BISF Documentation

Acumen, LLC

500 Airport Blvd., Suite 365

Burlingame, CA 94010

1 Overview of Beneficiary Summary Files

The Beneficiary Summary Files, derived from Medicare Parts A and B data, were created by Acumen to catalogue series of data across periods. These files, summarized at the beneficiary level for different time intervals, offer user-friendly data structures specifically designed for research purposes. These summary files come in three types – the Beneficiary Annual Summary File (BASF), the Beneficiary Quarterly Summary File (BQSF), and the Beneficiary Interview Gap Summary File (BISF). While the BASF and BQSF respectively use the year and quarter as units of time measurement, the BISF arranges the data from one Medicare interview to the next, a period referred to as the interview gap.

What all Beneficiary Summary Files have in common are the variables they summarize, as explained in Table 1. The variables are grouped into five principal categories, which contain coverage information, diagnosis-related group (DRG) information, beneficiary categorical condition group (CCG) information by year, expenditure and utilization information, and CMS Chronic Condition Warehouse information.

All three summary files condense Medicare Part A and B claims and enrollment information into one observation per beneficiary and unit of time. Individually as well as collectively, they provide comprehensive summary measures related to health and medical expenditures for each type of service and payer, along with utilization measures associated with these services. These files do so without revealing any personally or demographically-identifiable information.

Table 1: Beneficiary Summary File Variable Categories

Summary File Information	Brief Description
Coverage information	Coverage variables for every beneficiary summarize Medicare status code and months of enrollment.
Diagnostic-related groups	Diagnostic-related groups relate types of patients to the resources they consume. Currently, there are approximately 500 DRG's.
RDDC Categorical Condition Group	All valid diagnosis codes per beneficiary are mapped to 30 RDDC Categorical Condition Groups.
Expenditure and utilization	Expenditure and utilization variables describe fee-for-service reimbursements for services covered by CMS Standard Analytical File claim categories.
Chronic Condition Warehouse	CMS Chronic Condition Warehouse variables are constructed according to the algorithm provided by the Iowa Foundation for Medical Care; the algorithm consists of clinical and coverage criteria, as well as yearly and ever flag indicators

2 The BISF: Specific Features

2.1 The Interview Gap

What distinguishes the BISF from the other two MedRIC summary files is its connection to the HRS (Health and Retirement Survey) data, as the BISF is arranged by the interview start and end dates of the HRS, building upon its data. The beginning of an interview period of a beneficiary is noted by the variable Start_dt; correspondingly, the conclusion of a period is given by End_dt. While Start_dt and End_dt are formatted to include a day, month, and year, the days of each Start_dt and End_dt are always the first and last days of the month, respectively.

2.2 New Features: Gap Type Variables

In order to provide the potential for stronger and more meaningful analysis of this information, we have added two new variables to it which differentiate between the starting and closing dates of interviews, assigning them values based on the particular event which caused an interviewee to either first participate in the HRS or to complete the interview process. The variables Start_Gap_Type and End_Gap_Type correspond to a beneficiary's first and final dates of interview, respectively. Start_Gap_Type and End_Gap_Type may each take on only one of four potential values. These include 'N', indicating a new enrollee, 'I', signifying only an interview gap, 'C', showing that a gap has been summed up to the point of current data, and 'D', indicating death. While 'N' is only valid for Start_Gap_Type and 'C' and 'D' are only valid for End_Gap_Type, 'I' may be used for both. To illustrate, 'I' might be the value for the End_Gap_Type where the Start_Gap_Type is 'N', while also being the value for the Start_Gap_Type in another observation where the End_Gap_Type is 'D'. 'I' could even be the value for both Start_Gap_Type and End_Gap_Type. The new addition of gap type variables was initiated for its enrichment of analysis capability beyond the base framework of the claims data structure, adding a unique nuance which could be used by researchers to potentially explore cause and effect relationships originating from the Medicare data in a more complex and salient manner.

Table 2: Gap Type Variable Values

Gap Value	Description	Valid For
N	New Enrollment	Start_Gap_Type
I	Interview Gap	Start_Gap_Type, End_Gap_Type
D	Death Termination	End_Gap_Type
C	Summed to point of current data	End_Gap_Type