

Background

In an effort to assist parties interested in submitting papers who need immediate access to the public data files, the HSCRC has generated public use data files specifically for use in analyses for the white papers. The HSCRC White Paper Data Files are only available for FYs 2011-2013 and contain the following enhancements:

- These Inpatient and Outpatient files contain a limited number of variables that users will need for analysis and are grouped in latest grouper versions (see below for details).
- Inpatient dataset includes data elements from the 3M Potentially Preventable Conditions (PPC) grouper, AHRQ Prevention Quality Indicators (PQI) (v.4.4), and Maryland-specific Case Mix Index (CMI) logic.
- Both Inpatient and Outpatient datasets include variables that will allow users to perform patient-level analysis across hospital care settings (inpatient and outpatient) and across years within the same hospital and calculate readmissions using user-defined intervals (30 days, 60 days, etc.). Similar files with unique master patient ids that link records across hospitals are available only for CY2012.
- Data is available in flat (.txt) and SAS (.sas7bdat) file types.

Important Variables in the Enhanced Dataset

In the White Paper Data Files, there are 2 Medicare Provider ID variables (HOSPID and HOSPID_O) that can be used to include or exclude patients from Johns Hopkins Oncology Center, and University of Maryland Cancer Center.

- **HOSPID:** Hospital ID that aggregates the following hospital units:
 - Johns Hopkins Oncology Center patients under Johns Hopkins (210009)
 - University of Maryland Cancer Center under University of Maryland (210002)
- **HOSPID_O:** Hospital ID that carves out the following units into separate hospital IDs:
 - Johns Hopkins Oncology Center Hospital ID (210904)
 - University of Maryland Cancer Center Hospital ID (218994)

There are also 2 variables that are required to analyze admissions within the same hospital:

- **UNIQUEID:** Patient ID randomly assigned to each unique Medical Record Number by Hospital ID.
- **DAYS_TO_EVENT:** Number of days between the randomly assigned reference date and the actual discharge date. HSCRC revised the methodology developed by AHRQ HCUP to create this variable (See the link for AHRQ HCUP Revisit Methodology Documentation below for more information). This variable allows time between events to be calculated while keeping the dates of actual events confidential.

Using DAYS_TO_EVENT for Analysis of Discharges

DAYS_TO_EVENT is a timing variable used to identify intervals between inpatient and outpatient visits for each unique id. First, a random “start date” is assigned to each unique ID. The **DAYS_TO_EVENT** is calculated for all visits associated with each unique ID. The value of the **DAYS_TO_EVENT** is the difference between the visit discharge date and the random start date associated with the unique ID. The lowest value of **DAYS_TO_EVENT**

will be on the first or earliest event for a patient. To calculate the interval between visits, subtract the **DAYS_TO_EVENT** between two selected visits for the same unique ID.

Please be aware that the value of DAYS_TO_EVENT is not comparable across patients. In other words, the value of DAYS_TO_Event for Patient A has no relation to the value of DAYS_TO_Event for Patient B.

Data Limitations

There are several limitations to this data:

- HSCRC White Paper data sets are linked by a unique ID variable that is created for each unique medical record number for each hospital. The unique ID is will allow users to track a patient across FY 2011-2013, inpatient and outpatient, within the same hospital ID. Due to the inconstancy in the reporting of date of birth, age, and gender variables in the HSCRC abstract data, it is assumed that visits with the same medical record number and hospital ID belong to the same person.
- Users will not be able to track patients across hospitals using the unique ID variable.
- Users will not be able to link to other HSCRC research level files or prior years of data.

Groupers Specifications

- Inpatient: 3M APR-DRG, version 30; 3M PPC, version 30
- Outpatient: 3M EAPGs, version 3.8

For More Information

3M APR-DRG Grouper	APR DRG Software: 3M Health Information Systems - US
3M EAPG Grouper	Outpatient Groupers: 3M Health Information Systems - US
3M PPC Grouper	PPR and PPC Grouping Software: 3M Health Information Systems - US
AHRQ PQI	AHRQ.gov Agency for Healthcare Research & Quality (AHRQ)
AHRQ HCUP Revisit Methodology Documentation	HCUP Supplemental Variables for Revisit Analyses