TAF Technical Documentation: Claims Files

June 2022

Contents

I. Introduction 1

II. Records in the TAF claims files 2

III. Linking header-level and line-level records 4

IV. Linking claim and eligibility records 5

V. Identifying different types of records in the claims files 6

VI. Identifying and using managed care encounter data 9

VII. Identifying service use and payment records for specific Medicaid programs 9

VIII. Identifying service use and payment records for specific time frames 10

IX. Classifying and analyzing service use records 11

A. Distinguishing between institutional and professional claims 11

B. Identifying claims and encounters for specific procedures or services 12

C. Identifying services delivered by specific types of providers 13

D. Identifying services delivered in certain settings 15

E. Identifying services for the diagnosis and treatment of certain conditions 16

F. Identifying unique inpatient hospital stays 17

X. Assessing expenditures 19

A. FFS expenditures 19

B. Monthly beneficiary payments 20

C. Costs to managed care organizations 21

D. Service tracking claims 22

E. Supplemental payments 23

References 25

Appendix A: Revision History 26

Appendix B: Sample Claim Forms 28

Tables

Table 1. TAF claims files 2
Table 2. TAF inclusion criteria 3
Table 3. Records to use in various analyses 6
Table 4. FASC values 7
Table 5. Values for claims type code (CLM_TYPE_CD), by program type and record type 8
Table 6. Provider variables across the TAF files 15

I. Introduction

States administer Medicaid and the Children's Health Insurance Program (CHIP) and share the responsibility for funding and program administration with the federal government.[1] Each state compiles standardized data on Medicaid and CHIP enrollment, service utilization, payment, providers, managed care plans, and other information from its own eligibility and claims data systems into the federal Transformed Medicaid Statistical Information System (T-MSIS). States submit into T-MSIS both service use records (including fee-for-service claims and managed care encounters) and payment records (including capitation payments made to managed care plans and service tracking or supplemental payments made to providers) together in claims files organized by service and provider type.[2,3]

The Centers for Medicare and Medicaid Services (CMS) administers T-MSIS to improve quality of care and program integrity and to meet stakeholders' needs. Although states submit a wide variety of information to T-MSIS, the system is not optimized for conducting analyses. To meet this need, CMS constructs a research-optimized version of T-MSIS data called the T-MSIS Analytic Files (TAF).[4,5] Information on the completeness and quality of key TAF data elements can be accessed through *DQ Atlas,* available at https://www.medicaid.gov/dq-atlas/welcome. Specific topics relevant to each section of this technical documentation are noted in the footnotes.

The TAF are released as TAF Research Identifiable Files (RIF).[6] The TAF RIF include monthly claims files containing Medicaid and CHIP service use and payment records, as well as annual files containing demographic and eligibility data for all Medicaid- and CHIP-eligible beneficiaries and information on all Medicaid- and CHIP-enrolled providers and managed care plans. The four monthly claim files—the

[1] For more information about the Medicaid and CHIP programs, see the CMS website: https://www.Medicaid.gov.

[2] States submit service use and payment records to T-MSIS on a rolling basis after they are adjudicated and paid. One important implication of this is that an individual service may be represented by multiple records across different state T-MSIS submissions, in the form of original, void, or replacement claims.

[3] States have significant latitude in how they deliver and pay for covered services. They may reimburse health care providers for services delivered to Medicaid and CHIP beneficiaries by paying directly for each covered service on a fee-for-service (FFS) basis or by paying a flat monthly payment per beneficiary for a contracted set of services to another entity—such as a managed care plan—which then assumes responsibility for delivering care to the beneficiary. States also expend Medicaid funds on the following: Medicare Part A and Part B premiums for beneficiaries who are dually eligible for Medicare; premium assistance to enroll certain Medicaid beneficiaries into private coverage; supplemental payments above the standard fee schedule or other standard payment; supplemental lump sum payments to hospitals and other providers that are not tied to an individual service (often referred to as service tracking payments); and flat fees for providers for primary care case management. States report all of these service use and payment records to CMS in T-MSIS.

[4] For more information about TAF, see the T-MSIS Analytic Files website: https://www.medicaid.gov/medicaid/data-systems/macbis/medicaid-chip-research-files/transformed-medicaid-statistical-information-system-t-msis-analytic-files-taf/index.html

[5] More information on TAF production is available at https://www.medicaid.gov/dq-atlas/downloads/supplemental/9010_Production_of_TAF_RIF.pdf

[6] During the transformation into RIFs, some TAF data elements are suppressed, changed, or renamed. For more details on the difference between the pre-RIF and RIF version of the TAF data, including a crosswalk of variable names, see "Production of the TAF Research Identifiable Files (RIFs)," available in the Resources section of *DQ Atlas*.

inpatient file (IP), the long-term care file (LT), the pharmacy file (RX), and the other services file (OT)—are the focus of this technical documentation. Institutional inpatient services and payments are captured in the IP file; institutional long-term care services and payments are captured in the LT file; all other medical services and payments are captured in the OT file; and prescription drug fills and pharmacy payments are captured in the RX file (Table 1).

Table 1. TAF claims files

File	Type of claims, encounters, and payment records included
IP	• Hospital inpatient facility
LT	• Nursing facilities • Intermediate care facilities for individuals with intellectual disabilities • Mental health facilities • Independent (free-standing) psychiatric wings of acute care hospitals
RX	• Prescribed drugs and over-the-counter drugs filled at a pharmacy • Durable medical equipment (select types) • Other services provided by a pharmacy
OT	• All medical claims, encounters, and payments not captured in the IP or the LT files, including but not limited to: • Physician services (provided in inpatient and outpatient settings) • Outpatient hospital services • Dental services • Other professional services • Clinic services • Laboratory services • X-ray services • Sterilizations • Home health services • Personal support services • Durable medical equipment (select types) • Managed care capitation payments

II. Records in the TAF claims files

States submit service use and payment records into T-MSIS formatted as one header record and one or more line records that link to the header. Header records include summary information about the claim as a whole, whereas line records include detailed information about the individual goods and services billed as part of the claim.

CMS creates the TAF claims files by selecting the active T-MSIS header records that represent final-action claims. To do this, all versions of a claim must be grouped together into a "claim family," and the final-action algorithm selects the version that represents the final-action claim.[7] Next, all claim lines

[7] For more information about states in which the final-action algorithm cannot reliably identify final-action claims, see "Final Action Status in T-MSIS Claims," available in the Resources section of *DQ Atlas*.

affiliated with the selected final-action headers are selected for the TAF.[8,9] In some cases, final-action header records and their associated lines are excluded from the TAF when, for example:

- The final-action header represents a voided claim
- The final-action header represents a fully denied claim
- The header is an apparent duplicate of other header records in the claims family[10]

After records have been selected for the TAF, they are grouped into monthly files based on the service date (see Table 2 for which data elements are used to assign records to monthly files). When a service date cannot be determined because there are no data in the relevant date fields, the record is dropped from the TAF.

Table 2. TAF inclusion criteria

File	Inclusion criteria for monthly file
IP	Month/year of the discharge date; or, when the discharge date is unavailable, the most recent service end date on the claim line; or when the service end date among the claim lines is missing, the most recent service begin date on the claim line
LT	Month/year of the service end date
RX	Month/year of the prescription fill date
OT	Month/year of the service end date on the header; or, when the service end date on the header is missing, the service begin date on the header; or, when the service begin date is missing, the most recent service end date on the claim line

Expected variation in claims volume. States may choose the populations and benefit categories they cover, and as a result, their Medicaid and CHIP programs vary in the characteristics of their covered populations and in their benefit packages. The number of claims and average claims volume per beneficiary in the TAF claims files are therefore likely to vary somewhat by state. This is particularly true in the LT file, reflecting not only significant differences in how states instruct providers to bill for these

[8] Because of limitations in its claims processing system, Illinois captures adjustments to original claims as incremental credits or debits rather than voiding the original claim and submitting a replacement record with the new payment amount. As a result, the version of a record with the latest adjudication date may not represent the final action claim as it does in all other states. To ensure the TAF correctly captures all expenditures reported by Illinois into T-MSIS, all service use records are included in the IP, OT, LT, and RX files. This means that in some cases, the TAF will include multiple versions of a single claim for Illinois, so including all records in an analysis will overcount service utilization. For more information, see the technical documentation, "How to Use Illinois Claims Data," available in the Resources section of *DQ Atlas*.

[9] The TAF retains all lines associated with a header, including denied line-level records associated with a non-denied header claim. Denied line records can be identified by one of the following claims status codes (CLL_STUS_CD): 542, 585, and 654.

[10] T-MSIS claims are uniquely identified by a set of five data elements reported by the states, collectively called the "record key." The TAF variable names for these data elements include: SUBMTG_STATE_CD, ICN_ORIG, ICN_ADJ, ADJDCTN_DT, and LINE_ADJSTMT_IND. In some cases, states submit records to T-MSIS that have the same information in all five data elements; such records are considered duplicates even if other fields in the records have unique information. Because the TAF selection criteria for claims incorporates only the five data elements that make up the record-key, the TAF algorithm cannot identify which claim lines belongs to which claim header among a set of duplicate claim headers (that have identical record key data) should be selected. As a result, the TAF algorithm excludes all T-MSIS records that have identical record-key data.

services but also the extent to which state Medicaid programs have shifted toward providing home and community-based services for beneficiaries who need long-term services and supports.

