Medicare Data on Provider Practice and Specialty (MD-PPAS)
User Documentation

Version 2.4

Centers for Medicare & Medicaid Services

June 2021

Contents

Overview

CMS and ASPE have collaborated to update and distribute version 2.4 of the Medicare Data on Provider Practice and Specialty (MD-PPAS, pronounced "MD pass"), a data source that assigns Medicare providers to medical practices and elaborates on the CMS provider specialty classification system. The MD-PPAS are provider-level datasets built around two identifiers: the national provider identifier (NPI) and the tax identification numbers (TIN), which can be used to merge on other data.

The MD-PPAS version 2.4 includes annual files that span from 2008 to 2019. We used three CMS administrative data sets to generate this database:

- Medicare Provider Enrollment, Chain and Ownership System (PECOS) data extracted on June 2021 to obtain the birth date and specialty information for providers as well as the TIN legal name;
- National Plan and Provider Enumeration System (NPPES)[1] data extracted on June 2021 to obtain providers' names and sex; and
- Medicare fee-for-service (FFS) Part B non-institutional claims (100% file) to determine the pool of providers that are included in the MD-PPAS and identify the TINs that can be used to characterize provider group practices.[2] We used claims from the second half of 2008 (July-Dec) and the calendar years of 2009-2019 to capture TIN information.[3] We also use specialty information reported on claims to supplement missing PECOS data.

We chose to use PECOS specialty data rather than data from NPPES because providers must revalidate their PECOS data every five years, while NPPES requires providers to enter their information only once.

The MD-PPAS contains a record for any individual provider that had a valid NPI and submitted a Part B non-institutional claim for evaluation and management services, procedures, imaging, or non-laboratory testing with a positive allowed charges amount. Appendix A details how these sample criteria reduce the total provider count of the MD-PPAS relative to the provider count for the Part B non-institutional claims.

The creation of this database makes it easier for researchers to analyze several issues that involve TIN-based group practices or physician specialties.

[1] NPPES creates NPIs for providers. https://nppes.cms.hhs.gov/NPPES/Welcome.do

[2] These claims do not include physician services that are paid under the Optional Payment Method for Critical Access Hospitals (CAHs). Under this method, CAHs bill Medicare for both facility services and professional services rendered by a physician who has reassigned his or her billing rights to the CAH.

[3] We only used half the year of 2008 claims data because CMS required NPIs on all claims starting in May 2008.

Version Updates

Version 2.4 introduces the following changes from version 2.3

- The MD-PPAS data include a modified broad specialty variable that provides a more detailed breakdown of non-physician specialties.

- The data also include two additional variables related to the percent of providers' line items that are performed in a given place of service (POS). Specifically, MD-PPAS data now include variables that report the percent of line items performed in urgent care centers and retail clinics. Internal analyses revealed that providers affiliated with large retail clinics or urgent care centers listed the place of service as "office" as well as "retail clinic" or "urgent care center". Therefore, to help data users more accurately capture the POS for these providers, the data now report service lines associated with retail clinics and urgent care centers separately rather than collapsing them into the other POS category.

- The data use a new methodology to assign providers to CBSAs based on the zip code reported on their submitted claims. This change resulted in a small percentage of providers with different values for their CBSA code and type variable. Appendix F documents the percentage of providers with changes to these CBSA variables from version 2.3 to version 2.4.

- All previously available data years (2008 – 2017) were re-created to incorporate the above changes. As a result, data users will notice slight changes between versions 2.3 and 2.4 in PECOS-derived specialty and demographic data, because all data years now use the June 2021 PECOS extract to assign provider specialty.

- With the re-run of previous data years, there were negligible changes in the provider counts due to minor changes in the underlying claims data.

- Finally, the naming conventions for the MD-PPAS data sets were changed in order to conform to new internal documentation standards.

Version 2.3 introduces the following changes from version 2.2.

- The data now include variables on the percentage of a provider's line items that are performed in eight place-of-service categories.
- In 2017, PECOS introduced a new hospitalist specialty designation. To accommodate this change, version 2.3 changes the definition of hospitalists to include any provider who self-identifies as such in the PECOS data as well as primary care physicians who had at least 90% of their line items in an inpatient hospital setting.

- The data source for the part B non-institutional claims data changed the method by which it identifies final action claims. This modification can change the content of the claims data, e.g. procedure codes and allowed-charges dollar amounts. Consequently, there are some providers who no longer satisfy the criteria for inclusion in the MD-PPAS, leading the provider counts in the latest version of MD-PPAS to decline slightly.
- The data use a corrected form of the zip-code-to-CBSA crosswalk to assign providers to CBSA based on the zip code reported on their submitted claims. This correction resulted in a small percentage of providers that exist in both version 2.2 and version 2.3 with different values for their CBSA code and type variable.

Version 2.2 makes a single update to version 2.1.

- Providers are assigned to a Core Based Statistical Area using an updated zip code to CBSA crosswalk that corresponds to a given data year. Data from 2008 – 2015 have been updated using the new CBSA crosswalk tables.

Version 2.1 corrects two errors in version 2.0.

- Services are now limited to evaluation and management services, procedures, imaging, and non-laboratory tests. Version 2.0 inadvertently included Part B drugs (e.g., for chemotherapy) and excluded non-laboratory tests.
- A number of physicians (especially in large academic practices) billed using zip codes that represent an organization instead of a postal delivery area. As a result, those physicians were not assigned to a metropolitan area. This error has been corrected.

Version 2.0 updates.

