

SESUG Paper 287-2018  
**Analyzing Hospital Medicare Cost Report Data Using SAS®**  
Kimberly Andrews, Centers for Medicare and Medicaid Services

## ABSTRACT

Medicare-certified institutional health care providers are required to submit annual cost reports, which are maintained by the Centers for Medicare and Medicaid Services (CMS) in the Healthcare Cost Reporting Information System (HCRIS). Medicare Cost Reports (MCR) contain provider information such as facility characteristics, utilization data, total and Medicare costs, inpatient and outpatient charges, Medicare payment data, and financial statement data. HCRIS includes the following subsystems: Hospital, Skilled Nursing Facility (SNF), Home Health Agency (HHA), End-Stage Renal Disease (ESRD) Facility, Hospice, Community Mental Health Center (CMHC), and Rural Health Clinic (RHC)/Federally Qualified Health Center (FQHC). Our discussion focuses on the MCR for **Hospitals** (the most complex of the cost reports) explaining how to access the SAS files; describing the characteristics of the data, and providing basic SAS program code which can be used to analyze the data.

## BACKGROUND

When Medicare was established in 1965, Congress adopted the private health insurance sector's "cost-based reimbursement" system to pay for hospital services. Under this system, hospitals filed MCRs (mandatory per the Social Security Act), which were used to reconcile interim payments to hospitals with "allowable costs," as defined in regulation and policy. In 1982, Congress mandated the creation of the Inpatient Prospective payment system (IPPS), a method of reimbursement in which Medicare payment is made based on predetermined, fixed amounts (for example, inpatient hospital services are reimbursed using diagnosis-related groups (DRGs) on a per discharge basis). PPSs for other healthcare services occurred subsequently.

Prior to the implementation of the PPSs, MCRs were used to reimburse providers for Medicare services. Although MCRs are still used to reimburse providers for a portion of services that remain cost-based (e.g. organ acquisition, nursing anesthetists, and critical access hospitals), cost reports have evolved beyond their original purpose and continue to be a valuable resource to develop market baskets, investigate cost, payment, and utilization trends, and study elements of reimbursement.

## INTRODUCTION

Hospital MCR data is available on the CMS website in SAS datasets for years 2010-2017, one file per year, and in CSV files extracted from the HCRIS database for years 1996-2017. The SAS datasets are provider-level files, meaning each record contains the data fields for a single hospital for a single year (one record per hospital submission). Since there are approximately 6200 hospitals and 7000 data fields, each dataset contains about 6200 observations (rows) and 7000 variables (columns). The average hospital is required to complete about half of the fields.

One of the key aspects of the MCR is that it contains hospital characteristic, utilization, expense and revenue data making it possible to compute a myriad of measures on a national level, by state, by hospital category and for individual hospitals.

The hospital MCR form was revised in 2010. MCR data for 1996-2009 was reported using the CMS-2552-96 form; data for 2012-current was reported using the CMS-2552-10 form; and data for 2010-2011 was reported using both forms, depending on the provider. **This paper focuses on analyzing the Medicare cost report data for hospitals collected by the CMS-2552-10 form, using the SAS datasets.**

The material in this presentation is outlined as follows:

- SAS File Access and Documentation
- Worksheet Flow
- Analytics Using Key Fields
- Audits and Edits
- Outputting Your Results to Excel

## SAS FILE ACCESS AND DOCUMENTATION

The MCR data is on a fiscal year basis (i.e. the 2017 MCR file contains data for facilities with a cost reporting begin date between October 1, 2016 and September 30, 2017). The annual files are updated quarterly and each update includes all cost reports received as of the last day of that quarter. The updates are usually available on the CMS website approximately three weeks after the end of the quarter. Each quarters' annual files are stand-alone, as they typically contain a full year of data for the provider's "fiscal year". For example, one record in the 2017 MCR file would reflect expenses and revenue for a provider reporting on a calendar year 2017 period and another record with expenses and revenue reporting on a federal fiscal year 2017 period. Each record includes the provider's cost reporting period beginning and end dates, as well as a report record number (which is a unique identifier assigned to each MCR submission).

The cost report files are basically mature about 15 months after the year-end. For example, the 2016 cost report file was approximately 60% complete as of September 30, 2017; 90% complete as of December 31, 2017 and over 95% complete as of March 31, 2018.

