CMS 2008 Basic Stand Alone (BSA) Carrier Line Items Public Use File (PUF) Frequently Asked Questions (FAQ)

1. What is the *CMS 2008 BSA Carrier Line Items PUF*?

The *CMS 2008 BSA Carrier Line Items PUF* is a free downloadable file containing a subset of the information contained in carrier claims for the services provided by non-institutional providers (e.g., physicians) to a 5% sample of 2008 Medicare beneficiaries. Each of the 67,735,075 records in the file pertains to one line item of a carrier claim. Each record includes:

1. Gender of the beneficiary;
2. Age of the beneficiary at the end of 2008, reported as (1) under 65 years of age; (2) 65 to 69; (3) 70 to 74; (4) 75 to 79; (5) 80 to 84; and (6) 85 and older;
3. International Classification of Diseases, Clinical Modification (ICD-9-CM) for the line item;
4. Healthcare Common Procedure Coding System (HCPCS) code for the line item;
5. Berenson-Eggers Type of Service (BETOS) code for the line item;
6. Count of services associated with the line item;
7. Type of provider associated with the line item;
8. Type of service associated with the line item;
9. Place of service associated with the line item; and
10. Rounded Medicare payment for the line item.

2. How was this PUF created?

The *CMS 2008 BSA Carrier Line Items PUF* originates from a disjoint 5% random sample of beneficiaries from the 100% Beneficiary Summary File for 2008. To exclude any overlap with the beneficiaries in the existing 5% CMS research sample, the beneficiaries in that sample were excluded, and a 5-in-95 random draw was made of the remaining 95% of beneficiaries.[1] All carrier line items for the selected 5% of beneficiaries were then included in the sample from which the *CMS 2008 BSA Carrier Line Items PUF* was developed.

The selected line items were subjected to a thorough de-identification process. The methods used to protect the identity of beneficiaries are described in the answer to the next question.

---

[1] http://www.resdac.org/tools/TBs/TN-011_How5percentMedicarefilescreated_508.pdf

3. What has been done to protect the privacy of Medicare beneficiaries and?

Of paramount importance in the release of the PUF is the protection of beneficiary and provider confidentiality. To that end, all directly identifiable information has been removed in accordance with the HIPAA Privacy Rules.

Other important steps were taken as well:

- Only a small subset of possible variables was selected for inclusion in the file. This reduced the possible information that could be used to identify the beneficiaries and providers included in the new 5% sample.
- For the variables selected for inclusion, categorization was used to protect identities. For example, in place of date of birth or current age in years, age was categorized into six (6) intervals: (1) under 65 years of age; (2) 65 to 69; (3) 70 to 74; (4) 75 to 79; (5) 80 to 84; and (6) 85 and older. This categorization allows researchers to differentiate patterns in other data (e.g., in the frequency of a particular line item) between claims of younger and older beneficiaries but not to use age or date of birth as a highly identifying variable.
- No geographic information is provided.
- The final protection was provided by excluding some records from the final PUF -- those for which the combination of values for all ten (10) variables in the file was extremely uncommon in the Medicare population. No combination that occurred for fewer than 11 beneficiaries in the full Medicare population was allowed into the final PUF. This criterion disproportionately excluded line items with uncommon HCPCS codes, uncommon primary diagnosis codes, or uncommon values in the remaining variables. However, the criterion also assured that no record in the PUF could be linked to a particular beneficiary no matter how much information a user knew about any beneficiary.

4. How was provider confidentiality protected?

The *CMS 2008 BSA Carrier Line Items PUF* contains some information about providers, such as place of service, type of service and type of provider. Provider confidentiality was protected in the same manner as the beneficiary confidentiality: No combination that occurred for fewer than 11 providers in the full Medicare population was allowed into the final PUF.

5. Why is this file a line item-level file and not a claim-level or a beneficiary-level file?

It was determined that a claim-level or a beneficiary-level PUF carries a high risk of re-identification for a significant number of beneficiaries. The *CMS 2008 BSA Carrier Line Items PUF* reduces the risk by breaking the link between line items for a given claim and beneficiary.

6. What does the 3-digit ICD-9-CM primary diagnosis code (CAR_LINE_ICD9_DGNS_CD) mean?

CAR_LINE_ICD9_DIAG_CD provides the first three digits of the ICD-9-CM diagnosis code on the carrier line item (for V codes the first three alphanumeric characters and for E codes the first 4 alphanumeric characters are provided). The original ICD-9-CM diagnosis codes are composed of codes up to five (5) digits. CAR_LINE_ICD9_DGNS_CD provides information on the heading of a category, whereas the fourth and/or fifth digits (those to the right of the decimal point in the ICD-9-CM diagnosis codes) provide greater detail. The *CMS 2008 BSA Carrier Line Items PUF* SAS Data Users Guide provides the descriptions of these 3-digit ICD-9-CM diagnosis codes.