- This version of the MD-PPAS includes the first, middle, and last names of providers to make it easier to identify individual providers without looking up NPIs online.
- It includes the number of part B non-institutional line items, total Medicare allowed charges, and the number of unique patients, overall and by TIN, to give researchers a comprehensive view of the amount of Medicare services that are associated with a provider.
- It also includes monthly indicators that specify whether each TIN listed in the MD-PPAS was reported in the providers' Part B non-institutional claims for a given month. The monthly TIN indicators will help researchers decide how to treat providers that report multiple TINs during the year. Specifically, the indicators would help them determine whether multiple TINs are due to providers switching practices during the year or maintaining concurrent positions at two different practices.
- The MD-PPAS now includes all non-physician practitioners as well as physicians and non-physician practitioners whose primary "state" is Puerto Rico. These additions to the file will give researchers the flexibility to use these populations in their analysis.
- Lastly, we changed the methodology used to assign providers to TINs, their primary geographic location and their designation as a hospitalist. Previously, we assigned providers who report

multiple TINs or geographic locations in the claims data to the two TINs and the geographic location that represents the plurality of their allowed charges. Similarly, we identified hospitalists as primary care physicians who had at least 90 percent of their allowed charges from services delivered in an inpatient setting. In the current version of the MD-PPAS, we assign TINs, locations, and the hospitalist specialty designation using line items rather than total allowed charges because the latter metric can be sensitive to low-frequency services with high charges.[4]

Database Description

The following sections describe the major data elements of the MD-PPAS. A detailed file layout is provided in Appendix B.

Practice Variables

Payers such as Medicare are required by the Internal Revenue Service (IRS) to report payments made to each TIN, which necessitates having the TIN of the practice on the claim. All providers who bill under the same TIN have a financial organization in common, but they could be practicing at different locations. For the MD-PPAS, the number of providers assigned to a TIN represents the group's size.

We faced the problem of how to assign providers to a group practice when they report more than one TIN in the claims data. The MD-PPAS capture the two TINs that reflect the largest percentages of their service line items (henceforth "line items") for evaluation and management visits, procedures, imaging services, or non-laboratory tests with positive allowed charges amounts.[5] The primary TIN reflects the largest percentage of line items and the secondary TIN reflects that second largest percentage of line items.

For the vast majority of providers, the top two TINs reflect the bulk of their Medicare service volume. In 2019, 92.8% of providers billed under a single TIN or under one dominant TIN at a time. Specifically, 81.5% of providers billed under one TIN; 7.0%% of providers billed under multiple TINs but one TIN was dominant in the sense of having at least 90% (but less than 100%) of their line items; and 4.4% of providers billed under two TINs that were jointly dominant (i.e., the TINs together accounted for at least 90% of providers' line items), and they switched TINs once in the year. Only 5.8% of providers had jointly dominant TINs, and they billed under multiple TINs in the same month two or more times throughout the year. The remaining 1.4% of providers had less than 90% captured by the top two TINs, and only 0.32% of providers had less than 75% of line items captured by the two TINs.

For individual providers who bill under multiple TINs in a given year, the TIN service-month variables (TIN1_SRVC_MONTH and TIN2_SRVC_MONTH) are designed to help users tease out whether they are switching from one practice to another or concurrently billing under more than one TIN. If a provider

[4] Other CMS data products use number of line items to assign providers, for instance, to specialty.
[5] These services correspond to BETOS codes beginning with M, P, I and T2, respectively.

bills under one TIN early in the year, bills under a different TIN late in the year, and bills under both TINs in no more than a single month, one can infer the provider switched between practices. If a provider bills under both TIN1 and TIN2 in two or more months, then one can infer that the provider worked concurrently for multiple organizations during the year.

The database includes each TIN's legal name. The percentage of providers with missing legal names for TIN1 decreases monotonically from 13.0% in 2008 to 4.4% in 2019. Almost all of those providers with missing TIN names are in solo practices, that is, no other NPI billed under that TIN. In the last four years, approximately 99% were in solo practice. Previously, the percentage was slightly less.

It is legal for an entity such as a practice to use multiple TINs but aggregate the income across those TINs when filing federal tax returns. We have found apparent evidence of multiple-TIN practices—sets of TINs with names referring to the same university and a few pairs of TINs with many providers billing under both TINs. These multiple-TIN practices will be captured as separate groups in the database. While this evidence does not indicate that multiple-TIN practices are common, users should keep the issue in mind.

Finally, users should note that a provider billing under the TIN of a hospital might not necessarily be an employee of the hospital. Specifically, if a provider's group has a professional services arrangement (PSA) with a hospital (or other entity) its member providers can reassign their claims to the hospital.[6] Thus, some providers who bill under the TIN of a hospital may, in fact, be members of a provider group with a PSA with the hospital. There is no administrative data on the nature of the hospital-provider group affiliation. The hospital and provider group may be close—with the hospital owning the medical records of the practice and its office building—or the relationship may be less integrated.

Provider Demographic Variables

The NPPES data include information on a provider's self-reported sex and the PECOS data reports providers' birthdate. The percent missing for sex decreases monotonically from 0.6% in 2008 to 0.01% in 2019. The percent missing for birthdate decreases from 3.2% in 2008 to 0.01% in 2019.

Provider Specialty Variables

To participate in Medicare, physicians and other providers must enroll in PECOS. The PECOS maintains information on a provider's self-reported specialty over time. For each specialty, we retained only the primary specialty that was reported as of the PECOS extract date (June 2021). A small minority of providers (less than one percent in any given year) reported multiple primary specialties on the same most recent date. For providers that reported physician and non-physician specialties, the non-physician

[6] Medicare Claims Processing Manual, Chapter 1 – General Billing Requirements, (Rev. 2648, 02-01-13), Sec. 30.2.7, Payments for Service Provided Under a Contractual Arrangement – Carrier Claims Only.

specialty codes were deleted. Among the remaining providers, there were only at most two primary specialties reported in PECOS. The MD-PPAS reports both specialties. For the small proportion of providers for who PECOS lacked specialty data, we used the specialty with a plurality of line items as reported on claims.