Variation in claims volume that indicates a concern about data quality. Although some variation in claims volume is expected because of state-level variation in Medicaid eligibility and benefits, an extremely low or high volume of claims relative to most other states may indicate data completeness or quality issues in T-MSIS or in the TAF. States with an extremely low volume of claims are typically missing information in the T-MSIS data that leads those records to be excluded from the TAF. In some cases, the missing data are so extensive that a state's TAF claims file is unusable.

An unusually high volume of claims may indicate that a state is submitting certain types of claims in the wrong file, or it may be submitting line-level records as header records. TAF users may be able to correct for these data quality issues in their analyses.[11] For example, if a state has an unusually high volume of records reported in its IP or LT files, TAF users should consider using data elements such as type of bill to identify non-inpatient or non-long term care claims, respectively, and remove them from analyses of these files.[12]

States with multiple T-MSIS reporting entities. In some cases, more than one agency in a state reports eligibility and claims data to T-MSIS, and the data from each reporting entity have separate submitting state codes. As of 2021, four states have multiple reporting entities: Wyoming and Wyoming CHIP (SUBMTG_STATE_CD 56 and 93, respectively), Montana and Montana Third-Party Administrator (TPA) (SUBMTG_STATE_CD 30 and 94, respectively), Iowa and Iowa CHIP (SUBMTG_STATE_CD 19 and 96, respectively), and Pennsylvania and Pennsylvania CHIP (SUBMTG_STATE_CD 42 and 97, respectively). As part of the production of the TAF RIF, records from different data submitters in the same state are assigned the same state code. However, users of the non-RIF versions of TAF should make sure to include records with both submitting state codes for analyses of all Medicaid and CHIP beneficiaries in those states.

III. Linking header-level and line-level records

Each claim in the TAF is structured to capture one header record and one or more line records.[13] Header-level records capture data that apply to the entire claim. Line-level records capture data about the specific goods or services provided to a beneficiary as part of the overall service. A full claim record may

[11] For more information about states with an unusually low or high claim volume of claims for each file type, see "Claims Volume - IP," "Claims Volume - LT," Claims Volume - OT," and "Claims Volume - RX" in the Explore by Topic section of *DQ Atlas*.

[12] Header records in the IP file are expected to have type of bill values of 011*x*, 012*x*, or 085*x*, which correspond to inpatient hospital or critical access hospital services. Records with other type of bill values may represent outpatient hospital or outpatient facility records that a state erroneously included it its T-MSIS IP submissions. For information on the type of bill values expected in each claims file, see the Background and Methods materials in "Type of Bill - IP" in the Explore by Topic section of *DQ Atlas*.

[13] The TAF production algorithm includes final-action claim headers and all their associated line records but only for T-MSIS claim lines that can be linked to a T-MSIS claim header. If any of the variables that are used to link a claim line to a claim header are different, the claim line will become an "orphan" claim line, and the TAF excludes "orphan" claim lines.

have one claim header and many claim lines, or in the case of payment records, one claim header and no claim lines.

Similar to T-MSIS, the TAF organizes the claim header records in one file and the claim line records in a separate file. TAF header and line records must be linked to obtain complete information about a claim or payment record. They should be linked by using the TAF data analytic run ID (**DA_RUN_ID)**, which identifies the TAF production run that produced the TAF file, and the linking variable ***_LINK_KEY**, where * represents the claims file of interest (IP, LT, RX, or OT). The link key variables include the file version, the year/month of the file, the submitting state code, the original internal control number (ICN), the adjustment ICN, the adjudication date, and the adjustment indicator.

IV. Linking claim and eligibility records

Some analyses may require linking claim records to the eligibility record that represents the beneficiary who received the service. To do so, TAF users should first link line-level claim records to header-level records as detailed above, and then link the header record to an eligibility record in the TAF annual Demographic and Eligibility (DE) file. Header claim records and DE records should be linked using the submitting state code (**SUBMTG_STATE_CD**), unique beneficiary identifier[14] (**MSIS_IDENT_NUM** or **BENE_ID**), and the file year **(*_FIL_DT**, where * represents the file being linked such as DE, IP, OT, LT, or RX).

Not all records in the IP, OT, LT, and RX files can be linked to eligibility records. In many states, certain types of nonclaim financial transactions—including Disproportionate Share Hospital (DSH) or Upper Payment Limit (UPL) demonstration payments—are calculated and paid based on the total number of services delivered to all low-income or Medicaid-eligible individuals over a period of time. These payments are not expected to link to individual eligibility records. In other cases, many similar services delivered to a group of Medicaid beneficiaries (for instance, at a school-based vaccination event) may be billed on a single claim that cannot be linked to individual eligibility records. A small number of states may submit managed care capitation payments that cover many beneficiaries in a single T-MSIS record, and when this occurs the record cannot be linked to individual eligibility records. Finally, in some cases data quality issues may prevent the linkage of claims and eligibility records in TAF.[15]

As part of the TAF production process, dummy records are added to the DE file that represent unique beneficiary identifiers appearing on claims but for which states did not submit any eligibility information. If TAF users wish to remove these dummy records before linking claims and eligibility records, in order to

[14] The MSIS ID is the state-assigned unique beneficiary identifier present in the T-MSIS data submitted by states, while the BENE ID is the federally assigned unique beneficiary identifier that is added to the TAF RIF to allow linkage with Medicare data for dually eligible beneficiaries. For more information on these data elements, see “Unique Beneficiary Identifiers in the TAF RIF” in the Resources section of *DQ Atlas*.

[15] For more information on states with data quality problems that prevent accurate linkage of service use and payment records to individual beneficiaries, see the “Linking Claims to Beneficiaries” and “Linking Expenditures to Beneficiaries” topics in *DQ Atlas*.

retain only claims that match to beneficiaries reported by the state as being enrolled, they should exclude DE records where the missing eligibility data indicator (MSG_ELGBLTY_DATA_IND) is set equal to 1.

V. Identifying different types of records in the claims files

The TAF claims files include a wider variety of record types than were present in the predecessor MAX files, including both claims-based service use records and non-claims-based financial transactions. Service use records consist of both FFS claims, which represent billed services processed and paid by the state Medicaid or CHIP agency, as well as managed care encounters, which represent billed services processed and paid by a managed care organization. Non-claims-based financial transactions include payments that can be linked to individual beneficiaries, known as capitation payments, and payments that cannot be linked to individual beneficiaries, known as service tracking claims.[16]

Finally, states submit some "supplemental payments" that represent payments beyond the regular negotiated payment rate. For example, states might make payments, known as "wrap-around payments" to federally qualified health centers to bring the payment received from managed care plans for a visit up to the minimum level required by law under Medicaid. In some cases, supplemental payments include the same information present in other service use records, such as diagnosis codes and procedure codes. In other cases, supplemental payments are coded more like non-claim-based financial transactions.

For many analyses, TAF users will want to include or exclude certain types of records. Table 3 shows the types of records that should be used for different kinds of analyses.

Table 3. Records to use in various analyses

Analyses	Claim types to include	Claim types to exclude
Analyses of service use or population health	FFS claims Managed care encounters	Capitation payments Other non–claims-based financial transactions[a]
Analyses of state Medicaid or CHIP expenditures	FFS claims Capitation payments Supplemental payments Non–claims-based financial transactions[a]	Managed care encounters

[a] Most non–claims-based financial transactions that are not capitation payments are reported with a claim type code for service tracking or supplemental payments. Service tracking claims are available in the TAF and the TAF RIFs starting with Release 2 of the 2017–2018 TAF RIFs. They were excluded from earlier versions—specifically, the 2014–2016 TAF RIFs and Release 1 of the 2017–2018 TAF RIFs.

[16] Most per-member per-month payments that states make on behalf of individual beneficiaries, including managed care capitation payments, primary care case management fees, and premiums for private plans or Medicare, are reported with a record type of "capitation payment." However, some states report lump-sum payments or adjustments to Medicaid managed care plans as service tracking claims that cover multiple beneficiaries in a single transaction. Other types of payments represented as service tracking claims in the TAF include Disproportionate Share (DSH) payments to hospitals, supplemental payments made under the Upper Payment Limit (UPL) program, and certain services provided to a group of eligible beneficiaries and bulk billed.

TAF users have two options for selecting relevant records for their analysis: using either the federally assigned service category (FASC) or the claim type code. Both data elements are present on all header records in the IP, LT, OT, and RX files.