Table 1 identifies many valuable resources for cost report enthusiasts including links to the SAS datasets, worksheet forms and instructions, data dictionary and other documentation:

**Table 1 MCR Documentation**

Description	Website and Notes
CMS MCR Homepage	<a href="https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/Cost-Reports/">https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/Cost-Reports/</a>
Hospital MCR Homepage	From CMS MCR homepage, click on <a href="#">Hospital Form 2552-10</a>
SAS Files by Year	From Hospital MCR homepage, click on <a href="#">HOSPITAL-2010-SAS.ZIP (SAS datasets and documentation)</a>  The SAS datasets (prds_hosp10_yyyyy) are available for 2010-2017 and include a subset of key cost report fields.  Additional MCR fields and years (1996-2017) are available in CSV extracts from the HCRIS database, as identified below.
SAS File Record Layout	From Hospital MCR homepage, click on <a href="#">HOSPITAL-2010-SAS.ZIP (SAS datasets and documentation)</a>  The record layout and crosswalk to the CMS-2552-96 form are available in the Excel file:  <i>2552-10 SAS FILE RECORD LAYOUT AND CROSSWALK</i>

Provider Reimbursement Manual	<p>From Hospital MCR homepage, click on <a href="https://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/Paper-Based-Manuals-Items/CMS021935.html">https://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/Paper-Based-Manuals-Items/CMS021935.html</a></p> <p>Scroll down and click on Chapter 40 to open the ZIP file containing the worksheet forms and instructions PDFs.</p>
Worksheet Forms	<p>Upon opening the Chapter 40 ZIP file, click on <b>R14p240f.pdf</b></p> <p>Shading indicates the field is not collected. If the field is not shaded and not in the SAS file, it is available from the HCRIS database in CSV files.</p> <p>A note at the bottom of each form identifies the section of the instructions that applies to the fields in that form.</p>
Worksheet Instructions	<p>Upon opening the Chapter 40 ZIP file, click on <b>R14p240i.pdf</b></p>
Provider Number Intelligence	<p>From Hospital MCR homepage, scroll down and click Download <a href="#">HOSPITAL2010-DOCUMENTATION</a></p> <p>HCRIS_STATE_CODES identifies state in first two positions.</p> <p>HCRIS_FACILITY_NUMBERING identifies hospital type in last four positions.</p>
HCRIS Database CSV Files	<p>From Hospital MCR homepage, click on <a href="#">Cost Reports by Fiscal Year</a></p> <p>Use these files if you need a field or a year that is unavailable in the SAS files, yet available in the HCRIS database.</p>

## WORKSHEET FLOW

The hospital MCR is comprised of a series of worksheet forms each of which flows into the next (or sometimes a previous) worksheet. We discuss the flow of the worksheets below and will delve further into using the key fields of each worksheet later in the paper.

- S Series - Identification, Utilization and Wage Index Data
- A Series – Cost Finding
- B Series – Overhead Allocation
- C Series – Cost-to-Charge Ratios
- D Series – Medicare Apportionment
- E Series – Medicare Reimbursement Settlement
- G Series – Financial Worksheets
- H series – Hospital-based HHA
- I series – Hospital-based ESRD
- K series – Hospital-based Hospice
- L series – Capital Payment Worksheet
- M series – RHC/FQHC

## S SERIES – IDENTIFICATION, UTILIZATION AND WAGE INDEX

The first set of worksheets is the S series. These worksheets contain hospital identification data, characteristics, utilization measures and wage index information. Some of these fields trigger skip patterns to identify which other worksheets and fields the hospital should complete. Below is a table showing selected S worksheets and key fields.

**Table 2 Worksheet S Series Key Fields**

Form	Description	Key Fields / Notes
S, Parts I and II	Certification and Settlement	Provider Status Code (as submitted, settled w/ or w/out audit, reopened, amended)
S-2, Parts I and II	Identification and Characteristics	Provider number, hospital name, urban/rural, control type, teaching status
S-3, Part I	Statistics and Utilization Measures	Beds, days, discharges, FTEs
S-3, Part II - VI	Wage Index and Wage-Related	Salaries, benefits, contract labor
S-4	Hospital-Based HHA	Statistical and Utilization Data for Hospital Based Services
S-5	ESRD Dialysis	
S-6	Community Mental Health Center (CMHC) and Other Outpatient Rehab	
S-7	Hospital-Based SNF	
S-8	Hospital-Based RHC/FQHC	
S-9, Parts I - IV	Hospice	
S-10	Uncompensated and Indigent Care	Uncompensated care, charity care, bad debt

## A SERIES – COST FINDING

The next set of worksheets is the A series, described as the cost finding worksheets. In Worksheet A, overhead costs (e.g. capital) and revenue-generating costs (e.g. adults and pediatrics) are identified by cost center and split into their salary and non-salary components. Below is a table showing the key fields in Worksheet A.