We grouped physicians (allopathic and osteopathic) and non-physicians based on their primary specialty information into one of 11 broad specialties. For physicians the broad specialty categories are primary care, medical specialty, surgical specialty, psychiatry, obstetrics/gynecology, and hospital-based specialty. For non-physicians the broad specialty categories are nurse practitioners and physician assistants, other advanced practice nurses, limited liability physicians (LLP), physical, occupational, and speech therapists, and other non-physicians. Appendix C details how the broad specialty categories are defined based on primary specialty.

While most of the specialties used to develop the broad specialty categories rely on self-reported specialty data in the PECOS or claims data, assignment of the hospitalist specialty[7] involves using a more complicated methodology. Specifically, a provider is classified as a hospitalist in the first primary specialty variable (SPEC_PRIM_1) if they self-report that specialty in the PECOS data or meet the claims-based definition of being primary care physicians who had at least 90 percent of line items delivered in an inpatient setting.[8] For those hospitalists that are classified as hospitalist using the claims, their original PECOS specialty is captured in the second primary specialty variable (SPEC_PRIM_2). The MD-PPAS includes an indicator that flags whether the hospitalist designation in the first primary specialty variable is based on the self-reported designation in PECOS or the claims-based definition. Table 1 shows the distribution of hospitalists by data source and original specialty in 2019.

[7] Hospitalists are a subset of hospital-based physicians as defined for the electronic health record (EHR) incentive program, which also includes both surgeons and emergency room physicians.

[8] This claims-based definition of hospitalists is taken from Kuo YF, G Sharma, JL Freeman, and JS Goodwin, "Growth in the Care of Older Patients by Hospitalists in the United States," NEJM 360:11 (March 12, 2009): 1102-1112. Inpatient setting is identified as place of service = 21 on the claim.

Table 1. Distribution of Hospitalists by Data Source and Original Specialty (2019)

Data Source and Original Primary Specialty	Number of Hospitalists	Percent of All Hospitalists
Total	**37,878**	**100.0%**
PECOS Self-Report	16,348	43.2%
Claims-Based Definition [1]	21,530	56.8%
Internal Medicine	15,571	41.1%
Family Practice	2,801	7.4%
Pediatric Medicine	2,043	5.4%
Hospice and Palliative Care	829	2.2%
Geriatric Medicine	143	0.4%
General Practice	116	0.3%
Osteopathic Manipulative Medicine	24	0.1%
Preventive Medicine	3	0.0%

NOTES: [1] These providers did not report the hospitalist specialty in the PECOS data. MD-PPAS assigns them as hospitalists because they are primary care physicians with 90% or more of their line items in the inpatient hospital setting. The specialties listed are the designations found in the original first primary specialty variable from the PECOS data. The MD-PPAS reports these specialties in the second primary specialty variable (SPEC_PRIM_2) and reports the hospitalist specialty in the first primary specialty variable (SPEC_PRIM_1).

Hospitalists who self-report in the PECOS represent 43.2% of all hospitalists in the 2019 MD-PPAS data, an increase from 26.1% in the 2009 MD-PPAS data. This result is consistent with the fact that providers are required to revalidate their information in PECOS only every five years, suggesting that the percentage of providers identifying as hospitalists in the MD-PPAS will increase over time.[9] The remaining 56.8% of hospitalists are classified as such based on their claims and the vast majority of these physicians report Internal Medicine as their primary specialty.

Table 2 shows the number of providers by broad specialty and year.

[9] We extracted the PECOS data about eleven months after implementation of the hospital code. Some providers revalidate their information for reasons other than the five-year revalidation requirement, which could contribute to the self-report population being higher than one-fifth of all hospitalists.

Table 2. Distribution of Providers by Broad Specialty Category (2009, 2019)

Broad Specialty Category	2009		2019	
	Total Number of Providers	Percent of Total	Total Number of Providers	Percent of Total
Total Providers	830,369	100.0%	1,209,840	100.0%
Total Physicians	558,079	67.2%	671,135	55.5%
Primary Care	160,113	19.3%	186,330	15.4%
Medical Specialty	104,499	12.6%	127,131	10.5%
Surgical Specialty	94,898	11.4%	106,899	8.8%
Obstetrics/Gynecology	31,294	3.8%	35,093	2.9%
Hospital-Based Specialty	140,428	16.9%	187,678	15.5%
Psychiatry	26,847	3.2%	28,004	2.3%
Non-Physicians	272,070	32.8%	538,096	44.5%
Nurse Practitioners and Physician Assistants	84,353	10.2%	255,520	21.1%
Other Advanced Practice Registered Nurses	36,688	4.4%	55,132	4.6%
Limited Liability Physician (LLP)	45,237	5.4%	54,079	4.5%
Physical, Occupational, and Speech Therapists	39,943	4.8%	80,095	6.6%
Other	65,849	7.9%	93,270	7.7%
Unknown Specialty	220	0.03%	609	0.05%

NOTE: We show data from the 2009 MD-PPAS rather than the 2008 because 2009 is the first year where we have a full year of claims data. The MD-PPAS uses claims data to determine the universe of providers included in the file. Specifically, the data are restricted to individual providers that submitted a Part B non-institutional claim for evaluation and management services, procedures, imaging, or non-laboratory testing with a positive allowed charges amount.

