108
2010	109	110	111	112	113	114	115	116	117	118	119	120
2011	121	122	123	124	125	126	127	128	129	130	131	132
2012	133	134	135	136	137	138	139	140	141	142	143	144
2013	145	146	147	148	149	150	151	152	153	154	155	156
2014	157	158	159	160	161	162	163	164	165	166	167	168
2015	169	170	171	172	173	174	175	176	177	178	179	180
2016	181	182	183	184	185	186	187	188	189	190	191	192
2017	193	194	195	196	197	198	199	200	201	202	203	204
2018	205	206	207	208	209	210	211	212	213	214	215	216
2019	217	218	219	220	221	222	223	224	225	226	227	228
2020	229	230	231	232	233	234	235	236	237	238	239	240
2021	241	242	243	244	245	246	247	248	249	250	251	252
2022	253	254	255	256	257	258	259	260	261	262	263	264
2023	265	266	267	268	269	270	271	272	273	274	275	276

[ESTRING CONTINUES TO 2030]

If a participant is enrolled on any one day of a month, an "X" is populated for that month. The "-" ESTRING characters are used to populate any months prior to the date of first enrollment, at the conclusion of enrollment, and also for breaks between their first and last dates of enrollment.

Enroll First and Enroll Last

Enroll First (ENR_FRST) indicates the enrollees first date of health plan enrollment. Enroll Last (ENR_LAST) is this last date of enrollment. However, ENR_LAST is modified to correspond to the "complete thru" date of the database release. There is typically a 6-month lag on the "complete thru" date. A database delivered at the end of June will be considered "complete thru" the prior December. While claims will continue to flow into the database and we do not apply a cutoff date, any data beyond the "complete thru" date, and therefore beyond the ENR_LAST date for a patient, should be considered an incomplete view to that patient.

ENROLL2 File

In addition to the health plan enrollment information captured in the ENROLL file, a more granular view to enrollment is captured in the ENROLL2 file.

This file is also a patient-level file, but it contains 6 string-based records per individual providing a month-by-month indicator of benefit coverage.

Each position in the string represents a single month and the numbering begins as of January 2001 (same as ESTRING)

These 6 strings are:

- **Medical Benefit Indicator (Ben_Med):** Month by month indication of medical benefits
- **Pharmacy Benefit Indicator (Ben_Rx):** Month by month indication of pharmacy benefits
- **Pay Type (Pay_Type):** Month by month indication of payer type, e.g. Commercial, Self-Insured, etc.
- **Product Type (Prd_Type):** Month by month indication of type of plan, e.g. PPO, HMO, Indemnity, etc.
- **Medical Coordination of Benefits Indicator (Ben_MCOB):** Month by month indication of if medical coverage is primary or secondary
- **Pharmacy Coordination of Benefits Indicator (Ben_PCOB):** Month by month indication of if pharmacy coverage is primary or secondary

Appropriate codes are located in the data dictionary.

Continuous Enrollment

Many studies will require periods of what is commonly referred to as “continuous enrollment”, or periods of time during which a participant must have consecutive month by month flags indicating they were enrolled.

Continuous enrollment is applied to analyses to ensure a comprehensive view of the patient during the study period and/or look-back period. For example, an analysis focused on newly diagnosed patients would want to ensure the patients were continuously enrolled for a period of time prior to the first occurrence of the diagnosis in the database. This would enable the categorization of ‘newly diagnosed’ based on the diagnosis not appearing in the patient’s history during the prior X months.

There are a few different ways continuous enrollment is incorporated into projects.

For example:

- Require all individuals to be continuously enrolled in the calendar year.
- Require each individual to be enrolled for a minimum amount of time (e.g. at least 12 months of continuous enrollment during the project time period), recognizing these time periods can be different from person to person.
- Set enrollment based on an index date, typically some type of event like the appearance of a diagnosis or treatment (e.g. individuals must have at least 12 months of continuous enrollment prior to and after their diagnosis of ABC.DE).

Conversely, there are situations where requiring continuous enrollment could bias your findings. For example, if this disease area being studied had a high mortality rate, requiring patients to be enrolled for a long time period could eliminate the most severe patients from the analysis.