**Table 3 Worksheet A Key Fields**

Form	Description	Key Fields
A, Lines 1-23	Overhead	Capital, Administrative and General, housekeeping
A, Lines 30-46	Inpatient Routine	Adults & pediatrics and intensive care
A, Lines 50-77	Ancillary	Operating room, laboratory, therapy
A, Lines 88-93, 93.99	Outpatient	RHC, FQHC, clinic
A, Lines 94-101	Other Reimbursable	Ambulance, durable medical equipment
A, Lines 105-117	Special Purpose	Organ acquisition

A, Lines 190-201	Non-Reimbursable	Gift shop, research, physician offices
A, Column 1	Salary	Includes salary-related costs for each cost center
A, Column 2	Non-Salary	Includes non-salary-related costs for each cost center

## B SERIES – OVERHEAD ALLOCATION

In Worksheet B, the overhead costs (Worksheet A, lines 1-23) are allocated to the revenue-generating cost centers (Worksheet A, lines 30-46, 50-77, 88-93, 93-99, 94-101, 105-117, 190-201) using various statistics from Worksheet B-1, such as square footage to allocate capital and housekeeping costs, meals served to allocate cafeteria costs and pounds of laundry to allocate laundry costs. Worksheet B has a similar structure to Worksheet A.

## C SERIES - COST-TO-CHARGE RATIOS

Cost-to-charge ratios are computed by cost center and used to apportion a share of total ancillary costs to the Medicare program by cost center.

## D SERIES – MEDICARE APPORTIONMENT

In the D Series of worksheets, a share of the total costs is apportioned to the Medicare program by cost center. Medicare costs are not directly reported by providers, but are estimated by applying the following formulas to reported data:

- Routine Medicare Costs = Total Cost Per Day \* Medicare Days
- Ancillary Medicare Costs = (Total Ancillary Costs / Total Ancillary Charges) \* Medicare Charges

## E SERIES – MEDICARE REIMBURSEMENT SETTLEMENT

The E series includes the following worksheets that collect Medicare reimbursement settlement data.

**Table 4 Worksheet E Key Fields**

Form	Description
E, Part A	Inpatient Services
E, Part B	Outpatient Services
E-1, Part II	Health Information Technology (HIT) Payments
E-3, Part II	Inpatient Psychiatric Facility (IPF)
E-3, Part III	Inpatient Rehabilitation Facility (IRF)
E-3, Part IV	Long-Term Care Hospital (LTCH)

Revenue collected includes the following Medicare payments:

- Diagnosis Related Group (DRGs)
- Outlier
- Indirect Medical Education (IME)
- Graduate Medical Education (GME)
- Disproportionate Share (DSH)
- Uncompensated Care (UCC)
- End Stage Renal Disease (ESRD) Add-ons

## G SERIES – FINANCIAL WORKSHEETS

The Worksheet G Series contains the financial statements completed according to Generally Accepted Accounting Principles (GAAP), including the Balance Sheet and Statement of Revenue and Expenses. The total revenue and expense fields can be used to compute a total facility profit margin, as described later in this paper.

**Table 5 Worksheet G Series Key Fields**

Form	Description	Key Fields / Notes
G	Balance Sheet	Assets, liabilities
G-1	Statement of Changes in Fund Balances	Used to compute cash flow
G-2	Statement of Patient Revenues and Operating Expenses	Patient revenues, operating expenses
G-3	Statement of Revenue and Expenses	Total facility revenue and expenses

## OTHER WORKSHEETS

The following supplemental worksheets feed into the main worksheets described above:

**Table 6 Supplemental Worksheets**

H Series	Hospital-Based HHA
I Series	Hospital-Based ESRD
K Series	Hospital-Based Hospice
L Series	Capital Payments
M Series	RHC/FQHC

## USING THE WORKSHEET FORMS AND INSTRUCTIONS (S-3 EXAMPLE)

The figure below shows a section of Worksheet S-3, Part I, which contains the following:

- **utilization** including beds, days and discharges by cost center (rows) and by payer (columns)
- **full-time equivalents** (FTEs) by cost center (rows) and interns and residents, employees on payroll, and unpaid workers (columns)

The variable names in the MCR are defined by the worksheet, line and column of the field. Notice the field for total hospital general inpatient beds is found in column 2, line 14 (SAS file variable name = S3\_1\_C2\_14). Line 14 is the sum of lines 7 through 12. Similarly, column 2, line 27 contains total facility beds, including SNF, LTCH, IPF, and IRF beds as well as hospital general inpatient beds. Unshaded fields represent fields collected and available in the HCRIS database. A subset of key fields are included in the SAS files. See the cost report instructions as described below for a more detailed definition of each field.

**Figure 1 Extract of Worksheet S-3, Part 1**

Component	Worksheet A Line No.	No. of Beds	Bed Days Available	CAH Hours	Inpatient Days / Outpatient Visits / Trips				Full Time Equiv	
					Title V	Title XVIII	Title XIX	Total All Patients	Total Interns & Residents	Employees On Payroll
	1	2	3	4	5	6	7	8	9	10
1 Hospital Adults & Peds. (columns 5, 6, 7, and 8, exclude Swing Bed, Observation Bed and Hospice days) (see instructions for col. 2 for the portion of LDP room available beds)										
2 HMO and other (see instructions)										
3 HMO IPF Subprovider										
4 HMO IRF Subprovider										
5 Hospital Adults & Peds. Swing Bed SNF										
6 Hospital Adults & Peds. Swing Bed NF										
7 Total Adults and Peds. (exclude observation beds) (see instructions)										
8 Intensive Care Unit										
9 Coronary Care Unit										
10 Burn Intensive Care Unit										
11 Surgical Intensive Care Unit										
12 Other Special Care										
13 Nursery										
14 Total (see instructions)										

FORM CMS-2552-10 (03-2018) (INSTRUCTIONS FOR THIS WORKSHEET ARE PUBLISHED IN CMS PUB. 15-2, SECTION 4005.1)

The instructions for entering the cost report fields are found in Chapter 40 of the PRM, identified in the documentation section above. To identify the location of the instructions for a specific cost report form or field, see the note at the bottom of each form. In the case of Worksheet S-3, part 1, the instructions are in section **4005.1** of the PRM. Upon opening the PRM, search for section number **4005.1** and scroll through until you get beyond the table of contents and to the instructions for completing this worksheet.

Here are snippets of the instructions for column 2 and row 14 of worksheet S-3, Part I:

**Figure 2 Extract of Instructions for Worksheet S-3, Part 1**

Column 2--Refer to 42 CFR 412.105(b) and 69 FR 49093-49098 (August 11, 2004) to determine the facility bed count. Indicate the number of beds available for use by patients at the end of the cost reporting period.

A bed means an adult bed, pediatric bed, portion of inpatient labor/delivery/postpartum (LDP) room (also referred to as birthing room) bed when used for services other than labor and delivery, or newborn ICU bed (excluding newborn bassinets) maintained in a patient care area for lodging patients in acute, long term, or domiciliary areas of the hospital. Beds in post-anesthesia, post-operative recovery rooms, outpatient areas, emergency rooms, ancillary departments (however, see exception for labor and delivery department), nurses' and other staff residences, and other such areas that are regularly maintained and utilized for only a portion of the stay of patients (primarily for special procedures or not for inpatient lodging) are not termed a bed for these purposes. (See CMS Pub. 15-1, chapter 22, §2205.)

Line 14--Enter the sum of lines 7 through 13 for columns 2 through 8, and for columns 12 through 15, enter the amount from line 1. For columns 9 through 11, enter the total for each from your records. Labor and delivery days (as defined in the instructions for Worksheet S-3, Part I, line 32) must not be included on this line.

## ANALYTICS USING KEY FIELDS

Hospitals submit an annual cost report, usually including twelve months of data, corresponding with their fiscal year (i.e. the 2017 hospital cost report file contains cost reports with a begin date between October 1, 2016 and September 30, 2017). Since hospitals usually submit one cost report per year, the MCR SAS files typically contain one record per hospital.