In attempting to replicate these totals, users must make two decisions:

- Whether to maintain the classification of primary care physicians who meet the claims-based definition as hospitalists. The database includes an indicator that identifies the data source for the hospitalist classification.

- How to deal with providers who reported two primary specialties. Our method for classifying providers who reported two primary specialties is discussed in Appendix D.

There are two caveats about using the specialty data. First, in analyzing changes in specialties over time, users should keep in mind the impact of new codes. For instance, the measured increase in the number of cardiologists might understate the true increase because of the introduction of interventional cardiology. In general, the longer the time series, the more changes in health care organization and

technology, which might be reflected in changes in specialty categories. Appendix E lists the specialty codes introduced since 2011. Second, users should use the records for non-physicians with caution. These providers often submit claims using the NPI of their supervising physician.[10]

Place of Service Variables

Medicare requires that professional claims specify the type of location where the service was delivered. The place of service (POS) variables are included here to yield additional insights into how the provision of care has changed over time.

There are almost 50 POS codes in Medicare administrative data.[11] To make the set of POS variables easier to use, we generate an abbreviated set of categories, each representing the percentage of line items that represent a provider's services delivered in a given location. Specifically, we include four individual POS categories that represent a significant portion of total line items (office, inpatient hospital, emergency room) or have particular policy (ambulatory surgical center) or methodological (urgent care center or retail clinic)[12]significance. We have also combined three sets of POS codes into single categories:

- Two hospital outpatient department codes (on- and off-campus[13]) are combined into a single "hospital outpatient department" category;
- Two nursing home codes (skilled nursing facility and other nursing facility) are combined into a single "nursing facility" category; and
- Four codes representing residential arrangements without nursing care (patient home, assisted living facility, group home, and custodial care facility) are combined into a single "patient residence" category.

Finally, an "other" category captures the remaining POS codes (none of which individually have more than half of a percentage point of line items) plus line items without a valid POS code. Table 3 presents the distribution of line items by both the broad POS categories as well as some of the component categories over time.

---

[10] For instance, Medicare directs nurse practitioners and physician assistants: "You must bill under your own NPI for services furnished incident to your own professional services," but "a supervising physician must bill under his or her NPI for services you furnished incident to the physician's professional services." CMS, "Advance Practice Registered Nurses, Anesthesiologist Assistants, and Physician Assistants," Oct. 2016, https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNProducts/MLN-Publications-Items/CMS1244981.html, accessed, May 1, 2016.

[11] Place of Service Code Set, CMS, https://www.cms.gov/Medicare/Coding/place-of-service-codes/Place_of_Service_Code_Set.html, accessed, March 5, 2018.

[12] Internal analyses revealed that providers affiliated with large retail clinics or urgent care centers listed the place of service as "office" as well as "retail clinic" or "urgent care center". Therefore, to help data users more accurately capture the POS for these providers, the data now reports service lines associated with retail clinics and urgent care centers separately.

[13] The "hospital outpatient department, off campus" was implemented on January 1, 2016.

Table 3. Distribution of Claim Service Lines by Place of Service Categories (2009, 2019)

Place of Service Category	2009		2019	
	Total Number of Service Lines	Percent of Total Service Lines	Total Number of Service Lines	Percent of Total Service Lines
Total	903,944,471	100.0%	994,402,846	100.0%
Office	521,209,969	57.7%	571,424,322	57.5%
Retail Clinic	1	0.0%	63,553	0.0%
Urgent Care Center	1,495,408	0.2%	5,487,548	0.6%
Inpatient Hospital	181,909,397	20.1%	154,780,817	15.6%
Outpatient Department	93,828,409	10.4%	126,445,516	12.7%
On-Campus	93,828,409	10.4%	104,369,152	10.5%
Off-Campus[1]	--	--	22,076,364	2.2%
Emergency Room	43,444,621	4.8%	55,478,981	5.6%
Nursing Facility	31,835,973	3.5%	37,575,660	3.8%
Skilled	16,724,360	1.9%	20,765,681	2.1%
Other	15,111,613	1.7%	16,809,979	1.7%
Ambulatory Surgical Center	10,030,014	1.1%	12,873,597	1.3%
Patient Residence	10,520,395	1.2%	18,858,329	1.9%
Home	5,798,040	0.6%	9,935,915	1.0%
Assisted Living Facility	2,500,929	0.3%	7,455,413	0.7%
Group Home	109,537	0.0%	269,645	0.0%
Custodial Care Facility	2,111,889	0.2%	1,197,356	0.1%
Other	9,670,284	1.1%	11,414,523	1.1%
Valid	9,460,734	1.0%	11,331,532	1.1%
Missing	209,550	0.0%	82,991	0.0%

NOTE: We show data from the 2009 MD-PPAS rather than the 2008 because 2009 is the first year where we have a full year of claims data. The following is a list of CMS codes for the place of service locations listed above: 11=Office; 17=Retail Clinic; 20=Urgent Care Center; 21=Inpatient Hospital; 22=Outpatient Hospital, On-Campus; 19=Outpatient Hospital, Off-Campus; 23=Emergency Room; 24=Ambulatory Surgical Center; 31=Skilled Nursing Facility; 32=Nursing Facility; 12=Home; 13=Assisted Living Facility; 14=Group Home; 33=Custodial Care Facility.

[1] The hospital outpatient department - off-campus place of service category became effective on January 1, 2016.

Provider Geographic Location

Medicare claims also require providers to report the zip code in which the service was provided, because provider payment varies geographically. For each year in the 2008-2019 period, less than 4.7% of providers submit claims from multiple states and less than 14.0% submit them from multiple sub-

state geographic areas (Core Based Statistical Area, CBSA,). For those who submit claims from multiple areas, we first assigned each provider to the state with the plurality of his/her part B non-institutional line items. Within this state, we then identified the CBSA with the plurality of line items. Users should be aware that our assignment of CBSA changes over time as the Office of Management Budget changes the CBSA delineation files.