Record Types

Each record is categorized into one of 6 types, F, P, S, M, A, J, described below.

Record Type	Name	Description	General assignment logic
F	Facility	Identifies room and board accommodations in Inpatient setting	* Claims with room and board revenue codes * Claims without NDC, revenue or CPT codes with type of service equal to 2 (i.e. tos_flag = 2 - Room and Board)
P	Pharmacy	Retail Pharmacy drugs and drugs administered in office, Outpatient or Inpatient setting	* Claims with NDC code present * Claims with HCPCS codes for drugs, including J codes
S	Surgery	Surgical procedures provided by clinician or facility provider	*Claims with CPT/HCPCS of a surgical nature, performed by a clinician or facility provider (i.e. ptypeflg is 0 or 1)
M	Management	Procedures that evaluate a patient's condition	* Claims with CPT/HCPCS management codes - some codes require provider to be clinician or facility (i.e. ptypeflg is 0 or 1) and others do not * Claims with Management revenue codes and no CPT/HCPCS (e.g. clinic visit, professional fees)
A	Ancillary	Procedures for ancillary services or any other procedures or services that did not meet criteria of other record type assignments	* Claims for CPT/HCPCS codes for ancillary services * Claims with CPT/HCPCS for surgery or management procedures that did not meet provider type requirement for those services * Claims with revenue codes and no CPT code that did not meet Facility or Management record definition * Claims without any NDC, CPT or revenue code where type of service is not 2 (i.e. tos_flag <> 2)
J	Reconciliation	Cost reconciliation	Not based on revenue or procedure codes - only needed when total claim cost does not equal cost of individual claim lines

Note: Claim records that have negative costs or invalid procedure codes (CPT/HCPCS) will not get a record type

Record types are used to help segment the data. Some usage examples follow below:

Facility Records are reserved for inpatient room and board codes. These “F” records are used to calculate length of stay. They are also frequently used to isolate the last facility record within a confinement. The diagnosis on this record is typically a proxy for principal diagnosis.

Ancillary Records are records that did not meet the criteria for facility, pharmacy, surgery or management codes. Oftentimes, these are records associated with laboratory testing.

When defining a cohort with a certain disease, if the only time a diagnosis appears for a patient is on an ancillary record, it could be that the patient was being tested to “rule-out” a condition.

Diagnoses

There are 13 total diagnosis fields on each claim record. Diagnosis 1 thru 12 and Admit Diagnosis.

Prior to October 1, 2015, the diagnosis codes are ICD-9-CM codes. From October 1, 2015 forward, the codes are ICD-10-CM codes. The DIAGPRC_IND field indicates whether all codes on that claim line are ICD-9-CM or ICD-10-CM codes.

Admit Diagnosis

The admitting diagnosis is the condition identified by the physician at the time of the patient’s admission requiring hospitalization. This is considered the “unconfirmed” principal diagnosis. Principal diagnosis is considered the confirmed reason for admission and is typically reflected in the DIAG1 field.

While the field name is “Admit” diagnosis, the field is also populated for claims that were not part of an inpatient stay. For outpatient claims, the field is considered the patient’s reason for visit.

The vast majority of the non-inpatient claims with the Admit Diagnosis field populated come from either Outpatient Hospital or Emergency Room places of service.

The admit diagnosis field is only found on institutional claims (UB 04/CMS 1450).

Diagnosis (1-12)

There are 12 diagnosis code fields in the database (in addition to Admit diagnosis).

Guidance for working with diagnosis fields

Primary diagnosis is expected to be the most serious/resource intensive diagnosis.

While it is expected that the primary diagnosis is the same as the principal diagnosis and is in the DIAG1 field, typically all diagnosis fields are used when looking for the presence of a condition.

We advise researchers to use DIAG1 on the last “F” record of the confinement as the proxy for the principal diagnosis.

To create a study cohort based on diagnosis there are a few different techniques that are frequently used.

For example:

- do not include a patient if the only time the diagnosis of interest appears is on an Ancillary record type. This is where a ‘rule-out’ diagnosis could be picked up – where testing was done to rule out a certain disease.
- the analysis might require the patient have more than one occurrence of the diagnosis code. Or to be stricter, more than one occurrence of the diagnosis code, at least X days apart.