One of the most valuable aspects of the MCR is that it contains both detailed revenue and expenses for healthcare providers, making it possible to compute a myriad of health care measures from total cost by urban/rural to total revenue by control type (for-profit, nonprofit, government). In addition, the MCR contains hospital characteristics and utilization data, making it possible to compute such measures at an

aggregate level by hospital categories such as state, urban/rural, control type, teaching status, and chain/non-chain.

*Note: The methods and examples below are intended for training purposes only (not official estimates) and meant to demonstrate ways you can use the MCR data for research purposes.*

## HOSPITAL IDENTIFICATION DATA

### PROVIDER NUMBER

The unique identifier for hospitals in the MCR is the six-digit provider number (prvdr\_num). Intelligence is embedded within the provider number. The first two digits identify the state where the hospital is located and the last four digits indicate the type of hospital. See the “Provider Number Intelligence” section of the [DOCUMENTATION](#) table above for a list of state and hospital type codes. Note that if a hospital (provider) submits multiple cost reports for a given year, the report record number (rpt\_rec\_num) can be used as the unique identifier.

### HOSPITAL TYPE

The hospital MCR includes data for all types of hospitals including, but not limited to, acute-care inpatient (IPPS), LTCH, IRF, IPFs and Critical Access Hospitals (CAH). Each type of hospital has different characteristics, reimbursement methods, and cost structures; therefore, it may be necessary to subset the cost report SAS file by type of hospital before performing your analyses.

The code below demonstrates how to subset the SAS datasets by provider type:

```
data mcr2017;
  set hcris.prds_hosp10_yr2017;
  hospital_type = substr(prvdr_num,3,4); * extract hospital type;

  /* Execute the one IF statement that matches the hospital type to be analyzed. */
  if hospital_type <= '0879';          * acute-care inpatient hospitals;
  if '2000' <= hospital_type <= '2299'; * long-term care hospitals (LTCHs);
  if '3025' <= hospital_type <= '3099'; * inpatient rehab facilities;
  if '4000' <= hospital_type <= '4499'; * inpatient psychiatric facilities;

  if substr(prvdr_num,3,2) = '13'      * critical access hospitals1;

run;
```

### STATE

The HCRIS data includes providers from all states and United States (U.S.) territories. A provider's state or territory is identified using the first two digits of the provider number (See the “Provider Number Intelligence” section). In the example SAS code below, we show how to identify and exclude hospitals in Maryland and in the U.S. territories. Both groups have a waiver from the Medicare IPPS and are, therefore, sometimes excluded from acute-care hospital analyses.

---

<sup>1</sup> As a result of the Balanced Budget of 1997 aimed at reducing financial vulnerability and improving access to healthcare by keeping essential services in rural communities, many rural hospitals (which experienced a period of increased closures in the 1980's and 1990's) became designated CAHs. This policy change granted these hospitals exemption from the IPPS method of Medicare reimbursement and offered: reimbursement based on the actual costs of treating Medicare patients instead of based on a fee-for-service (FFS) schedule, completion of an abbreviated cost report, similar to the report completed by no or low Medicare utilization facilities. To be eligible for the CAH designation, hospitals must provide emergency care services 24 hours a day, 7 days a week; have 25 or fewer acute care inpatient beds; maintain an annual average length of stay of 96 hours or less for acute care inpatient patients; be located more than 35 miles from another hospital (unless designated as a necessary provider and be in a rural area and more than a 35-mile drive from another hospital or more than a 15-mile drive from another hospital in an area with mountainous terrain or only secondary roads.



```
data MCR2017;
  set hcris.prds_hosp10_yr2017;
  state = substr(prvdr_num,1,2) ;
  if state not in ('21','40','48' , '64','65','80' );
run;
```

### URBAN RURAL

Worksheet S-2 (part 1, column 1, lines 26 or 27) contains the fields needed to identify a hospital's geographic status. Line 26 defines the geographic status at the beginning of the period and line 27 defines the status at the end of the period. You can use the code below to define geographic status:

```
if S2_1_C1_27 = '1' then urbrur = 'URBAN';
else if S2_1_C1_27 = '2' then urbrur = 'RURAL';

else if S2_1_C1_26 = '1' then urbrur = 'URBAN';
else if S2_1_C1_26 = '2' then urbrur = 'RURAL';
```