A data element on claims is likely to be reliable to the extent that it is used for payment. Payment for provider services varies across payment localities. For claims processing, zip codes are assigned to these localities, which cross state lines only in one locality (Washington, DC area). Two-thirds of the states have a statewide locality and half of the rest have just two localities, suggesting that the state variable is more reliable than the CBSA variable.

Although no systematic review has been performed, some large practices with multiple sites appear to bill Medicare using fewer zip codes than are listed on their websites. We caution users that claims from large practices may overstate the number of their providers in a CBSA relative to the number in surrounding non-CBSA areas and encourage users to compare a practice's data to its website, which often includes a map of its clinics.[14]

Utilization Summary Measures

To provide researchers with a comprehensive view of the amount of Medicare services billed by a provider, we include the following summary measures obtained from provider claims for their evaluation and management, procedure and non-laboratory test services with positive allowed charges amounts:

- Number of line items,
- Total allowed charges, and
- Number of unique beneficiaries

Each of these variables is aggregated to the NPI-level and for each NPI-TIN combination.

Atypical FFS Medicare Providers

Users should note that since the MD-PPAS include any provider who submitted at least one fee-for-service claim to Medicare during the year, the data contain records for providers that are not typically considered providers in fee-for-service Medicare. For example, about 8,500 pediatricians (less than 10 percent of the estimated pediatricians in practice) submitted a claim in 2012, reflecting the fact that although Medicare primarily enrolls the elderly, Medicare also covers several thousand disabled children. Additionally, many health maintenance organizations (HMOs) do not routinely bill fee-for-service for Medicare, but on occasion, many of their providers may submit at least one fee-for-service

[14] This issue is less of a concern for smaller practices, which are more likely to have a single site.

claim to Medicare for a non-enrollee who seeks emergency care. As a result, the database includes a majority of the providers in a large group-model HMO.

Illustrative Uses of the MD-PPAS

The creation of this database makes it easier for researchers to analyze several issues that involve TIN-based group practices or physician specialties. For example, CMS payments made to providers for having electronic medical records (EMR) are based on TINs, and NCQA's medical homes are defined in terms of NPIs. The MD-PPAS could be combined with these other databases to examine the extent to which small versus large group practices are adopting EMR systems or the specialty mix of group practices that are serving as medical homes. Below are some other illustrative uses of the MD-PPAS

- **Practice size**: A practice can be operationalized as the physicians who bill under a given TIN. Using the number of physicians as a measure of the size of the practice, ASPE has found that practice size has been increasing.[15]

- **Health systems**: The Agency for Healthcare Research and Quality (AHRQ) has created a Compendium of U.S. Health System by linking across multiple data sources.[16] To be considered a health system, a group of providers must have at a minimum one acute care hospital and one group of physicians that provides both primary and specialty care. The providers must be connected via common ownership or joint management. The database is designed to facilitate studies of the impact of vertical integration on quality of care and market consolidation.

  The compendium database is available for download.[17] The file that links each group practice to a system (if any) has the legal name of the TIN but not the TIN itself.[18] To link MD-PPAS to this practice-system file, a TIN-level file (with TIN name) might be created from MD-PPAS. Most of the physicians in MD-PPAS are in practices that link to non-missing values for system (including no system). Many of the remaining ones are solo practitioners whose practice TIN lacks a name in MD-PPAS.[19]

---

[15] WP Welch, AE Cuellar, SC Stearns, and AB Bindman, "Proportion of Physicians in Large Group Practices Continued to Grow in 2009-11," Health Affairs, 32:9 (Sept. 2013).

[16] Furukawa MF, Machta RM, Barrett KA, Jones DJ, Shortell SM, Scanlon DP, Lewis VA, O'Malley AJ, Meara ER, Rich EC. Landscape of health systems in the United States. Medical Care Research and Review. 2019 Jan 23:1077558718823130.

[17] https://www.ahrq.gov/chsp/data-resources/compendium.html, assessed 2-14-2020.

[18] https://www.ahrq.gov/sites/default/files/wysiwyg/chsp/compendium/group-practice-linkage-public-file.csv, assessed 2-14-2020.

[19] Most of TINs with non-unique names can be assigned using one of two approaches: First, typically all of the TINs with the same name in the compendium file are assigned to the same system (or are deemed to belong to no system). Second, in a number of cases (often associated with a medical school), MD-PPAS has several TINs with the same name (e.g., "Mount Sinai School of Medicine"). In the compendium, some these TINs are assigned to a system with a similar name (e.g., "Mount Sinai Health System") but other TINs are unassigned. A reasonable approach would be to assign all of the "Mount Sinai" TINs to the "Mount Sinai" system. .

- **Physician affiliation with specific hospitals:** Ho et al. has used claims to link physician groups to hospitals.[20]

Contact Information

Please contact or Alshadye Yemane, CMS/OEDA alshadye.yemane@cms.hhs.gov or Pete Welch, ASPE/HP, pete.welch@hhs.gov to ask questions on the methodology used to develop the MD-PPAS, make suggestions for improving the methodology, or to report errors in the data file.