Treatments

Treatments captured in the database represent dispensed or administered therapies. The database will not capture written prescriptions that were not filled, nor items like over-the-counter drugs or inexpensive drugs that were paid for in cash and did not generate any claim submission to the patient’s payer.

Treatments are captured via either NDC or HCPCS codes.

Please note that drugs administered in the inpatient setting are not consistently captured. Their bundled costs will be presented via Revenue codes.

National Drug Code (NDC)

The National Drug Code (NDC) will be captured in the data. This is an 11 digit number that represents the labeler (manufacturer, repackager or distributor), product name (including strength and dosage form), and package size.

The medication reference file contains reference information for all of the NDC codes found in the database. This will include product names, dosage form, route of administration, strength, and generic product identifier codes (GPI).

Healthcare Common Procedure Coding System (HCPCS)

The CMS has developed the Healthcare Common Procedure Coding System (HCPCS) for reporting medical procedures and services. Some of the most commonly used HCPCS Level II Codes, J-codes are used for non-orally administered medication, chemotherapy and immunosuppressive drugs, and inhalation solutions as well as some orally administered drugs. These Level II codes are five (5) characters in length and are comprised of one (1) letter and four (4) numbers.

Note, there are also medications captured in the temporary HCHPS codes (C/Q).

For example:

- Ustekinumab, for intravenous injection 1mg had the code C9487 from 4/1/2017-6/30/2017.
- The C code was discontinued and from 7/1/2017-12/31/2017 the code was Q9989.
- Beginning January 1, 2018, the permanent J code of J3358 was assigned.

When querying the database, make sure to include all relevant codes that were in effect during the analysis time period. The same drug may have multiple J codes if there are different types of administration options (e.g. IV vs. subcutaneous injection).

The HCPCS codes are captured in the field named PROC_CDE or the field ICDPRC(1-12). These same fields will also contain CPT-4 codes, described in the Procedures section.

Provider Specialty

Provider specialties are mapped to one of IQVIA's standard specialties for PharMetrics Plus. There can be more than one provider specialty on a claim record.

The rendering and billing provider's specialties are most frequently populated, and they are often the same specialty. However in some settings of care, the billing specialty might be "hospital" for example, while the rendering specialty will be associated with the treating provider.

The Attending provider specialty is only present on inpatient claims. Note that the Attending provider may be directing the care provided by students and not physically rendering care.

The vast majority of the time, NDC records do not contain the specialty of the provider. Prescriber specialty is not required to fill a prescription claim. When the specialty is populated, it is often 'pharmacy' or 'other' for an NDC claim.

Procedures

Inpatient and outpatient procedures are captured via different standard coding schemes and captured in multiple fields in the database.

Procedure Code (PROC_CDE) Field

This field contains a mix of CPT and HCPCS codes.

CPT codes are 5 digit codes. They start with a number and are typically numeric, but there are a subset of codes that have 4 numbers and end with a letter. They are used to capture medical services and procedures performed by physicians and other health care professionals. The CPT system is maintained by the American Medical Association.

HCPCS codes are a single letter followed by 4 numeric digits. This coding scheme is intended to capture products, supplies and services not included in the CPT codes. The HCPCS system is maintained by CMS.

HCPCS codes are also referenced under the Treatment section as this is how many non-orally administered treatments are captured.

Procedure codes are on both professional and institutional claims. This field is at the claim detail level (vs. header level)

Note that while we do capture some HCPCS during inpatient stays, the presence/absence of these codes will depend on how billing is done at that institution. If billing is based on case rate or DRGs, the codes may not be present. In addition, some claim lines may only list the revenue codes for these services and not the more specific CPT/HCPCS codes.

Inpatient Procedure (ICDPRC1 thru ICDPRC12) Fields

Inpatient Procedures are captured via ICD-9-Procedure Coding System (PCS) codes until October 1, 2015 or ICD-10-PCS beginning October 1, 2015.

ICDPRC1 is the primary procedure on an institutional claim. ICDPRC2 thru ICDPRC12 represent additional inpatient procedures on the claim.