### CONTROL TYPE

Worksheet S-2 (part 1, column 1, line 21) contains the field that identifies the hospital's type of control. You can use the code below to define a hospital's control type:

```
if S2_1_C1_21 in (1,2) then CtrlStat = 'Non-Profit';
if S2_1_C1_21 in (3,4,5,6) then CtrlStat = 'Profit';
if S2_1_C1_21 in (7,8,9,10,11,12,13) then CtrlStat = 'Government';
```

### TEACHING STATUS

Worksheet S-3, part 1 contains the fields that identify a hospital's teaching status. The code below defines teaching status, by first computing the interns and residents to beds ratio and then calculating the teaching status (major, minor, none) based on that ratio:

```
if S3_1_C2_27 > 0 then IRsPerBed = S3_1_C9_27/S3_1_C2_27;

if IRsPerBed GE 0.25 then Teach = 'Major Teaching';
if IRsPerBed GT 0 AND IRsPerBed LT 0.25 then Teach = 'Minor Teaching';
if IRsPerBed = . then Teach = 'Non-Teaching';

IME_Pay = E_A_HOS_C1_29;
GME_Pay_PtA = E4_C1_49;
GME_Pay_PtB = E4_C1_50;
```

## UTILIZATION MEASURES

### OCCUPANCY RATE

Hospital occupancy rate is defined as the percentage of hospital beds that are in use at a given time and can be computed by dividing **hospital days** by **hospital bed days available**.

Hospital days are broken out by cost center/row (adults and pediatrics, intensive care) and payer/column (Medicare (XVIII), Medicaid (XIX), Children's Health (V), and Residual (which would mostly reflect private-pay and self-pay patients)). Total hospital days are collected on line 14 and total facility days are collected on line 27.

Bed Days Available are calculated in the MCR as the *number of beds times the number of days in the cost reporting period*. Hospital bed days available are reported on line 14 and total facility bed days available are on line 27.

Hospital Occupancy Rate is computed in the example below:

```
data prds_hosp10_yr2017;
set hcris.Prds_hosp10_yr2017;
if substr(prvdr_num,3,4) <= '0879';
    hospital_days = S3_1_C8_14;
    bed_days_avail= S3_1_C3_14;
    occupancy_rate = hospital_days / bed_days_avail;
run;
```

### **LENGTH OF STAY**

Average hospital length of stay, defined as the average number of days a patient spends in the hospital, is computed by dividing hospital days (S3\_1\_C8\_14) by hospital discharges (S3\_1\_C15\_14), as illustrated below:

```
data mcr2017;
set hcris.prds_hosp10_yr2017;
if substr(prvdr_num,3,4) <= '0879';
    hospital_days = S3_1_C8_14;
    hospital_discharges = S3_1_C15_14;
if hospital_days > 0 and hospital_discharges > 0 then
    average_los = hospital_days / hospital_discharges;
run;
```

### **MEDICARE AND MEDICAID PERCENT OF TOTAL**

The days and discharges fields in Worksheet S-3 can be used to determine which percentages of a hospital's general acute care inpatient business comes from the Children's Health Insurance Program (CHIP) (Title V), Medicare (Title XVIII), Medicaid (Title XIX), and other sources. Days from private health insurance are included in the residual of total hospital days and the sum of the CHIP, Medicare, and Medicaid days. The code below shows how to determine the percent of total days and discharges from CHIP, Medicare and Medicaid:

```
data mcr2017;
set hcris.prds_hosp10_yr2017 (KEEP=prvdr_num S3_1_C6_14 S3_1_C7_14 S3_1_C8_14);
if substr(prvdr_num,3,4) <= '0879';

    CHIP_Days      = S3_1_C5_14;
    Medicare_Days  = S3_1_C6_14;
    Medicaid_Days  = S3_1_C7_14;
    Hospital_Days  = S3_1_C8_14;

    CHIP_Days_Pct   = CHIP_Days/Hospital_Days;
    Medicare_Days_Pct = Medicare_Days/Hospital_Days;
    Medicaid_Days_Pct = Medicaid_Days/Hospital_Days;

    CHIP_Discharges = S3_1_C12_14;
    Medicare_Discharges = S3_1_C13_14;
    Medicaid_Discharges = S3_1_C14_14;
    Hospital_Discharges = S3_1_C15_14;

    CHIP_Days_Pct      = CHIP_Days/Hospital_Days;
    Medicare_Discharges_Pct = Medicare_Discharges/Hospital_Discharges;
    Medicaid_Discharges_Pct = Medicaid_Discharges/Hospital_Discharges;
run;
```

## COST MEASURES

Detailed hospital cost information is found throughout the cost report. Worksheet A contains overhead costs (general service cost centers) and revenue-generating costs (inpatient routine, ancillary, outpatient, other reimbursable, special purpose, and non-reimbursable cost centers), broken out by their salary and non-salary components.