[20] Ho V, Tapaneeyakul S, Metcalfe L, Vu L, Short M, (2020). Using Medicare data to measure vertical integration of hospitals and physicians. Health Services and Outcomes Research Methodology. Advance online publication. doi: 10.1007/s10742-020-00207-7

Appendix A: MD-PPAS Sample Inclusion Criteria Impact

(Return to Main Text)

The MD-PPAS only includes providers that meet the following criteria:

- Has a valid NPI value
- Has submitted a Part B non-institutional claim for evaluation and management services, procedures, imaging, or non-laboratory testing
- Has a Part B non-institutional claim with a positive allowed charges amount
- Is an individual provider

Exhibit A1 details how the criteria used to select providers for inclusion in the MD-PPAS produces differences in the total provider count between the part B non-institutional data and the final MD-PPAS production tables. The exclusion criteria are applied in the order that they appear in the table. The majority of providers are excluded because they do not have claims for evaluation and management visits, procedures, imaging or non-laboratory tests. However, because of the hierarchical nature of the way the exclusion criteria are applied a large proportion of these providers may also be organizational providers as well.

Exhibit A1. Impact of MD-PPAS Provider Inclusion Criteria on Final Provider Count

Data Year	Total Providers	Failed Inclusion Criteria Categories				Final MD-PPAS Population
		Invalid NPI Value	No Claims for Evaluation and Management Visits, Procedures, Imaging, or Non-Laboratory Tests	No Claims with Positive Allowed Charges Amounts	Organization Provider	
2008	1,295,484	1,394	500,100	2,182	6,081	785,727
2009	1,367,159	243	526,800	3,640	6,107	830,369
2010	1,405,360	160	532,075	3,430	5,914	863,781
2011	1,456,586	78	543,195	3,635	5,401	904,277
2012	1,504,204	73	553,622	4,071	5,342	941,096
2013	1,548,775	40	564,094	4,930	5,166	974,545
2014	1,587,122	32	570,234	5,887	4,685	1,006,284
2015	1,633,160	40	580,212	6,306	4,266	1,042,336
2016	1,684,063	9	591,033	6,360	4,012	1,082,649
2017	1,741,175	21	604,873	6,749	4,549	1,124,983
2018	1,797,481	15	619,269	7,453	4,408	1,166,336
2019	1,851,494	16	629,968	7,298	4,372	1,209,840

Users should note that differences in the final action algorithm used for the source Part B non-institutional claims data may cause differences in the final provider count when trying to compare the MD-PPAS to a reproduced provider sample created using the internal Part B non-institutional data they have received from CMS.

Appendix B: MD-PPAS File Layout

(Return to main text)

Variable Name	Description	Data Source
ID variables		
npi	National provider identifier (NPI)	Claims
name_last	Provider last name	NPPES
name_first	Provider first name	NPPES
name_middle	Provider middle name	NPPES
Demographic variables		
sex	1=Male; 2=Female	NPPES
birth_dt	Birth date	PECOS
Specialty variables		
spec_broad	Broad specialty based on spec_prim_1. 1 = Physician - Primary care 2 = Physician - Medical specialty 3 = Physician - Surgical specialty 4 = Physician - Obstetrics/gynecology with no primary care specialty. 5 = Physician - Hospital-based specialty (includes designated hospitalists) 6 = Physician - Psychiatry 7 = Non-physician – Nurse Practitioner and Physician Assistant 8 = Non-physician – Other Advanced Practice Registered Nurses (APRN) 9 = Non-physician – Limited Liability Physicians (LLP) 10 = Non-physician – Physical, Occupational, and Speech Therapists 11 = Non-physician – Other 99 = Specialty Unknown	PECOS/claims
spec_prim_1	Primary specialty (the most recently reported in PECOS)	PECOS/claims
spec_prim_1_name	Name of primary specialty	
spec_prim_2	Concurrently reported primary specialty	PECOS/claims
spec_prim_2_name	Name of concurrently reported primary specialty	
spec_source	Source data for specialty 1=PECOS 2=claims	PECOS/claims
spec_source_hosp	Source data for hospitalist specialty designation 1=PECOS 2=claims	PECOS/claims
Place of service (POS) variables		
pos_office	% of line items delivered in office	Claims
pos_retail	% of line items delivered in retail clinics	Claims
pos_urgent	% of line items delivered in urgent care centers	Claims
pos_inpat	% of line items delivered in inpatient hospital	Claims
pos_opd	% of line items delivered in hospital outpatient department (OPD)	Claims

Variable Name	Description	Data Source
pos_er	% of line items delivered in emergency room (ER)	Claims
pos_nursing	% of line items delivered in nursing facility or skilled nursing facility	Claims
pos_asc	% of line items delivered in ambulatory surgery center (ASC)	Claims
pos_resid	% of line items delivered in the patient's residence (i.e., home, assisted living facility, custodial care facility, or group home)	Claims
pos_other	% of line items delivered in other places of service	Claims
Geographic location		
state	State abbreviation with the most line items for that NPI 99=missing	Claims
state_multi	Multiple state indicator (1=multiple states; 0=single state)	Claims
cbsa_type	Type of CBSA for physician 1=Metropolitan area 2=Micropolitan area 3=non-CBSA 9=missing CBSA code	Claims
cbsa_cd	CBSA code with the most allowed line items for that NPI 00000=non-CBSA 99999=missing CBSA code	Claims
cbsa_name	CBSA name	Claims
cbsa_multi	Multiple CBSA indicator (1=multiple CBSAs; 0=single CBSA)	Claims
Utilization summary measures		
npi_srvc_lines	Count of line items billed by NPI	Claims
npi_allowed_amt	Total allowed charges billed by NPI	Claims
npi_unq_benes	Number of unique beneficiaries for whom the NPI billed	Claims
TIN1 variables		
tin1	Tax identification number (TIN) with the most service lines	Claims
tin1_legal_name	TIN1 legal name	PECOS
tin1_srvc_month	Twelve monthly flags for whether the NPI billed for any services under TIN1. Position 1 pertains to January; position 12 to December. 1= billed 0=did not bill	Claims
tin1_srvc_lines	Count of line items billed under TIN1	Claims
tin1_allowed_amt	Total allowed charges billed under TIN1	Claims
tin1_unq_benes	Number of unique beneficiaries for whom the NPI billed under TIN1	Claims
TIN2 variables		
tin2	Tax identification number (TIN) with the most service lines	Claims
tin2_legal_name	TIN2 legal name	PECOS
tin2_srvc_month	Twelve monthly flags for whether the NPI billed for any services under TIN2. Position 1 pertains to January; position 12 to December. 1= billed 0=did not bill	Claims