ICDPRC codes are on institutional claims, but there are times when institutional claims are submitted for outpatient services, for example where outpatient hospital is the place of service.

Note that ICDPRC codes are captured at the ‘header’ record level, are repeated for each claim line within an individual claim and cannot be associated with individual lines or their costs. In addition, these are surgical procedures and will not include drugs, lab tests, etc.

For claims that occur in an outpatient setting, these fields will sometimes contain CPT/HCPCS codes.

Revenue Code (REV_CODE) Field

The Revenue Center Code field represents a high-level description of services performed by a hospital/other facility. Revenue codes are present on Institutional claims (UB 04/CMS 1450) and can be used to capture both inpatient and outpatient services. Revenue codes are maintained by CMS. A reference file for Revenue Codes is included with delivery of a transaction level extract.

Confinements

Confinements are identified within the database via the presence of a confinement number. All of the claims with the same patient ID and confinement number belong to the same unique inpatient stay.

Within a patient ID, a confinement number is assigned by using the dates of the first room and board code (identified via the record type of F) and the last room and board code in a series of records that have overlapping or contiguous dates and the same provider ID (the “provider” for a room and board code is the billing provider – typically this will indicate “hospital”).

The minimum date (FROM_DT) on all claims within the confinement will be the start of the confinement. The maximum date (TO_DT) will be the end of the confinement.

Once the bounds of the confinement are established, all records for that patient that fall within those bounds, regardless of record type, are assigned the same confinement number.

A given confinement can have multiple F records as long as there is no change in the assigned billing provider ID. Therefore, a patient transferred from a CCU room to a standard room at the same facility may have multiple ‘room and board’ codes but all codes would be within the same confinement if there is no gap in the dates of service.

Costs During Confinements

All costs during a confinement should be summed. Not all providers will allocate charges the same way. For example, one hospital might show daily room and board charges while another might lump all the daily charges into one entry. In addition, some claim lines from the hospital may not show any costs as the costs are allocated to only one or two lines.

Revenue Codes (rev_code)

Revenue codes are captured on Institutional claims. They are descriptions and dollar amounts charged for hospital services provided to a patient. The revenue code indicates where the service was performed or indicates a type of service.

For example:

- **0113** Room and Board, Private – Pediatric
- **0291** Durable Medical Equipment - Rental
- **0324** Radiology, Diagnostic – Chest x-ray

Patient Discharge Status Codes (patstat)

Patient discharge status codes are captured on Institutional claims. A patient discharge status code is a two-digit code that identifies where the patient is at the conclusion of a health care facility encounter (this could be a visit or an inpatient stay) or at the time end of a billing cycle (the 'through' date of a claim).

CMS requires patient discharge status codes for: Hospital Inpatient Claims; Skilled Nursing Claims; Outpatient Hospital; All Hospice and Home Health Claims

Discharge Status codes - examples:

- **01** Home/Self care
- **07** Left against medical advice
- **09** Admitted as Inpatient

***Note*:** Our current privacy certification requires us to exclude discharge status codes indicative of death to protect patient privacy.

Physician Services During Confinements

A physician might bill separately for their services during the course of a confinement. Due to this, CPT/HCPCS codes can appear during the confinement as the way some procedures are identified, in addition to ICD9/10 PCS codes.

Treatment During Confinements

The database does not consistently capture drug therapy administered during an inpatient confinement. This is based on how the treatments are billed – we may receive only revenue codes indicating drugs were administered.

Emergency Services

Emergency Services are captured in the database several different ways. When we think about Emergency Services, we typically think about three distinct scenarios.

In the first scenario, the individual visits the Emergency Room and is discharged from the Emergency Room.

More complex than this, the second scenario is where an individual visits the Emergency Room and ends up in “Observation Status”. Observation status according to CMS is a well-defined set of specific, clinically appropriate services, which include ongoing short-term treatment, assessment, and reassessment, that are furnished while a decision is being made regarding whether patients will require further treatment as hospital inpatients or if they are able to be discharged from the hospital. In most cases a patient does not remain in observation status for more than 48 hours.

These patients can have claims for consecutive days but they are not considered inpatients and these records will not be assigned a confinement number.