In Worksheet B, the overhead costs are “stepped down” to the revenue-generating cost centers using the appropriate statistics (e.g. accumulated costs, square footage, meals served, and pounds of laundry). Total costs by cost center (reflecting direct costs as well as overhead costs) are available in Worksheet B, column 26. Total facility costs are apportioned to the Medicare program in Worksheet D by cost center.

When analyzing MCR costs (or payments), it is common to compute a cost/payment **ratio** in lieu of a cost/payment **level**. Computing ratios instead of levels is more robust and allows analysts to compare costs or payments across hospitals, regardless of differences in utilization and other factors.

## OVERHEAD COSTS

The first twenty-three lines on Worksheet A represent the overhead cost centers. The following example demonstrates how to compute overhead costs per bed:

```
data mcr2017;
  set hcris.prds_hosp10_yr2017 (keep = prvdr_num A_C1_4 - A_C1_23
                                A_C2_1 - A_C2_23 S3_1_C2_27);

  OH_salary = sum (OF A_C1_4-A_C1_23);      * salary component of overhead;
  OH_nonsalary = sum (OF A_C2_1-A_C2_23);    * non-salary component of overhead;
  total_beds= S3_1_C2_27;
  overhead_per_bed = sum(OH_salary,OH_nonsalary)/total_beds;

run;
```

## REVENUE-GENERATING COSTS

Worksheet A, lines 30-117, contain the revenue-generating cost centers. These are the fields such as adults and pediatrics, intensive care, and operating room that provide patient services resulting in patient revenue.

## COMPENSATION (WAGE INDEX AND WAGE-RELATED COSTS)

Worksheet S-3, Parts II and III contain hospital wage index information and wage-related information, including salaries, contract labor and other wage-related costs, such as employee benefits. Worksheet S-3, part II and III are only required for IPPS hospitals.

The hospital salary costs in Worksheet S-3, part II, column 1, line 1, match the salary costs computed by summing the costs in Worksheet A, column 1. In Worksheet S-3, column 4, salary costs are adjusted for reclassifications identified in Worksheet A-6. Notice the compensation costs are separated into categories like non-physician anesthetists, physician, and teaching costs, allowing you to categorize the costs into PPS-included and PPS-excluded components.

The code below shows one way to compute compensation costs as a percent of total costs:

```
data compensation2017;
  set hcris.prds_hosp10_yr2017;

  if substr(prvdr_num,3,4) <= '0879';

  salary = S3_2_C4_1;
```

```

contract_labor = sum(of S3_2_C4_11- S3_2_C4_16);
benefits = sum(OF S3_2_C4_17-S3_2_C4_25,
               S3_2_C4_2201, OF S3_2_C4_2550- S3_2_C4_2553);

total_compensation =sum(salary, contract_labor, benefits);
total_costs = sum(A C1 200,A C2 200);
comp_pct = total_compensation/total_costs;

run;

```

### ***UNCOMPENSATED CARE COSTS***

Detailed uncompensated care (UCC) costs, including charity care and bad debt (defined below), are located in worksheet S-10:

Charity care includes services for which hospital policies determine the patient is unable to pay.

Bad debt is the unpaid dollar amount for services rendered to a patient or third party payer for which the provider expected payment.

Worksheet S-10 is only applicable to acute-care inpatient hospitals. A portion of UCC costs are reimbursed through UCC payments, established by the Affordable Care Act (ACA). These payments are discussed in more detail below.

## **REVENUE MEASURES**

### ***MEDICARE PAYMENTS***

Medicare payment information can be found in Worksheet E, which has five parts:

- Part A - Inpatient Hospital Services Under PPS
- Part B - Medical and Other Health Services

Worksheet E, Part A contains inpatient payments, including the following:

- Diagnosis Related Group (DRG) payments for inpatient services to Medicare patients
- Outlier payments
- Managed Care (MC) payments
- IME (Indirect Medical Education) payments to teaching hospitals
- DSH (Disproportionate Share) payments to hospitals that treat a disproportionate share of indigent patients
- UCC (Uncompensated Care) payments, established by the ACA to help offset a mandated reduction to Medicare DSH payments
- Special add-on payments for new technologies
- Extra payments to hospitals that treat a high percentage of ESRD beneficiaries
- Graduate Medical Education (GME) Part A and Part B payments.