Variable Name	Description	Data Source
tin2_srvc_lines	Count of line items billed under TIN2	Claims
tin2_allowed_amt	Total allowed charges billed under TIN2	Claims
tin2_unq_benes	Number of unique beneficiaries for whom the NPI billed under TIN2	Claims

Appendix C: Definition of included provider types and broad specialty categories

(Return to main text)

PECOS recognizes more than 50 physician specialties as well as a number of limited license physician (LLP) and non-physician specialties. The exhibit below shows how the PECOS specialties were collapsed into 11 broad physician and non-physician specialties.

Exhibit C.1: Broad specialties defined in terms of PECOS specialty codes

Broad category	CMS Designations		
	Practitioner category	PECOS code	Description
Primary care	Physician	01	General Practice
Primary care	Physician	08	Family Practice
Primary care	Physician	11	Internal Medicine
Primary care	Physician	12	Osteopathic Manipulative Medicine
Primary care	Physician	17	Hospice And Palliative Care
Primary care	Physician	37	Pediatric Medicine
Primary care	Physician	38	Geriatric Medicine
Primary care	Physician	84	Preventative Medicine
Medical specialty	Physician	03	Allergy/Immunology
Medical specialty	Physician	06	Cardiovascular Disease (Cardiology)
Medical specialty	Physician	C3	Interventional Cardiology
Medical specialty	Physician	07	Dermatology
Medical specialty	Physician	10	Gastroenterology
Medical specialty	Physician	13	Neurology
Medical specialty	Physician	21	Cardiac Electrophysiology
Medical specialty	Physician	29	Pulmonary Disease
Medical specialty	Physician	39	Nephrology
Medical specialty	Physician	44	Infectious Disease
Medical specialty	Physician	46	Endocrinology
Medical specialty	Physician	66	Rheumatology
Medical specialty	Physician	79	Addiction Medicine
Medical specialty	Physician	82	Hematology
Medical specialty	Physician	83	Hematology/Oncology
Medical specialty	Physician	90	Medical Oncology
Medical specialty	Physician	C0	Sleep Medicine
Medical specialty	Physician	C7	Advanced Heart Failure and Transplant Cardiology
Medical specialty	Physician	C8	Medical Toxicology
Medical specialty	Physician	C9	Hematopoietic Cell Transplantation and Cellular Therapy

Broad category	CMS Designations		
	Practitioner category	PECOS code	Description
Medical specialty	Physician	D3	Medical Genetics and Genomics
Medical specialty	Physician	D7	Micrographic Dermatologic Surgery
Medical specialty	Physician	D8	Adult Congenital Heart Disease
Surgery specialty	Physician	02	General Surgery
Surgery specialty	Physician	04	Otolaryngology
Surgery specialty	Physician	14	Neurosurgery
Surgery specialty	Physician	18	Ophthalmology
Surgery specialty	Physician	20	Orthopedic Surgery
Surgery specialty	Physician	23	Sports Medicine
Surgery specialty	Physician	24	Plastic And Reconstructive Surgery
Surgery specialty	Physician	28	Colorectal Surgery
Surgery specialty	Physician	33	Thoracic Surgery
Surgery specialty	Physician	34	Urology
Surgery specialty	Physician	40	Hand Surgery
Surgery specialty	Physician	76	Peripheral Vascular Disease
Surgery specialty	Physician	77	Vascular Surgery
Surgery specialty	Physician	78	Cardiac Surgery
Surgery specialty	Physician	91	Surgical Oncology
Obstetrics-Gynecology	Physician	16	Obstetrics/Gynecology
Obstetrics-Gynecology	Physician	98	Gynecological/Oncology
Hospital based	Physician	05	Anesthesiology
Hospital based	Physician	09	Interventional Pain Management
Hospital based	Physician	22	Pathology
Hospital based	Physician	25	Physical Medicine And Rehabilitation
Hospital based	Physician	30	Diagnostic Radiology
Hospital based	Physician	36	Nuclear Medicine
Hospital based	Physician	72	Pain Management
Hospital based	Physician	81	Critical Care (Intensivists)
Hospital based	Physician	92	Radiation Oncology
Hospital based	Physician	93	Emergency Medicine
Hospital based	Physician	94	Interventional Radiology
Hospital based	Physician	C6	Hospitalist [1]
Psychiatry	Physician	26	Psychiatry
Psychiatry	Physician	27	Geriatric Psychiatry
Psychiatry	Physician	86	Neuropsychiatry
Other	Physician	99	Undefined Physician Type
Nurse Practitioner/ Physician Assistant	Non-physician	50	Nurse Practitioner
Nurse Practitioner/	Non-physician	97	Physician Assistant