The FASC variable **(FED_SRVC_CTGRY_CD)** is available for all years and versions of TAF files that are produced in 2022 or later, starting with Release 1 of the 2020 TAF RIF.[17] The variable assigns IP, LT, OT, and RX header records into 1 of 21 distinct service categories (Table 4). Using the FASC code is the preferred method for differentiating between non–claims-based financial transactions and claims-based service use records, because it allows users to identify capitation payments to managed care plans more consistently across states than the claim type code does. The FASC also provides information on the type of claim form that was likely used in submitting the claim, which can help users determine which data elements are expected to be populated on the record.[18]

Table 4. FASC values

Type of record	Federally assigned service categories
Nonclaim financial transactions	• Managed care capitation payments • Other per-member per-month payments • DSH payments • Other financial transactions
Service use (submitted on institutional claim forms)	• Inpatient hospital • Nursing facility • Intermediate care facility • All other overnight facilities • Hospice • Outpatient facility • Clinic • All other outpatient institutional claims
Service use (submitted on noninstitutional claim form)	• Radiology • Laboratory • Home health • Transportation services • Dental • Home and community-based services, not otherwise specified • Durable medical equipment and supplies • Physician and all other professional claims
Service use (prescription drug)	• Prescription drug

[17] For more information on the FASC code, see the methodology brief "Assigning TAF Records to a Federally Assigned Service Category," available in the Resources section of *DQ Atlas*.

[18] For instance, type of bill and revenue center codes are available only on institutional claim forms. Place of service is available only on professional claim forms. Prescription drug fills are submitted on claim forms that do not include diagnosis codes.

The claim type code **(CLM_TYPE_CD)** is an alternative method for differentiating among the types of records in the TAF claims files.[19] It is also the only data element that can be used to differentiate between FFS claims that are processed and paid by the state and managed care encounters that are processed and paid by a Medicaid managed care plan. Table 5 shows the claim type code values that should be used to identify specific types of records.

Table 5. Values for claims type code (CLM_TYPE_CD), by program type and record type

Record type	Definition	Claim type code values by program		
		Medicaid or Medicaid-expansion CHIP	Separate CHIP	Other
FFS	Claims paid by the state to the provider for services rendered	1	A	U
Capitated payment	Fixed per beneficiary per month payments made by states on behalf of individually identified Medicaid and CHIP beneficiaries[a]	2	B	V
Managed care encounter	Records submitted by managed care organizations that represent claims submitted by providers to a managed care organization	3	C	W
Service tracking claim[b]	Lump sum payments to providers that cannot be attributed to a specific Medicaid or CHIP beneficiary[c]	4	D	X
Supplemental payment	Payments above the capitation fee or set payment rate for services provided to a specific Medicaid or CHIP beneficiary[d]	5	E	Y

[a] Payments with claim types 2, B, and V include those made to Medicaid managed care plans, to providers for primary care case management, to Medicare for premium payments on behalf of dually eligible beneficiaries, and to other private plans for premium assistance.
[b] Service tracking claims are available in the TAF and the TAF RIFs starting with Release 2 of the 2017–2018 TAF RIFs. They were excluded from earlier versions—specifically, the 2014–2016 TAF RIFs and Release 1 of the 2017–2018 TAF RIFs.
[c] The Medicaid identification number (MSIS_IDENT_NUM) on service tracking claims usually begins with "&" to indicate that the payment record cannot be linked to a specific beneficiary identifier.
[d] Supplemental payments can be attributed to a specific person but not always to a specific service.

In some cases, TAF users may want to restrict their analyses to claims covered under a specific program (for example, they may need to exclude S-CHIP claims). The claim type code (CLM_TYPE_CD) can also be used to do this. TAF users should also note that the claim type code includes values (U, V, W, X, and Y) for "other" types of records that the state did not classify as either Medicaid or CHIP payment records. These records may or may not represent services that qualify for federal matching funds under Title XIX

[19] For more information about the quality of the claim type code variable, see "Supplemental Payments" and "Non-Program (Other) Claims" in the Explore by Topic section of *DQ Atlas*.

or Title XXI. As states appear to use these records for different reasons, TAF users will need to decide on a state-by-state basis whether to include the records in their analyses.[20]

The claim type code cannot be used to identify claims that are partly covered by Medicare. TAF users who want to identify such claims should use the crossover claims indicator variable (XOVR_IND). The crossover claims indicator variable is useful when conducting analyses of beneficiaries who are dually eligible for Medicare and Medicaid.

VI. Identifying and using managed care encounter data

The claim type code (CLM_TYPE_CD) is one of the best ways to identify managed care encounter records. However, for more detailed information on the entity that adjudicated and paid for services delivered under managed care, TAF users can also rely on the managed care plan ID number (MC_PLAN_ID), which is a unique, state-assigned number that represents the health plan under which the service was provided. The managed care plan ID on encounter records can be linked to eligibility records through the annual DE managed care supplemental file—or it can be linked to records in the annual managed care plan file—to find out more about the characteristics of the plan.

TAF users might want to identify all managed care encounters for a specific type of managed care, such as all services provided by behavioral health plans in the state. To do so, they must link the managed care plan ID in the encounter record with the same plan ID in the annual managed care plan file to determine the managed care plan type associated with the plan ID.

The quality of managed care encounter data is generally highest in the IP and OT files and lowest in the LT file.[21] TAF users should be mindful that some states do not require comprehensive managed care plans to cover services that would be included in the LT or RX files and might therefore have few or no encounters in these files. Similarly, the services that states require behavioral health organizations to cover vary considerably by state. Therefore, plans or states with no encounters in certain files may not necessarily indicate a data quality issue.

VII. Identifying service use and payment records for specific Medicaid programs

The TAF claims files include two variables that indicate whether services were covered under a specific Medicaid program or a state waiver program. TAF users interested in identifying claims related to types of Medicaid programs should rely on the program type code (**PGM_TYPE_CD**) or the waiver type code (**WVR_TYPE_CD**) as follows:

[20] For more information about the states that use the “other” claim type code, see “Non-Program (Other) Claims” in the Explore by Topic section of *DQ Atlas*.

[21] For more information on the quality of encounter data for comprehensive managed care plans, see “CMC Plan Encounters - IP,” “CMC Plan Encounters - LT,” “CMC Plan Encounters - OT,” and “CMC Plan Encounters - RX” in the Explore by Topic section of *DQ Atlas*.

- **PGM_TYPE_CD** indicates whether a service was provided under a special Medicaid program, such as Indian Health Services or Money Follows the Person.
- **WVR_TYPE_CD** specifies the waiver type under which an eligible individual is covered and received services. These waivers are 1915(b), 1915(c), 1115, and 1332.[22]

VIII. Identifying service use and payment records for specific time frames

TAF users may be interested in identifying claims and encounter records for services that occurred during a specific time frame. To do so, users can subset files by date by relying on different date variables, depending on the claim file and whether claim header-level or line-level variables are required for the analysis. The IP and LT header-level records include the admission date (**ADMSN_DT**), which represents the date on which the beneficiary was admitted to the facility, and the discharge date (**DSCHRG_DT**), which represents the date on which the beneficiary was discharged from the facility.[23] The OT header-level records include service beginning dates (**SRVC_BGNNG_DT**) and service ending dates (**SRVC_ENDG_DT**).

Some states are known to incorrectly report discharge dates on LT header records even when beneficiaries are not discharged and remain at the facility. In 2016, as many as 15 states had a non-missing discharge date on more than 98 percent of their LT headers, even though these claims often represent weekly, biweekly, or monthly interim bills for extended stays in long-term care facilities that should not have a discharge date. TAF users should be cautious in relying solely on the admission and discharge dates reported on individual LT headers to determine the number of stays or average length of stay in long-term care facilities, as these dates may not be accurate.[24]

The IP, LT, and OT line-level records include service beginning dates (**SRVC_BGNNG_DT)** and service ending dates (**SRVC_ENDG_DT)** for the specific service represented by the line. If a service is received during a single visit with a provider, the service beginning and ending dates will generally be the same. If a service involved multiple visits on different days, or if the period of care extended for two or more days, then the service beginning date is the date on which the service covered by the claim began, and the service ending date is the date on which the service covered by the claim ended.

To account for missing service dates, the IP and OT header-level records also include a derived service ending date (**SRVC_ENDG_DT_DRVD**). In the IP file, the value for this derived variable is set to the discharge date if it is available. If the discharge date is missing, the derived service ending date is then

[22] For more information about how well the number of 1915(c) participants compare to the CMS 372 report, see "1915(c) Participation" in the Explore by Topic section of *DQ Atlas*. There are additional data elements in the TAF Annual Demographic and Eligibility file to indicate whether a beneficiary received coverage through other state plan options including 1915(a), 1915(i), 1915(j), and Community First Choice/1915(k).

[23] For more information about the validity of admission and discharge dates, see "Admission Date - IP," "Admission Date - LT," "Discharge Date - IP," and "Discharge Date - LT" in the Explore by Topic section of *DQ Atlas*.

[24] TAF users can review the recommended method for identifying unique inpatient stays in Section IX.F of this documentation. TAF users who wish to analyze long-term care stays can apply this approach to the claims in the LT file.

set to the most recent service ending date among the corresponding claim line records. If the service ending date is missing on all claim lines, the derived servicing ending date is then set to the most recent service start date among the corresponding claim line records. In the OT file, the derived service ending date is set to the service ending date from the claim header if it is available. If the service ending date is missing, the derived service ending date is then set to the service start date from the claim header. If both the service ending date and the service start date are missing from the claim header, the derived service ending date is then set to the most recent service ending date among the corresponding claim lines. To obtain the most complete date information available the TAF, users can use the derived service ending date in place of the discharge date in the IP file and the service ending date in the OT file to assign service dates to records.