In the third scenario, the patient originally presents at the Emergency Room and is admitted. In this scenario, all of the services in the Emergency Room would be assigned the same confinement number as the admission.

Costs

There are a few different fields capturing information related to costs. Those are summarized below:

Field	Description
Allowed Amount	The contracted or accepted reimbursable amount for covered medical services or supplies the health plan agrees to pay to service providers. RECOMMENDED
Charge Amount	This is the amount charged/billed for services provided by the servicing provider or facility. This amount is highly variable and not typically close to what is paid.
Co-insurance Amount	The amount the insured individual pays to the provider out of pocket. Typically derived as a set percentage of the cost of covered medical services. The amounts in this field are the dollar amounts.
Coordination of Benefit Amount (COB Amount)	If this field is populated, it means the payer reflected in the data is **not** the primary payer, and the amount in this field reflects what was paid by the primary payer.
Copayment	Amount an insured individual pays directly to a provider at the time the services or supplies are rendered. Typically, a fixed amount per service, such as $15.00 per office visit. Amounts may include any penalty for an insured's noncompliance such as lack of prior authorization for services rendered out of network. Also includes any amount paid as copayment or coinsurance for prescription medications.
Deductible	The portion of this service that the enrollee must pay which is applied to the total period deductible. Deductibles are usually applied over a specific time period, such as per calendar year, per benefit period, or per episode of illness.
Paid	This is the actual amount paid to the provider by the health plan/payer. RECOMMENDED

Guidance for working with cost fields

Allowed amount and Paid amount are the recommended variables to use when looking at costs. The different between allowed amount and paid amount typically represents the patient's liability.

There may be situations where all of the cost fields on a claim are zeros. There are a number of different possible reasons for this:

- Claims from capitated plans will typically appear with zeros in the cost fields.
- On Outpatient/office claims it is possible to have items not paid explicitly by insurance carrier as they are considered included in the rest of the costs.
- If a claim is paid on case/DRG basis (typically Inpatient claims), the costs will be on one line of the claim so other lines will show zero amounts.
- Costs paid by someone else (for example, children's vaccine supplied by the state); the provider will put the item on the claim but not charge the insurance carrier since the item is paid for by another entity.

RESEARCH CONSIDERATIONS

Listed below are items to consider when working with the data.

Patient Privacy Modifications

Variable	Modification
Patient Age	Patients that would appear over age 85 in current year have their year of birth changed to '0000'
Zip Codes	Those outside the U.S. or with less than 20,000 residents are set to '000'
Place of Service	Place of service codes identified as 'high risk' are set to 'ZZ'
Death and Morbidity	Diagnosis codes indicating a patient has died and other highly sensitive codes are removed from the database. Discharge codes indicating a patient has expired are removed from the database

Provider Dimension Considerations

Item	Details
Database is payer (patient) centric	Database does not provide a comprehensive view to individual physicians; unknown percentage capture of any individual provider or practice
Limited information on physician	Summary specialty information
Contractual restrictions	Payer/plan cannot be identified, within the database as a standalone database or combined with any other data

Patient Dimension Considerations

Item	Details
Only covered services captured	Plans may have carve outs (e.g. Mental Health)
Demographics	HIPAA restrictions around patient age, zip 3, etc
Database is primarily a commercially insured population	Under-representation of 65+ population and therefore diseases in that demographic Certain diseases covered by government programs, e.g. ESRD, will be under-represented

Disease Dimension Considerations

Item	Details
Only covered services captured	Non-covered services (e.g. cosmetic surgery), any services paid by cash not captured
Data is captured for reimbursement purposes	Symptoms/side effects less likely to be captured Lifestyle issues (e.g. smoker, diet/exercise) not captured No vitals or lab results (tests performed but no results) No measures of severity (e.g. tumor stages) unless reflected in Diagnosis code

Treatment Dimension Considerations

Item	Details
Inpatient Treatment	Limited visibility to drugs administered in the inpatient setting
Drug route of administration	Injectable drugs may be captured via NDC or HCPCS codes, multiple codelists will need to be created
Only covered services	Over the counter treatments will not be captured Low cost drugs, like $4 generics will not be captured