Broad category	CMS Designations		
	Practitioner category	PECOS code	Description
Physician Assistant			
Other Advanced Practice Nurses	Non-physician	42	Certified Nurse Midwife
Other Advanced Practice Nurses	Non-physician	43	Certified Registered Nurse Anesthetist
Other Advanced Practice Nurses	Non-physician	89	Clinical Nurse Specialist
LLP	Non-physician	19	Oral Surgery (Dentists Only)
LLP	Non-physician	35	Chiropractic
LLP	Non-physician	41	Optometry
LLP	Non-physician	48	Podiatry
LLP	Non-physician	85	Maxillofacial Surgery
LLP	Non-physician	C5	Dentist
Physical, Occupational, Speech Therapists	Non-physician	65	Physical Therapist
Physical, Occupational, Speech Therapists	Non-physician	67	Occupational Therapist
Physical, Occupational, Speech Therapists	Non-physician	15	Speech Language Pathologist
Other	Non-physician	32	Anesthesiology Assistant
Other	Non-physician	62	Psychologist Billing Independently
Other	Non-physician	64	Audiologist
Other	Non-physician	68	Clinical Psychologist
Other	Non-physician	71	Registered Dietitian Or Nutrition Prof
Other	Non-physician	73	Mass Immunization Roster Biller
Other	Non-physician	80	Clinical Social Worker
Other	Non-physician	88	Undefined Non-Physician Type

Source: CMS specialty values come from the U.S. Dept. of Health and Human Services, "CMS Manual System, Pub 100-06 Medicare Financial Management, Transmittal 209," April 27, 2012, subsections 4-5.
http://cms.gov/Regulations-and-Guidance/Guidance/Transmittals/2012-Transmittals-Items/R209FM.html.
Notes: [1] The hospitalist specialty designation includes both providers that self-report and those primary care providers who report 90% or more of their line items are delivered in the inpatient hospital setting.

Appendix D: Classifying physicians with multiple primary specialties

(Return to main text)

The PECOS database includes variables on whether the provider self-reported a given specialty to be his/her primary specialty and the date on which specialty was entered. For each specialty, we retained only the primary specialty that was reported on the most current date. However, a small minority of physicians reported multiple primary specialties on the same most recent date. For these physicians, we retained the first two reported primary specialties.

If the specialties spanned more than one broad category, we applied our understanding of training pathways and clinical judgment to create a set of rules to classify each of these physicians into only one broad specialty.

- Primary care overrides obstetrics/gynecology and general surgery.
- Medical specialty overrides primary care.
- Surgical specialty overrides all of the above.
- Obstetrics/gynecology overrides all of the above except for primary care.
- Hospital-based specialty overrides all of the above.
- Psychiatry overrides all of the above.

Appendix E: Specialty Codes Introduced Since 2011

(Return to main text)

The table below contains the implementation dates, codes, and names for specialties introduced since 2011.

Exhibit E.1: Implementation Dates for Medicare Specialty Codes

Implementation Date	Specialty Code	Specialty Description
October 2012	23	Sports Medicine
October 2012	C0	Sleep Medicine
January 2015	C3	Interventional Cardiology
January 2017	C5	Dentist
April 2017	C6	Hospitalist
October 2017	C7	Advanced Heart Failure and Transplant Cardiology
October 2017	C8	Medical Toxicology
October 2017	C9	Hematopoietic Cell Transplantation and Cellular Therapy
October 2018	D3	Medical Genetics and Genomics
January 2019	D4	Undersea and Hyperbaric Medicine
May 2020	D7	Micrographic Dermatologic Surgery
May 2020	D8	Adult Congenital Heart Disease

NOTE: The breaks in the series (C1, C2, and C4) do not represent provider specialties.

SOURCE: MLN Matters articles: for 23 and C0, MM7600 (revised 4-27-2012) https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNMattersArticles/Downloads/MM7600.pdf; for C3, MM8812 (revised 9-26-2014) https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNMattersArticles/downloads/MM8812.pdf; for C5, MM9355 (revised 6-24-2016) https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNMattersArticles/Downloads/MM9355.pdf; for C6, MM9716 (revised 11-28-2016), and for C7, C8, and C9, MM9957 (4-26-2017) https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNMattersArticles/downloads/MM9957.pdf; for D3, MM10457 (10-1-2018), https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNMattersArticles/downloads/MM10457.pdf; for D4, MM10666 (revised 12-19-2018), https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNMattersArticles/downloads/MM10666.pdf; for D7 and D8 https://www.cms.gov/files/document/r10124cp.pdf (accessed 06-21-2021).

Appendix F: CBSA Variable Changes in Version 2.4

(Return to Main Text)

Version 2.4 of the MD-PPAS data uses a new methodology to assign providers to CBSA based on the zip code reported on their part B non-institutional claims. This change resulted in a small percentage of providers with different values for their CBSA code and type variable. . Exhibit F1 shows the percentage of providers that had their values for CBSA code and CBSA type change when the updated crosswalk was used.

Exhibit F1. CBSA Variable Value Changes Between MD-PPAS Version 2.3 and Version 2.4

Year	Number of Providers in Both Versions	Percent of Providers with Changes in CBSA Code	Percent of Providers with Changes in CBSA Type
2008	785,726	0.06%	0.02%
2009	830,368	0.06%	0.02%
2010	863,781	0.06%	0.02%
2011	904,277	0.06%	0.02%
2012	941,096	0.003%	0.003%
2013	974,544	0.002%	0.002%
2014	1,006,282	0%	0%
2015	1,042,336	0.10%	0.10%
2016	1,082,649	0.20%	0.20%
2017	1,124,981	0.01%	0.01%

For changes to the CBSA code, the percentage of providers with updates to their values range from 0.06% and 0%. For changes to the CBSA type variable, the percentage of providers with updates to their values range from 0.2% to 0%. In both instances, earlier data years had a higher percentage of providers with value changes than later years.