In the RX file, the header-level fill date (**RX_FILL_DT**) represents the date on which a drug, device, or supply was dispensed by a provider. TAF users who want to approximate the time frame during which a beneficiary received drug treatment can use a combination of the fill date (**RX_FILL_DT)** and the number of days' supply dispensed by the provider (**SUPLY_DAYS_CNT**). For example, if a beneficiary was dispensed a seven-day (**SUPLY_DAYS_CNT**) supply of a particular drug on January 1, 2016 (**RX_FILL_DT**), a TAF user may choose to assume that the final day of treatment was on January 7, 2016. Note that this method is a proxy for identifying the time frame, and it relies on the assumption that a beneficiary took the medication as prescribed, which may not be correct.

IX. Classifying and analyzing service use records

A. Distinguishing between institutional and professional claims

Providers submit all medical claims on either an institutional claim form or a professional claim form. Institutional claims are often referred to as "UB-04 claims" when submitted in paper form or as "837I claims" when submitted in electronic form. Professional claims are referred to as "CMS-1500 claims" when submitted in paper form or as "837P" when submitted in electronic form. Appendix Figures A.1 and A.2 show each form.

In general, institutions such as hospitals, nursing facilities, intermediate care facilities for individuals with intellectual or development disabilities, rehabilitation facilities, home health agencies, and clinics (including federally qualified health centers and rural health clinics) submit institutional claims. Physicians (both individual and groups), dentists, other clinical professionals, free-standing laboratories and outpatient facilities, ambulances, and suppliers of durable medical equipment submit professional claims. Patient visits to a facility, such as an inpatient hospital, generally result in both an institutional claim and a professional claim—for example, for the physician services.

It is important for TAF users to be able to distinguish between institutional and professional claims because the standardized fields in institutional and professional forms, and therefore the information available for each type of claim, differ slightly. Type of bill codes and revenue codes are only present on the institutional claim form, and these data elements should take a missing value in the TAF for claims submitted on professional or other non-institutional claim forms. In contrast, place of service is only present on the professional claim form, and this data element should take a missing value in the TAF for

claims submitted on an institutional claim form. Procedure codes are expected on every line of a noninstitutional claim, but they may be present or missing on some or all lines of an institutional claim.

The IP and LT files include only institutional claims. The OT file contains a mix of institutional and professional claims. The FASC code is the preferred method for differentiating between claims that were submitted on an institutional claim form and a noninstitutional claim form.[25]

B. Identifying claims and encounters for specific procedures or services

TAF users may want to identify records for specific procedures or services, such as claims for the treatment of a substance use disorder or for a Caesarean section. The most reliable way to identify claims for specific services is to use the procedure codes, which capture the CPT, HCPCS or ICD-10-PCS code that describes a service or good delivered by a provider to a beneficiary on the specified date of service. TAF users will need to create their own list of procedure codes that pertain to the specific service(s) they are investigating. To identify procedure codes for specific analyses, TAF users can obtain CPT codes from the American Medical Association, and they can access Level II alphanumeric HCPCS codes and ICD-10-PCS codes on the CMS website.[26]

TAF files produced in 2022 or later years, starting with Release 1 of the 2020 TAF RIF, include information on the Agency for Healthcare Research and Quality's (AHRQ) Clinical Classification Software for Services and Procedures (CCS–Services and Procedures) category to which the procedure code on the line record is assigned. The CCS categories can be used to roll up procedure codes to clinically meaningful categories for analytic purposes.[27] TAF line records that use state-specific procedure codes will not have CCS–Services and Procedures information, because these nonstandard codes are not included in the software.

Records in the IP file can have up to six procedure codes, which are captured on the header record. The principal header-level procedure code (**PRCDR_1_CD**) is intended to be used for definitive treatment, not for diagnostic or exploratory purposes. Hospitals can use the additional procedure codes (**PRCDR_2_CD – PRCDR_6_CD**) and the related data elements (**PRCDR_*_CD_DT, PRCDR_*_CD_IND**) to record additional procedures. Records in the OT file have one procedure code (**PRCDR_CD**) on each line record, which captures the service or good delivered by a provider to a beneficiary on a specified date of service. The LT and RX files do not have procedure codes.

Some states use state-specific procedure codes rather than the nationally recognized HCPCS, CPT, or ICD-10-PCS procedure codes. State-specific procedure codes can make it challenging for TAF users to

[25] For more information on the FASC code, see the methodology brief "Assigning TAF Records to a Federally Assigned Service Category," available in the Resources section of *DQ Atlas*.

[26] The HCPCS code set is available at https://www.cms.gov/Medicare/Coding/HCPCSReleaseCodeSets/index.html. The ICD-10 code set is available at https://www.cms.gov/Medicare/Coding/ICD10/index.html.

[27] For more information on the mapping of CPT and HCPCS codes to the CCS–Services and Procedures category, see https://www.hcup-us.ahrq.gov/toolssoftware/ccs_svcsproc/ccssvcproc.jsp.

identify the services received by beneficiaries in these states because there is no general catalogue or listing of these codes and their meaning.[28]

Some institutional claims may not have a procedure code. However, TAF users may be able to rely on the revenue code **(REV_CD)** to identify the general type of service associated with a line record on an institutional claim. For example, there are revenue codes that identify emergency room services, laboratory services, and services provided in intensive care units.

The state-assigned type of service code (**TOS_CD**) is available on the line records. This data element, which the states report in their T-MSIS claims records, is intended to map each service provided to a Medicaid or CHIP beneficiary to standardized service categories. Although most states submit valid type of service codes on nearly all records in the IP, LT, OT, and RX files, there is substantial variation across states in the frequency with which various types of service codes are used, suggesting that states differ in how they apply the type of service codes to the same type of record.[29] This is particularly apparent for services that meet the definition of both the older, broadly defined codes used in MSIS (such as inpatient hospital) as well as the more granular codes newly available in T-MSIS (such as critical access hospital inpatient services, which is a subset of inpatient hospital services). Many states appear to use the most broadly defined codes from among the set of applicable type of service codes, which makes it difficult or impossible to use the type of service code to make meaningful comparisons of the use of specific services across states. TAF users should be very cautious when using the type of service data element to systematically identify a specific service across different states without including information from other fields, such as procedure or revenue codes.

The FASC code **FED_SRVC_CTGRY_CD** is applied to header records during TAF production, using a consistent set of rules across all states that rely on data elements that providers include on the claim, such as bill type, revenue code, procedure code, and National Provider Identifier (NPI). The FASC code will identify a more comparable set of services across states than the type of service code, since the rules for classifying records are the same for all states. However, the FASC code includes only 21 distinct categories and is much less granular than the state-assigned TOS code, which means it may not be usable for all analyses.

C. Identifying services delivered by specific types of providers

TAF users may want to identify the providers or categories of providers who deliver or bill for services rendered to Medicaid and CHIP beneficiaries. A “provider” could be an institution, a group of individual clinicians, or an individual clinician. A record in the TAF could include information for up to six different providers affiliated with the service (for example, the servicing provider and the billing provider), depending on the file type. Table 6 shows the TAF provider variables and indicates the claims files in

[28] For more information about the accuracy of procedure code variables, see “Procedure Codes - IP,” “Procedure Codes - OT Professional,” and “Procedure Codes – OT Institutional” in the Explore by Topic section of *DQ Atlas*.

[29] For more information about how often the type of service code is missing on claim records, see “Type of Service - IP,” “Type of Service - LT,” “Type of Service - OT,” and “Type of Service - RX” in the Explore by Topic section of *DQ Atlas*.

which each variable appears. Not all provider variables are present in every TAF claims file. For example, the IP and LT files have variables for the admitting provider, whereas the OT and RX files do not.

The TAF FFS and encounter records contain limited information about the providers affiliated with the claim, such as provider identifiers and, for certain providers in a claim, classification information. TAF users can obtain a wide range of additional information on providers affiliated with a given service by linking claims or encounter records to the Annual Provider (APR) file.[30] The TAF APR offers more details on the characteristics, locations, classifications, affiliated groups, affiliated programs, licensing/accreditations, and (for facility providers) bed types for Medicaid- or CHIP-eligible providers, as well as other identifiers associated with providers.[31] Unlike the claims file, the APR is designed to capture all taxonomy codes applicable to a provider.[32]

TAF users who want to link claims files to the APR base file should do so using the file date year, submitting state code, and the state-assigned provider ID.[33] If linking claims to APR records using the state-assigned provider ID yields an inadequate linkage rate, TAF users could use the NPI to link the remaining records. However, users should exercise caution when doing so given the variation across states in the percentage of providers either missing an NPI (for some states, the APR has particularly high rates of providers missing an NPI) or having more than one NPI.[34]

The information in FFS and encounter records in the TAF claims files about providers who render and bill for the care provided to beneficiaries include several variables that describe each provider type. Two of these variables can be used to identify a unique provider across claims: (1) the NPI, which is the unique, 10-digit identification number that the National Plan and Provider Enumeration System (NPPES) assigns to each HIPAA-covered health care provider; and (2) the state-assigned unique identifier used in the state's Medicaid Management Information System. The variables for servicing and billing NPI providers are generally well-reported by the states but NPIs are not required for atypical providers such as those that offer taxi services, home and vehicle modifications, and respite services. TAF users should note,

[30] TAF users will need to determine whether the provider information in claims is sufficient for their analysis or whether linking to the APR is necessary. For example, for users who want to examine beneficiary receipt of services delivered by primary care providers, the provider information in the claims may be sufficient. Users who want to examine beneficiary receipt of services delivered by primary care providers by the provider's geographic location (for instance, to examine the adequacy of provider networks) would need to link to the APR.

