

CMS 2008 Basic Stand Alone (BSA)
Home Health Agency (HHA) Beneficiary Public Use File (PUF)
Frequently Asked Questions (FAQ)

1. What is the *CMS 2008 BSA HHA Beneficiary PUF*?

The *CMS 2008 BSA HHA Beneficiary PUF* is a free downloadable file containing a subset of the information contained on home health agency (HHA) claims provided to a 5% sample of 2008 Medicare beneficiaries. Each of the 124,829 records in the file pertains to one Medicare beneficiary. 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. Total number of admissions;
4. Count of therapy visits;
5. Count of skilled nursing care visits;
6. Count of home health aide visits;
7. Rounded Medicare payment for HHA claims of the beneficiary.

2. How was this PUF created?

The *CMS 2008 BSA HHA Beneficiary 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,¹ the beneficiaries in that sample were excluded, and a 5-in-95 random draw was made of the remaining 95% of beneficiaries. All HHA claims for the selected 5% of beneficiaries were then included in the sample from which the *CMS 2008 BSA HHA Beneficiary PUF* was developed. To increase the utility of the file, beneficiaries who were enrolled in Medicare Part A and Part B for less than twelve (12) months in 2008 were excluded.

The selected beneficiaries 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?

Of paramount importance in the release of the PUF is the protection of beneficiary 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 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, the file was created with age 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., counts of therapy visits) between 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 from the final PUF -- those for which the combination of values for all seven (7) 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 beneficiaries with uncommon values in variables such as count of different types of visits and Medicare payment. 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?

There is no risk of provider identification in the *CMS 2008 BSA HHA Beneficiary PUF* as the PUF does not contain any information about individual providers.

5. Why is there no diagnosis information in the *CMS 2008 BSA HHA Beneficiary PUF*?

Information on diagnosis at admission is not included in the *CMS 2008 BSA HHA Beneficiary PUF* to protect the privacy and confidentiality of Medicare beneficiaries. The distribution of beneficiaries across diagnosis codes in the initial 5% sample is highly dispersed, leaving very few beneficiaries per each 5-digit International Classification of Diseases, Clinical Modification (ICD-9-CM) code. Even after coarsening these codes to 3-digit ICD-9-CM or to Major Diagnostic Classification (MDC), the distribution was considered unsafe for release in a PUF.

6. Why are the Health Insurance Prospective Payment System (HIPPS) codes not included in the *CMS 2008 BSA HHA Beneficiary PUF*?

Similar to the issue with diagnosis codes, the distribution of the HIPPS codes was very sparse increasing the risk of the PUF significantly. Hence, they were not included in the PUF.

7. How are the number of visits categorized in the *CMS 2008 BSA HHA Beneficiary PUF*?

Each of the three counts of visits (therapy, skilled nursing care and home health aide) variables is categorized into four: (0) 0 visits, (1) 1 - 13 visits, (2) 14 - 19 visits, and (3) 20 or more visits. This assumption is based on the definition of episodes for HHA claims in the Medicare Claims Processing Manual.

8. How is the *HHA_PMT_AMT* variable calculated?

This variable is the sum of all payments made by Medicare for a beneficiary's HHA claims ending in 2008. It is also the sum over all admissions if a beneficiary has multiple HHA admissions in 2008.

9. How is the *CMS 2008 BSA HHA Beneficiary 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 HHA Beneficiary PUF*. These two 5% samples are disjoint.

10. What are the limitations of the *CMS 2008 BSA HHA Beneficiary PUF*?

The *CMS 2008 BSA HHA Beneficiary PUF* is intended to give researchers a convenient initial look at data drawn from CMS HHA claims. The file contains measures of demographic characteristics of beneficiaries, count of HHA care visits, and Medicare payment amount. In order to preserve confidentiality, suppression criteria have been applied to variables and beneficiaries on 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.

11. 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.

12. 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, Outpatient, and Carrier*. CMS intends to release additional PUFs in 2011-2012.

13. How may I provide feedback on the CMS 2008 BSA HHA Beneficiary 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.