7. What does the HCPCS code variable (CAR_LINE_HCPCS_CD) mean?

CAR_LINE_HCPCS_CD provides the HCPCS codes that identify line items (or procedures) in carrier claims. The descriptions for Level II HCPCS codes are available on the CMS website.[2] The descriptions for Level I HCPCS codes are available from American Medical Association.[3] To learn more about this variable, refer to the ResDAC website.[4]

8. What does the BETOS code variable (CAR_LINE_BETOS_CD) mean?

CAR_LINE_BETOS_CD corresponds to groupings of HCPCS codes which consist of readily understood clinical categories and were developed to analyze the growth of Medicare expenditures. The BETOS codes cover all HCPCS codes and assign each HCPCS code to only one BETOS code. The description of BETOS codes and the HCPCS-BETOS crosswalk are available on CMS website.[5]

9. What does the service count variable (CAR_LINE_SRVC_CNT) mean?

CAR_LINE_SRVC_CNT provides the count of the total number of services of a particular line item listed in the *CMS 2008 BSA Carrier Line Items PUF*. To learn more about this variable, refer to ResDAC website.[6]

10. What does the type of provider variable (CAR_LINE_PRVDR_TYPE_CD) mean?

CAR_LINE_PRVDR_TYPE_CD indicates the type of provider that furnished the medical service. These types range from individual physicians to groups, clinics, and laboratories. For more information refer to the Chronic Condition Data Warehouse for more information.[7]

---

[2] https://www.cms.gov/HCPCSReleaseCodeSets/ANHCPCS/list.asp
[3] http://www.ama-assn.org
[4] http://www.resdac.org/ddvh/NewFiles/HCPCS_CD.htm
[5] https://www.cms.gov/HCPCSReleaseCodeSets/20_BETOS.asp
[6] http://www.resdac.org/ddvib/Files/SRVC_CNT.htm
[7] http://ccwdata.org/downloads/variables/datadictionary/claims_b/PRV_TYPE.TXT

11. What does the type of service variable (CAR_LINE_CMS_TYPE_SRVC_CD) mean?

CAR_LINE_CMS_TYPE_SRVC_CD is a variable that provides information regarding the type of service performed on a patient. Examples include medical care, surgery, consultation, diagnostic radiology, anesthesia and ambulance. For more information refer to the Chronic Condition Data Warehouse for more information.[8]

12. What does the place of service variable (CAR_LINE_PLACE_OF_SRVC_CD) mean?

CAR_LINE_PLACE_OF_SRVC_CD indicates the place of service where the medical service was provided. Examples include physician's office, patient's home, inpatient hospital, emergency room, skilled nursing facility, and outpatient hospital. Please refer to the Chronic Condition Data Warehouse for more information.[9]

13. How is Medicare payment amount represented in the *CMS 2008 BSA Carrier Line Items PUF*?

Payments were rounded according to the following rules:

Value of Medicare Payment Amount ($)	Rounding Rule
0 - 100	Round to the nearest multiple of 5
100 - 500	Round to the nearest multiple of 25
500 - 1,000	Round to the nearest multiple of 100
Greater than 1,000	Round to the nearest multiple of 1,000

14. What data cleaning steps were performed to obtain the initial 5% carrier sample?

The 5% carrier sample was cleaned by removing line items with invalid HCPCS codes and denied line items. Invalid ICD-9-CM diagnosis codes were replaced with missing values. De-identification procedures were then performed on the resulting initial 5% carrier line item sample (see FAQ #3 above).

15. Can I know which line items belong to the same Medicare beneficiary?

The *CMS 2008 BSA Carrier Line Items PUF* does not allow users to link multiple line items on the file for those beneficiaries with more than one line item in 2008. The record identification field on the PUF contains a new series of random numbers generated just for the *CMS 2008 BSA Carrier Line Items PUF,* which is used to sort line items in a random order. Users wishing to work with a subsample of the line items on the file can use the record identifier to draw a random subset of records.

---

[8] http://ccwdata.org/downloads/variables/datadictionary/claims_b/TYPSRVCB.TXT
[9] http://ccwdata.org/downloads/variables/datadictionary/claims_b/PLCSRVC.TXT

16. Can I know which line items belong to the same claim for a Medicare beneficiary?

A carrier claim (for a beneficiary) might include more than one line item. However, the *CMS 2008 BSA Carrier Line Items PUF* does not allow users to link multiple line items for a single claim. (Also, see FAQ #15 above).

17. How is the *CMS 2008 BSA Carrier Line Items PUF* different from the 5% CMS standard research sample?

There is no overlap in terms of beneficiaries between the 5% CMS standard research sample and the *CMS 2008 BSA Carrier Line Items PUF*. These two 5% samples are disjoint.

18. What are the limitations of the *CMS 2008 BSA Carrier Line Items PUF*?

The CMS 2008 BSA Carrier Line Items PUF is intended to give researchers a convenient initial look at data drawn from CMS carrier claims. The file contains measures of demographic characteristics of beneficiaries, diagnosis, services provided, count of services, information about the type of provider and service, place of service, and Medicare payment amount for the line item. In order to preserve confidentiality, suppression criteria have been applied to variables and line items in the initial file. Some variables are rounded or categorized. Researchers should read the General Documentation and the Data Dictionary and Codebook to determine the appropriateness of this PUF for addressing specific research questions.

19. How may I request additional data?

See the Files for Order section of the CMS Web site http://www.cms.gov/home/rsds.asp. This site lists available CMS data files, data file properties, information about data-use agreements, as well as ordering and payment information.

20. What is the plan for future data releases?

The *CMS 2008 BSA Inpatient Claims PUF* was released in February 2011. This release contains one BSA PUF for each remaining type of service: *Durable Medical Equipments, Prescription Drug Events, Hospice, Home Health Agency, Skilled Nursing Facility, Carrier, and Outpatient.* CMS intends to release additional PUFs in 2011-2012.

21. How may I provide feedback on the *CMS 2008 BSA Carrier Line Items PUF*?

Questions and comments can be submitted to the Research Data Assistance Center (http://www.resdac.org/) via resdac@umn.edu or 1-888-9RESDAC.