[31] For more information about the completeness and usability of APR data elements, see "Facility/Group/Individual Code," "Facility Characteristics," "Group and Individual Providers— Classification Types," "Facilities—Classification Types," "Provider Location," and "National Provider Identifier - APR" in the Explore by Topic section of *DQ Atlas*.

[32] The APR file also contains providers who are authorized but did not bill for or render services during the year, as well as providers whose authorization is pending, denied, or terminated.

[33] For more information about how well the state-assigned provider identifiers in the TAF claims files can be linked with provider identifiers in the TAF APR, see "Linking Claims to Providers" in the Explore by Topic section of *DQ Atlas*.

[34] For more information about the completeness of the provider NPI on claims, see "Billing Provider NPI - IP," "Billing Provider NPI - LT," "Billing Provider NPI - OT," "Billing Provider NPI - RX," "Servicing Provider NPI - OT," "Prescribing Provider NPI - RX," and "Dispensing Provider NPI - RX" in the Explore by Topic section of *DQ Atlas*. For more information about the percentage of APR records with an NPI, see the "National Provider Identifier - APR" topic in *DQ Atlas*. For additional information about providers, see the NPPES NPI Registry.

however, that a small number of states are struggling to report provider NPIs, especially for dispensing providers and prescribing providers. In some states, provider NPIs are commonly available on FFS claims but not on managed care encounters, or vice versa. TAF users interested in using provider NPIs should assess the level of missingness for the specific claims needed for their analyses.

Table 6. Provider variables across the TAF files

Provider	Variable prefix	Description	IP	OT	LT	RX
Admitting	ADMTG_PRVDR_*	The provider, hospital, or other institution responsible for admitting a patient	X		X	
Billing	BLG_PRVDR_*	The entity responsible for billing for services	X	X	X	X
Dispensing	DSPNSNG_PD_PRVDR_*	The provider responsible for dispensing a prescription drug				X
Health Home	HH_PRVDR_*	A provider enrolled in a Health Home care model		X		
Operating	OPRTG_PRVDR_*	The provider who performed the surgical procedure	X			
Referring	RFRG_PRVDR_*	The provider who recommended the servicing provider to the patient	X	X	X	
Servicing	SRVCNG_PRVDR_*	The provider who delivers or completes a particular medical service or non-surgical procedure	X	X	X	X
Prescribing	PRSCRBNG_PRVDR_*	The provider who prescribed a drug, device, or supply				X
Provider under direction	PRVDR_UNDER_DRCTN_*	The provider who directed the care that another provider administered		X		
Provider under supervision	PRVDR_UNDER_SPRVSN_*	The provider who supervised another provider		X		

Beyond these identifiers, the TAF claims files have several other provider-specific variables that describe a provider, such as specialty code and taxonomy code. However, not every provider-specific variable is present for all providers. For example, every provider field has a corresponding variable that captures the provider NPI number (***_NPI_NUM**), but not every provider field has a corresponding provider specialty code (***_PRVDR_SPCLTY_CD**) or state-reported provider taxonomy code (***_TXNMY_CD**). In addition, the quality of provider-specific variables can vary significantly across states, providers, and files.[35] To supplement the existing provider information reported on claims, the TAF IP, OT, and LT files produced in 2022 and later years, starting with Release 1 of the 2020 TAF RIF, also include a constructed variable for billing and servicing providers that uses the NPI on a claim to pull the provider's primary taxonomy information directly from NPPES (***_NPPES_TXNMY_CD**).

D. Identifying services delivered in certain settings

TAF users may need to identify service use records for care delivered in a certain health care setting, such as an inpatient hospital or residential facility. Certain settings can be inferred from the file in which a

[35] For more information on the quality of the variables for the billing provider type, see "Billing Provider Type - IP," "Billing Provider Type - LT," "Billing Provider Type - OT," and "Billing Provider Type - RX" in the Explore by Topic section of *DQ Atlas*.

service use record is located. By design, all services in the IP file are delivered in an inpatient setting, all services in the LT file are delivered in a long-term care facility, and all services in the RX file are delivered by a pharmacy. However, the OT file includes claims for facility services delivered in an outpatient setting and for professional services delivered across all types of settings, including inpatient, outpatient, and long-term care facilities. As a result, to access all records from a particular setting type, TAF users will frequently need to combine records from the OT claims file with records from another claims file. For example, to summarize all hospital utilization or costs, TAF users will need records from both the IP and OT claims files.[36]

E. Identifying services for the diagnosis and treatment of certain conditions

TAF users may need to identify records and services for specific health conditions if, for example, they want to estimate the number of beneficiaries who received treatment for diabetes or to calculate the total cost of care for sepsis. To identify records for services pertaining to specific conditions, TAF users can rely on three groups of variables: (1) diagnosis codes, (2) procedure codes, and (3) NDC codes.

Diagnosis codes. Records in the IP file can have up to 12 diagnosis codes, those in the OT file can have up to 2 codes, and those in the LT file can have up to 5 codes (**DGNS_*_CD**). Records in the RX file do not have diagnosis codes, since diagnoses are not recorded on pharmacy claims.[37] Some states may not require providers to submit diagnosis codes on claims for certain types of services captured in the OT file, such as medical supplies, prosthetic equipment, or nonemergency medical transportation (NEMT) services, because the providers billing for these services may not be in the best position to know and record an accurate diagnosis for the beneficiary. In addition, many states do not require providers to submit diagnosis codes on dental claims. However, if a provider includes any diagnosis codes on those claims, CMS instructs states to pass them through to T-MSIS as reported on the claim, even if the diagnosis code is inaccurate or invalid. For this reason, TAF users may want to exercise caution in using diagnosis codes on certain types of OT claims even if they appear to be valid ICD-10 diagnosis codes.

TAF users will need to create their own diagnosis code lists that pertain to the specific condition(s) they are investigating. To identify diagnosis codes for specific analyses, TAF users can access ICD-10-CM codes on the CMS website.[38]

Procedure codes. Since procedures are sometimes condition-specific, TAF users may in some cases be able to rely on procedure codes to identify the treatment for specific conditions. Records in the IP file have up to six procedure codes (**PRCDR_*_CD**), which are located on the IP claim header. In the OT file there is one procedure code (**PRCDR_CD**) on each OT claim line.[39] Records in the LT and RX files do

[36] For more information on service setting, see “Place of Service” in the Explore by Topic section of *DQ Atlas*.

[37] For more information about the accuracy of diagnosis codes, see “Diagnosis Code - IP,” “Diagnosis Code - LT,” and “Diagnosis Code - OT” in the Explore by Topic section of *DQ Atlas*.

[38] The ICD-10 code set is available at https://www.cms.gov/Medicare/Coding/ICD10/index.html.

[39] For more information about the usability of procedure codes in the TAF, see “Procedure Codes - IP,” “Procedure Codes - OT Professional,” and “Procedure Codes – OT Institutional” in the Explore by Topic section of *DQ Atlas*.

not have procedure codes. TAF users will need to create their own list of procedure codes that pertain to the specific service(s) they are investigating. To identify procedure codes for specific analyses, TAF users can obtain CPT codes from the American Medical Association, and they can access Level II alphanumeric HCPCS codes and ICD-10-PCS codes on the CMS website.[40]

NDC codes. The 11-digit NDC code (**NDC_CD**) that is present on pharmacy claims indicates the drug, device, or medical supply covered by a claim. As with procedure codes, specific drugs, devices, and medical supplies may be condition-specific and could therefore be used to identify beneficiaries who receive treatment for specific conditions. The NDC code is on every claim in the RX file and on a small number of claims in the IP, OT, and LT files. TAF users will need to create their own NDC list that pertains to the specific condition(s) they are investigating. To identify codes, they can access the National Drug Code Directory on the Food and Drug Administration website.[41]

F. Identifying unique inpatient hospital stays

TAF users might need to identify unique inpatient hospital stays for research or to calculate quality measures. In general, the IP file should contain acute inpatient stays at general hospitals, and the LT file should contain nonacute stays as well as stays at non-general hospitals, such as specialty psychiatric hospitals.[42] In some cases, one inpatient stay will generate several claims, each with its own header and lines. Users must consolidate these claims to the stay level if they want to examine Medicaid and CHIP-funded inpatient services at the stay level.

The approach to identifying a stay involves rolling up multiple IP and LT claim records for the same beneficiary if the beneficiary's claim durations overlap or immediately follow one another, as indicated by the admission and discharge dates.

