



Medicare Current Beneficiary Survey Public Use File

Lisa B. Mirel, Shannon Corcoran, Jian Zhang

Presented by: Debra Reed-Gillette

Office of Enterprise Data & Analytics

Centers for Medicare & Medicaid Services

Outline

- Background
- Why release an MCBS PUF?
- Designing the MCBS PUF
- Disclosure review process
- Key elements of the MCBS PUF
- Next steps

Background 1

- MCBS is a continuous, multipurpose, in-person survey of a representative national sample of the Medicare population (n~14,000, both aged and disabled)
 - Complex probability design
 - Sample includes beneficiaries living in the community (noninstitutionalized) and in health care facilities
 - Panel survey and overlapping panels create annual files
 - Survey data are combined with CMS claims and administrative data to make a more complete dataset
- Provides important information on Medicare beneficiaries that is *not available* in CMS administrative data;
 - Satisfaction, access to care, and quality of care from the beneficiary perspective
 - Insurance coverage and payments from non-Medicare sources (e.g. out of pocket (OOP))

Background 2

- Began in 1991
- Consists of three interviews a year for up to four years

	2010			2011			2012			2013			2014		
	56 Jan- Apr	57 May- Aug	58 Sep- Dec	59 Jan- Apr	60 May- Aug	61 Sep- Dec	62 Jan- Apr	63 May- Aug	64 Sep- Dec	65 Jan- Apr	66 May- Aug	67 Sep- Dec	68 Jan- Apr	69 May- Aug	70 Sep- Dec
Panel 2010															
Panel 2011															
Panel 2012															
Panel 2013															
															ATC 2013

- Follows beneficiaries into and out of health care facilities
- Oversamples beneficiaries aged less than 65 and 85 and older
- Data are used to inform policy and assess current and future needs of Medicare Beneficiaries

Background 3

- There are two data files from the MCBS for each data year
 - Access to Care (ATC)
 - Cost and Use files (CAU)
- Currently, the files are only available as Limited Data Sets and require a signed agreement between the researcher and CMS (Data Use Agreement)
 - LDS cost \$600 per module per year
 - Data Use Agreement process takes time

Why Release an MCBS PUF?

- Better access to health system information is a key goal of the Department of Health and Human Secretary's Delivery System Reform initiative
 - PUF will open up access to important data on Medicare beneficiaries care
- Public release of de-identified survey data is standard by other HHS agencies
 - AHRQ – MEPS
 - CDC/NCHS - NHANES, NHIS, NAMCS, NHAMCS

Benefits of a MCBS PUF

- Increase policy-relevant analyses
 - Attract new researchers and policy-makers, for whom the cost and time associated with accessing the MCBS LDS can pose significant deterrents to use
 - Promote research to improve the Medicare program
- Provide publically available data for key indicators of the Medicare population for benchmarking comparisons with other data sources

Designing the MCBS PUF

Goal was to release a dataset that includes:

- Demographic data
- Self-reported data on
 - health status and functioning
 - use of and access to preventive services
 - access to and satisfaction with, usual source of care
- Health insurance coverage (both from survey reported and administrative claims)
- Household characteristics

Disclosure Risk and Privacy Protection

- Conducted a thorough evaluation to mitigate disclosure risks (k-anonymity)
- Evaluation included well-established protocols used by other federal agencies that release public-use survey data
 - Federal Committee on Statistical Methodology (FCSM) guidance for determining whether health information is “not personally identifiable” - Statistical Policy Working Paper 22
 - NCHS checklist on disclosure potential of data

Steps Taken to Address Disclosure Risk

- HIPAA Safe Harbor: removal of specific information that could identify an individual
- Variable reduction: releasing limited number of variables
- Coarsening: re-code continuous variables to limit disclosure, but remain useful
- k-anonymity criteria, $k=5$
 - Data for at least 5 respondents are present in each cross-tabulated cell between the selected identifying variables and the variable of interest
 - Consistent with other federal agencies that release de-identified survey-reported health data

Identifying variables

- To conduct an evaluation using k-anonymity we created race, ethnicity, age, and sex categories
- These were referred to as the identifying variables
- All other variables were cross classified with the identifying variables
- Categories for identifying variables:

non-Hispanic white	non-Hispanic black	Hispanic	Other
<65	<65	<65	<65
65-74	65-74	65-74	65-74
75-84	75-84	75+	75+
85+	85+		

Evaluation and Decisions

- Only beneficiaries residing in the community are included.
- Single year of age is not released and only certain age categories are released for Hispanics and Other race/ethnicity groups with smaller number of beneficiaries.
- No dates of any kind are included.
- Monthly insurance coverage is summarized at an annual level.
- No relationships are included in the household characteristics.
- Variables that are missing for a large proportion of respondents are evaluated on a case by case basis to assess if the “missing” is in fact classified as a true value and if it raised a disclosure risk.
- All payment fields have been dropped as well as other continuous variables with substantially skewed distributions.
- All other variables were collapsed to prevent cell sizes <5 when crossed with the identifying variables.

Additional Steps to Prevent Disclosure

- A separate randomly generated identifier was created for the PUF so that no respondent could be linked back to their LDS file or their Medicare Claims records that are available through existing data use agreements.
- No geographic information is on the file so that no linkages can be made to geographic PUFs and other PUFs produced elsewhere within CMS.
- No identifiable plan related information for Medicare Advantage or Medicare Part D is being released in order to eliminate linkages to those quality measure files produced by other areas of CMS.
- The MCBS PUF includes 13,924 beneficiaries. This is less than ½ a percent of all the almost 49 million Medicare beneficiaries in 2013.

Key Elements – MCBS PUF vs. LDS

MCBS PUF	MCBS LDS
Population – community dwelling	Population – community and facility
Number of variables - 472	Number of variables – 1,742
ID – new PUFID Randomly generated, can't be linked back to BASEID	ID – BASEID Randomly generated, can't be linked back to HIC number
Date fields - NO	Date fields - YES
Geographic identifiers – NO	Geographic identifiers - YES
Cost/payment data – NO	Cost/payment data – YES
Demographic data – YES All variables are categorical	Demographic data – YES Continuous age variables available
Insurance coverage – YES Summarized to annual level	Insurance coverage – YES Monthly level
Identifiable plan-related information for MA or Part D - NO	Identifiable plan-related information for MA or Part D - YES

Potential Research Using the MCBS PUF

- Compare access to care for different age and race/ethnicity groups
- Compare health status for those enrolled in Medicare Advantage all year to those who are not
- Assess satisfaction with care for those enrolled in Medicare Advantage by demographic characteristics like education

Next Steps



- Release of the MCBS PUF in May 2016
- Plan to release update annually
- Incorporate user feedback for next release

Thank you!