Because inpatient hospital stays might span more than one month, we recommend using at least 12 months of data when identifying and collapsing multiple-claim inpatient hospital stays.[43] Users should first merge the line-level claim files with header-level claim files, using (1) the maximum value of production data run ID (**DA_RUN_ID**) available in both files and (2) the same unique linking variable (**IP_LINK_KEY**). Next, users should retain only header-level records in which the claim type code

[40] The HCPCS code set is available at https://www.cms.gov/Medicare/Coding/HCPCSReleaseCodeSets/index.html. The ICD-10 code set is available at https://www.cms.gov/Medicare/Coding/ICD10/index.html.

[41] The National Drug Code Directory is available at https://www.fda.gov/drugs/drug-approvals-and-databases/national-drug-code-directory.

[42] In practice, there is variation in the file into which states report psychiatric and other specialty hospital stays. We therefore recommend using both the IP and LT file for any studies of acute-care stays that might sometimes be delivered at facilities other than general hospitals.

[43] Using 12 months of data could result in a slight overcount of inpatient hospital stays if claims have a discharge date in the measurement year but the constructed stay-level discharge date is in the following year.

indicates an FFS or encounter claim and the FASC code (FED_SRVC_CTGRY_CD) indicates an inpatient hospital stay.[44]

Complete start and end dates in the relevant claims are required to identify unique inpatient hospital stays. If available, TAF users should use the discharge date (**DSCHRG_DT**) at the header level to determine the end date of the claim. If the discharge date is missing, TAF users should assign the most recent service ending date (**SRVC_ENDG_DT**) as the discharge date. Similarly, TAF users should use the original admission date (**ADMSN_DT**) to establish the start date of the claim or, if that date is missing, use the earliest service beginning date (**SRVC_BGNNG_DT**) among the lines associated with the header record. Users should exclude header records and their associated lines if the header-level admission or discharge dates are missing and cannot be filled with line-level service beginning and ending dates.

IP and LT claims can be considered part of the same inpatient stay if they have (1) the same state code and beneficiary identification number (**SUBMTG_STATE_CD** and **MSIS_IDENT_NUM**) or the same federally assigned beneficiary identifier (**BENE_ID**) in the TAF RIF, (2) the same billing provider number,[45] and (3) service dates that represent continuous or overlapping periods according to one of the following conditions:

- The admission date or discharge date are the same.
- The admission and discharge dates of one claim completely overlap the admission and discharge dates of another claim.
- The admission date of one claim is equal to or one day after the discharge date of another claim, and the patient status (PTNT_STUS_CD) of the preceding claim is 30 (still a patient) or null (unknown).[46,47]

TAF users can then determine the characteristics of each stay, based on all linked claims it comprises. For example, users should select the earliest admission date among all linked claims included in the stay as the stay-level admission date, and the latest discharge date as the stay-level discharge date. TAF users can also bring diagnosis, procedure, or revenue codes from the claim level to the stay level.

[44] For more information on the FASC code, see the methodology brief "Assigning TAF Records to a Federally Assigned Service Category," available in the Resources section of *DQ Atlas.*

[45] TAF users should use the state-assigned billing provider number (BLG_PRVDR_NUM)—or the billing provider NPI (BLG_PRVDR_NPI_NUM) if the state-assigned provider number is not available. Users should exclude header records and associated line records if both the billing provider number and NPI of the billing entity is missing.

[46] Contiguous claims, in which the admission date of one claim equals the discharge date of another claim, should only be consolidated into a unique inpatient stay if the preceding claim indicates the beneficiary remained a patient and was not discharged from the hospital (that is, if nothing in the preceding claims indicates the beneficiary was actually discharged).

[47] TAF users should count a transfer as a separate IP stay if the transfer resulted in a new billing provider number or if the patient was discharged during the transfer.

X. Assessing expenditures

Many TAF users will be interested in using the data to understand Medicaid and CHIP expenditures. TAF users should always be careful to use the claim type code (**CLM_TYPE_CD**) to subset to only those records in which the expenditure information is meaningful in the context of their analysis. In particular, TAF users should exclude managed care encounter records if they are examining payments made by state Medicaid and CHIP agencies, since the payment information on encounter records does not represent payments made by the state on behalf of beneficiaries. (For more information, see Section X.C, Costs to managed care organizations.)

The major types of expenditures captured in the TAF are described in the remainder of this section.

A. FFS expenditures

Payments made by state Medicaid agencies to providers can be identified by using the FFS claim type (**CLM_TYPE_CD** = A for Medicaid and Medicaid-expansion CHIP FFS claims, **CLM_TYPE_CD** = 1 for separate CHIP FFS claims)[48] and the Medicaid paid amount variables. Payment variables appear on both the header and line-level records in the TAF claims files. The header-level total Medicaid paid amount (**TOT_MDCD_PD_AMT**) represents the total Medicaid payment associated with the entire claim, whereas the line-level Medicaid paid amount (**MDCD_PD_AMT**) represents payments made for individual line items enumerated within the claim or payment record. The line-level Medicaid paid amounts should always sum to the header total Medicaid paid amount, although in some cases this does not occur. Payment information on the line and header records is most likely to be consistent on FFS claims that are processed and paid at the line level, which includes many of the services in the OT and RX files. When the payment information is consistent, TAF users can elect to sum payments from either line or header records (but should not use both). In contrast, payment information on FFS claims that are processed and paid at the claim header level—including most IP and LT services in many states—is more likely to be inconsistent.[49] In cases where the sum of the line-level payments does not equal the header-level payment, TAF users may consider using the payment level indicator (**PYMT_LVL_IND**) to decide which payment amount to use. States use this indicator to document whether the claim was processed and paid at the header or the line level. When header- and line-level payment amounts are inconsistent, TAF users

[48] Claim type codes with values of “U” represent FFS claims for “other” types of records that the state did not classify as either Medicaid or CHIP payment records. These records may or may not represent services that qualify for federal matching funds under Title XIX or Title XXI. As states appear to use these records for different reasons, TAF users will need to decide on a state-by-state basis whether to include the records in their analyses.

[49] Long-term care facilities’ payment structures often do not easily translate to claim lines. For example, nursing facilities’ payment policies usually involve per diem rates, as well as adjustments that reflect residents’ medical acuity, the facility’s peer group, and the burden of certain conditions that are expensive to treat. As a result, one common pattern in the TAF is positive payments on LT claim headers and zero or missing payments on LT claim lines, which suggests that some states do not successfully disaggregate payments to long-term care facilities across the lines of the claim. Similarly, many state Medicaid agencies have adopted fixed payment policies based on diagnosis-related groups for inpatient services; these are standardized payments designed to cover all services provided during an inpatient stay. This payment arrangement is difficult to disaggregate to the line level, which may be why some states report positive payment on IP claim headers but zero or missing payments on IP claim lines.

should conduct additional sensitivity analyses to understand how the selection of line or header payment amounts impacts the results of the analysis.

Some states also report wraparound payments on supplemental payment records (**CLM_TYPE_CD** = 5). These are considered FFS expenditures because they are paid by the state, related to a particular service, and can be linked to a specific beneficiary. However, these records do not represent the full payment received by the provider for the service.

TAF users who conduct analyses based on FFS data should know that FFS claims with high rates of missing, zero, or negative payment data may preclude them from being able to accurately analyze health care costs. In addition, file submissions with a high proportion of line- and header-level FFS payments that are inconsistent may indicate a data quality issue, and users should exercise caution when relying on payment data from these claims.[50]

When analyzing FFS costs, TAF users may want to assess the volume of claims records for the relevant state or states as a measure of FFS claim completeness in the TAF. States with implausibly high or low volume of claims may have data quality issues that preclude accurate analysis of health care costs.[51] Some states are known to have FFS expenditures in the TAF that do not align well with external benchmarks of state Medicaid program spending, and users should exercise caution when using FFS payment data in those states.[52] Because the data elements available to classify FFS claims in the TAF do not align perfectly with the service categories used in external benchmarks such as the CMS-64, a state may have high alignment in total Medicaid expenditures or total FFS expenditures, but low alignment in specific service categories (for example, inpatient expenditures or long-term care expenditures).[53] This pattern suggests that the TAF expenditure data are complete but other TAF data elements do not support accurate partitioning into the same service categories as used by the benchmark data.

B. Monthly beneficiary payments

Monthly beneficiary payments made by state Medicaid agencies can be identified by using the capitated payment claim type (**CLM_TYPE_CD** = 2 for Medicaid and Medicaid-expansion CHIP records,

[50] For more information about the quality of payment data in header- and line-level records, see “Missing Payment Data – FFS Claims,” “Payment Data Consistency - IP,” and “Payment Data Consistency - OT,” “Payment Data Consistency - LT,” and Payment Data Consistency - RX” in the Explore by Topic section of *DQ Atlas*.

[51] For more information on states that may have this issue, see “Claims Volume - IP,” “Claims Volume - LT,” Claims Volume - OT,” and “Claims Volume - RX” in the Explore by Topic section of *DQ Atlas*.

[52] For more information, see “Total FFS Expenditures,” “FFS Inpatient Expenditures,” “FFS Long-Term Care Expenditures,” “FFS Other Medical Expenditures,” and “FFS Prescription Drug Expenditures” in the Explore by Topic section of *DQ Atlas*.

[53] For more information on the differences between the TAF and CMS-64 data on Medicaid expenditures, see the methodology brief “Medicaid Expenditure Data: TAF and the CMS-64,” available in the Resources section of *DQ Atlas*.

CLM_TYPE_CD = B for separate CHIP records).[54] Some states also report monthly beneficiary payments that may or may not be linkable to individual beneficiaries using other claim type codes (**CLM_TYPE_CD** = 4 for service tracking claims or **CLM_TYPE_CD** = 5 for supplemental payments).[55] Users can identify monthly beneficiary payments across all of these claim type codes using the FASC code (**FED_SRVC_CTGRY_CD** = 11 for managed care capitation payments and **FED_SRVC_CTGRY_CD** = 12 for other per-member per-month payments).

These records represent a variety of monthly payments. When a state contracts directly with a managed care plan, the covered services may either be comprehensive benefits contracted from a Medicaid managed care organization or narrower sets of inpatient or outpatient services contracted from a prepaid health plan.[56] States also make other monthly payments on behalf of Medicaid beneficiaries, including the following: a flat fee paid to a primary care provider for primary care case management plan services; Medicare Part A and Part B premiums for beneficiaries who are dually eligible for Medicare; and in some cases, premium assistance for enrolling Medicaid beneficiaries into private coverage.[57]

Most states accurately report monthly payments made to managed care plans, especially comprehensive managed care plans. However, capitation payments in the TAF were deemed unusable in some states because they did not align well with a benchmark, and in general, most states do not appear to be reliably capturing payments for Medicare premiums made on behalf of dually eligible beneficiaries.[58]

C. Costs to managed care organizations

Under Medicaid managed care arrangements, providers bill managed care plans, which process and pay the claims. In turn, these plans submit to states the cost and service information as encounter records to be included in T-MSIS submissions to CMS. These managed care encounter records can be identified by using the managed care encounter claim type (**CLM_TYPE_CD** = 3 for Medicaid and Medicaid-expansion

[54] Claim type codes with values of "V" represent capitated payment records for "other" types of records that the state did not classify as either Medicaid or CHIP payment records. These records may or may not represent services that qualify for federal matching funds under Title XIX or Title XXI. As states appear to use these records for different reasons, TAF users will need to decide on a state-by-state basis whether to include the records in their analyses.

[55] TAF users may want to include service tracking claims or supplemental payment records in analyses of managed care expenditures if the claims have a type of service code that indicates an additional capitation payment (TOS_CD = 119, 120, 121, or 122) or "other" monthly payments (TOS_CD = 138, 143, or 144).

[56] Prepaid health plans often cover a specific type of service, such as behavioral health care or dental care.

[57] Some Medicaid beneficiaries may qualify for employer-based coverage, and if this coverage is less costly than enrolling the beneficiary in traditional Medicaid, states have the option of paying the premium on behalf of the eligible beneficiary. Some states have used other policy options, such as 1115 waiver demonstrations or the Basic Health Program, to enroll eligible beneficiaries into private plans available through the state or federal Health Insurance Exchange. Note: the services provided to beneficiaries with private coverage are processed through a private health insurer, and the claims may or may not be submitted to T-MSIS and therefore may not be in the TAF.

[58] For more information about the quality of TAF-based capitated payment data, see "Total Monthly Beneficiary Payments," "CMC Payments," "PHP Payments," "PCCM Fees," and "Premium Assistance Payments" in the Explore by Topic section of *DQ Atlas*. For more information on the differences between the TAF and CMS-64 benchmark data on Medicaid expenditures, see the methodology brief "Medicaid Expenditure Data: TAF and the CMS-64," available in the Resources section of *DQ Atlas*.

CHIP records, **CLM_TYPE_CD** = C for separate CHIP records).[59] The Medicaid payment information on encounter records does not represent the same type of payment information that is on FFS claims. Instead, the Medicaid payment amount on encounter records represent payments made by managed care entities to institutions and providers; it does not represent a Medicaid or CHIP payment by the state (as it does on FFS claims). States report their payments for managed care services in capitation records, which represent the per-beneficiary-per-month premium payment from state Medicaid agencies to managed care entities.

As of 2019, the Medicaid and CHIP Managed Care Final Rule requires states to report to T-MSIS the amount managed care entities pay to institutions and providers for services. Historically, however, these data were suppressed by many managed care organizations because they considered the information proprietary. The payment information on managed care encounters in the TAF is therefore often less complete than it is on FFS claims,[60] and the payment data is masked on managed care encounter records in the TAF RIF.

As with FFS claims, high rates of missing, zero, or negative payment data on encounter records may preclude TAF users from accurately analyzing health care costs. Provider payment information on managed care encounters is missing or unusable in several states. However, the missing information may be specific to individual managed care plans. If users are interested only in specific managed care plans in a state, they can limit claims to the managed care plans of interest by using **MC_PLAN_ID** and then reassess the rate of missing provider payment information at the plan level. If users have the flexibility to exclude certain plans, they could also reassess the missing payment information for each managed care plan to evaluate which plans have data that are sufficiently complete and otherwise accurate to include in their study.

When analyzing costs to managed care organizations, TAF users may want to assess the volume of encounter records for the relevant state or states. States with implausibly high or low volume of encounter records may have data quality issues that preclude accurate analysis of health care costs.[61]

D. Service tracking claims

Service tracking claims are available in the TAF and are included in the publicly available TAF RIF for years and versions starting with Release 2 of the 2017 and 2018 TAF RIFs.[62] Service tracking claims include payments for services rendered to groups of beneficiaries; they cannot be attributed to a specific beneficiary, provider, or visit (CMS 2019b). Disproportionate Share Hospital (DSH) payments, Upper

[59] Claim type codes with values of “W” represent encounter records for “other” types of records that the state did not classify as either Medicaid or CHIP payment records. These records may or may not represent services that qualify for federal matching funds under Title XIX or Title XXI. As states appear to use these records for different reasons, TAF users will need to decide on a state-by-state basis whether to include the records in their analyses.

[60] For more information about the quality of payment data in encounter records, see “Missing Payment Data—Encounters” in the Explore by Topic section of *DQ Atlas*.

[61] For more information, see “CMC Plan Encounters - IP,” “CMC Plan Encounters - LT,” “CMC Plan Encounters - OT,” and “CMC Plan Encounters - RX” in the Explore by Topic section of *DQ Atlas*.

[62] Service tracking claims are redacted from the 2014–2016 TAF RIFs and Release 1 of the 2017–2018 TAF RIFs.

Payment Limit (UPL) supplemental payments, cost-settlement payments, drug rebates, and other lump-sum payments are most often captured as service tracking claims in the TAF. Service tracking claims could also represent additional capitation payments.[63] Although nearly all states report these types of payments in the CMS-64 expenditure reporting used to draw down federal matching funds, less than two-thirds of states reported any service tracking claims in 2016, suggesting these types of payments tend to be less complete in the TAF than FFS claims and monthly beneficiary payments.[64]

The best way to identify service tracking claims is the claim type code (**CLM_TYPE_CD** = 4 for Medicaid and Medicaid-expansion CHIP service tracking claims, **CLM_TYPE_CD** = D for separate CHIP service tracking claims).[65] The Medicaid paid amount (**TOT_MDCD_PD_AMT**) on the header should be zero, and the service tracking payment amount (**SRVC_TRKNG_PYMT_AMT**) should have a non-zero value, although not all states are moving the Medicaid paid amount to the service tracking payment amount properly. In addition, the Medicaid identification number (**MSIS_IDENT_NUM**) on service tracking claims usually begins with "&" to indicate that the payment record cannot be linked to a specific beneficiary identifier, and the service tracking type should be a value other than zero.

E. Supplemental payments

Supplemental payment records represent payments made in addition to a capitation payment or negotiated rate for services provided to a specific beneficiary; the payment can be attributed to a specific person but not always to a specific service. Supplemental payments that can be linked to both a specific beneficiary and a particular service might be considered an FFS expenditure. The best way to identify supplemental payments is claim type code (**CLM_TYPE_CD** = 5 for Medicaid and Medicaid-expansion CHIP supplemental payments, **CLM_TYPE_CD** = E for separate CHIP supplemental payments).[66] The payment amount for these records should appear in the Medicaid paid amount (**TOT_MDCD_PD_AMT**) on the header record. For the few states that submit these records, supplemental payment records represent a mix of financial transactions and claims that include information on service use. TAF users might want to use the presence of information typically coded in claims forms (including diagnosis code,

[63] Service tracking claims with type of service code (TOS_CD) 119, 120, 121, or 122 indicate an additional capitation payment, and those with type of service code 138, 143, or 144 indicate "other" monthly payments.

[64] For more information on the differences between the TAF and CMS-64 data on Medicaid expenditures, see the methodology brief "Medicaid Expenditure Data: TAF and the CMS-64," available in the Resources section of *DQ Atlas*.

[65] Claim type codes with values of "X" represent service tracking claims for "other" types of records that the state did not classify as either Medicaid or CHIP payment records. These records may or may not represent services that qualify for federal matching funds under Title XIX or Title XXI. As states appear to use these records for different reasons, TAF users will need to decide on a state-by-state basis whether to include the records in their analyses.

[66] Claim type codes with values of "Y" represent supplemental payments for "other" types of records that the state did not classify as either Medicaid or CHIP payment records. These records may or may not represent services that qualify for federal matching funds under Title XIX or Title XXI. As states appear to use these records for different reasons, TAF users will need to decide on a state-by-state basis whether to include the records in their analyses. For more information on the states that report these "other" records, see "Non-Program (Other) Claims" in the Explore by Topic section of *DQ Atlas*.

procedure code, and NDC) to differentiate between supplemental payments that appear to be financial transactions and those that appear to be service use records.

All supplemental payment records are available in the TAF; however, the TAF RIF excludes supplemental payment records in which the Medicaid identification number begins with an “&” because they cannot be attributed to a specific beneficiary.

Some states and policy analysts refer to payments on service tracking claims as supplemental payments, but the service tracking claims in T-MSIS/TAF are specifically for lump sum payments and the supplemental payments are for the additional payments provided to a specific beneficiary.

References

Centers for Medicare & Medicaid Services. “April 2019 Medicaid and CHIP Enrollment Data Highlights.” Available at https://www.medicaid.gov/medicaid/program-information/medicaid-and-chip-enrollment-data/report-highlights/index.html. Accessed July 19, 2019.

Centers for Medicare & Medicaid Services. “CMS MACBIS T-MSIS Reporting Reminder: Payment Amounts for Service Tracking Claims.” n.d. Available at https://www.medicaid.gov/medicaid/data-and-systems/macbis/tmsis/tmsis-blog/88451. Accessed August 12, 2019.

Centers for Medicare & Medicaid Services. *DQ Atlas*. Available at https://www.medicaid.gov/dq-atlas/.

Appendix A: Revision History

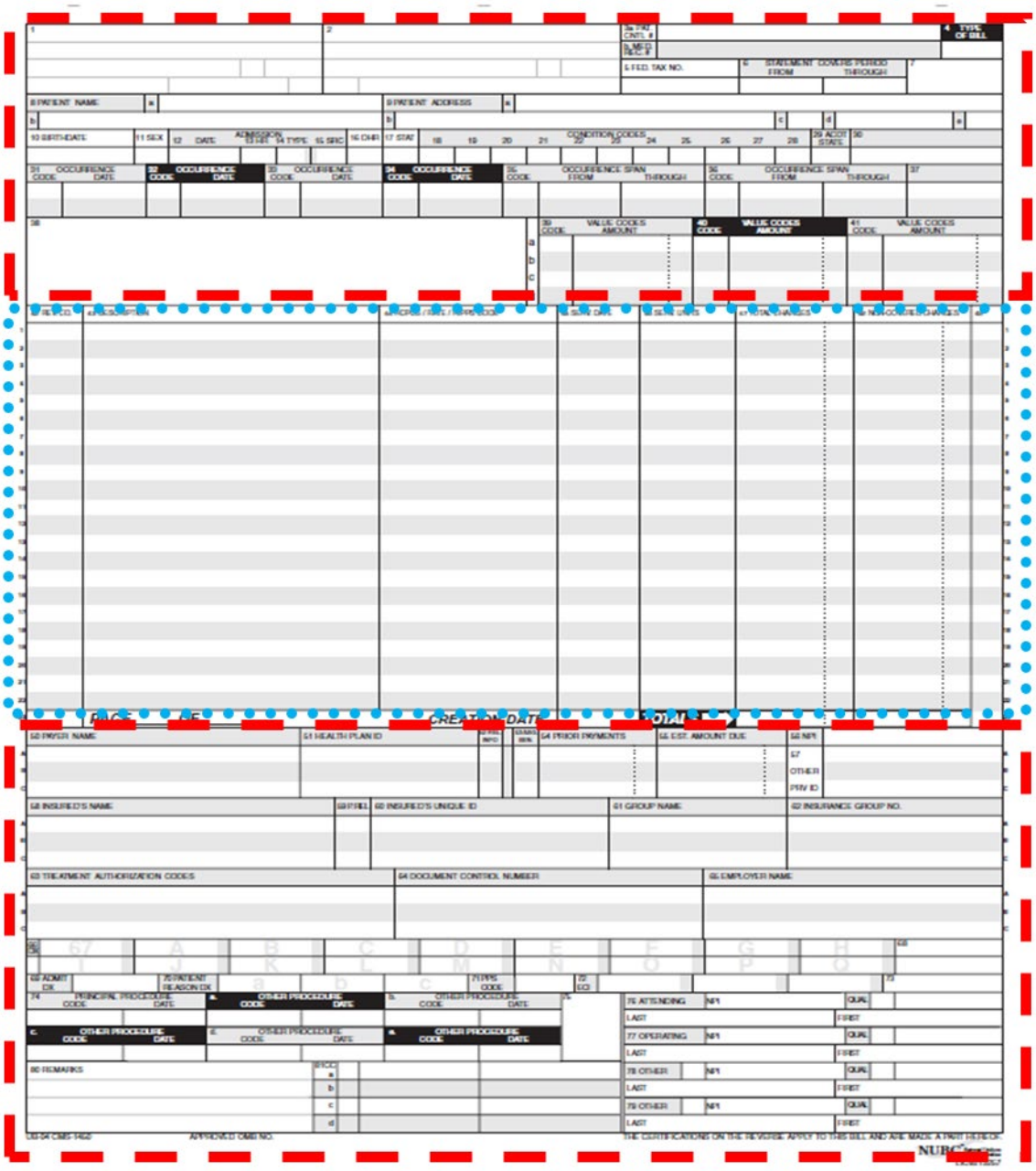
Table A.1. Revisions to technical documentation, by version

Version of documentation	Description of revision
August 2021	• Revised guidance on how to identify denied claim line records to use only claims status codes (CLL_STUS_CD) = 542, 585, and 654. Previously, this documentation incorrectly stated that claims status codes 026, 26, 087, and 87 could be used to identify denied line records. • Added information to account for new valid values of SUBMTG_STATE_CD for non-Medicaid entities from states that submit CHIP or TPA separately from Medicaid. • Clarified that service tracking claims are available in the TAF and TAF RIFs starting with 2017–2018 Release 2. Previously, this documentation said service tracking claims were not available in the TAF RIFs. • Added information about a new derived service ending date variable (SRVC_ENDG_DT_DRVD) in the IP and OT files, which became available in 2021. • Revised guidance on how to identify institutional claims and added information on how to identify professional claims in the OT file. • Added guidance on how to identify unique inpatient hospital stays. This method has been revised from the original recommendation issued in November 2019. In the revised method, (1) claims related to an inpatient stay are identified using revenue codes and type of bill codes instead of type of service codes, (2) billing provider NPI is used when the state-assigned billing provider number is missing, and (3) claims from the LT file are included along with those from the IP file to account for inpatient psychiatric stays. • Clarified that some states use supplemental payment claims (claim type code = 5 or E) to report FFS “wraparound” and monthly beneficiary payments. • Clarified that some states use service tracking claims (claim type code = 4 or D) to report monthly beneficiary payments. • Added information about how TAF users can link claims to the annual provider file to obtain a wider range of provider information.
June 2022	• Provided additional context about linking claims and eligibility records • Updated guidance for identifying different types of records in the claims files, unique inpatient stays, and monthly beneficiary payments to include use of the new FASC code, available in TAF files generated in 2022 or later years • Provided additional information about the new NPPES primary taxonomy variable in TAF for identifying services delivered by specific types of providers, available in TAF files generated in 2022 or later years.

Appendix B: Sample Claim Forms

Figure B.1. Sample institutional claim form (UB-04)

The information in the top and bottom sections of the form (surrounded by the red, dashed line border) is captured on the header-level claim record. The information in the middle section of the form (surrounded by the blue, dotted line border) is captured on the line-level claim records.



Source: CMS. "CMS-1450." 2019. Available at https://www.cms.gov/Regulations-and-Guidance/Legislation/PaperworkReductionActof1995/PRA-Listing-Items/CMS-1450. Accessed August 26, 2019.

Figure B.2. Sample professional claim form (CMS-1500)

The information in the top and bottom sections of the form (surrounded by the red, dashed line border) is captured on the header-level claim record. The information in the middle section of the form (surrounded by the blue, dotted line border) is captured on the line-level claim records.

Source: CMS. "CMS 1500." Available at https://www.cms.gov/Medicare/CMS-Forms/CMS-Forms/Downloads/CMS1500.pdf. Accessed August 26, 2019.

Anna Christensen[1], Andres Arguello[1], Lauren Hula[1], Linda Nguyen[1], Allison Barrett[1], Kimberly Proctor[2], and Jessie Parker[2]. “TAF Technical Documentation: Claims Files.” Baltimore, MD: CMS, 2021.

Reviewers: Laura Nolan[1]; Julia Baller[1]; Julie Sykes[1]; Carol V. Irvin[1]; Cara Stepanczuk[1]

[1]Mathematica, [2]Centers for Medicare & Medicaid Services, Center for Medicaid and CHIP Services, Data and Systems Group, Division of Business and Data Analytics

Correspondence should be addressed to DataConnectSupport@cms.hhs.gov