The following example depicts how to compute the mean and median DRG, outlier, MC, IME, DSH, and extra ESRD payments:

```
data payments2017;
  set hcris.prds_hosp10_yr2017 (keep= prvdr_num
                                E_A_HOS_C1_1 E_A_HOS_C1_101-E_A_HOS_C1_104
                                E_A_HOS_C1_2 E_A_HOS_C1_201 E_A_HOS_C1_202
                                E_A_HOS_C1_3 E_A_HOS_C1_29 E_A_HOS_C1_34
                                E_A_HOS_C1_36 E_A_HOS_C1_54
                                E4_C1_49 E4_C1_50);

  Hospital_Type = SUBSTR(prvdr_num,3,4);
  if Hospital_Type <= '0879';      * subsets acute-care inpatient hospitals;
    DRG_Pay = SUM (E_A_HOS_C1_1, OF E_A_HOS_C1_101-E_A_HOS_C1_104);
    Outlier_Pay = SUM (E_A_HOS_C1_2, E_A_HOS_C1_201, E_A_HOS_C1_202);
    Mngd_Care_Pay = E_A_HOS_C1_3;
    IME_Pay = E_A_HOS_C1_29;
    DSH_Pay = E_A_HOS_C1_34;
    UCC_Pay = E_A_HOS_C1_36;
    New_Tech_Pay = E_A_HOS_C1_54;
    GME_Pay_A = E4_C1_49;
    GME_Pay_B = E4_C1_50;

  keep prvdr_num DRG_Pay Mngd_Care_Pay Outlier_Pay IME_Pay DSH_Pay UCC_Pay
      New_Tech_Pay GME_Pay_A GME_Pay_B;

run;

proc summary data= payments2017;
  var DRG_Pay Mngd_Care_Pay Outlier_Pay IME_Pay DSH_Pay UCC_Pay New_Tech_Pay
      GME_Pay_A GME_Pay_B;

  * The dataset Payments_ds will contain one record with the mean and median of
    each of these payment types;
  output out=Payment_ds mean= median= ;
run;
```

## FINANCIAL WORKSHEETS

The G Series of worksheets contains the following hospital financial statements:

- Balance Sheet (Worksheet G)
- Statement of Revenue and Expenses (Worksheet G-3).

One of the key measures computable from Worksheet G-3 is the total facility profit margin, which can be used as a tool to analyze the financial health of a hospital. The following SAS code shows how to compute a total facility profit margin for individual hospital, directly from the variables in Worksheet G-3:

```
total_revenues = sum(G3_C1_3,G3_C1_25);
total_expenses = sum(G3_C1_4,G3_C1_30);
total_margin = sum(total_revenues-total_expenses)/total_revenues;
```

Profit margins can also be calculated at an aggregate level as equally-weighted and revenue-weighted as defined below:

- Equally-weighted margin is computed by taking the mean of the individual hospital margins. Hospitals of all sizes contribute equally to this measure.
- Revenue-weighted margin is calculated by summing the revenues and expenses for all hospitals and computing the margin using the sums of the revenues and expenses. Larger hospitals contribute more to this measure.

The code below computes a total facility equally-weighted margin (total\_margin\_mean):

```
proc means data=margins;
var total_margin;
output out=mean_margin_dataset mean=total_margin_mean;
run;
```

This code computes a revenue-weighted total facility margin (total\_margin\_wgt) on an aggregate level:

```
proc summary data=margins;
var total_revenues total_expenses;
output out=sumrevexp sum=;
run;
data revwgtmargin;
set sumrevexp;
total_margin_wgt = sum(total_revenues-total_expenses)/total_revenues;
run;
```

The MCR also contains the necessary data to compute profit margins for overall Medicare, inpatient Medicare, outpatient Medicare, and hospital-based units. Computing Medicare margins is more complex than computing total facility margins since, unlike total facility margins, there are not specific variables for Medicare revenues and Medicare expenses. There are multiple Medicare payment variables in the E series of worksheets and many Medicare cost variables throughout the D series of worksheets. Thus, there are numerous ways to measure Medicare margins by summing the appropriate Medicare costs from worksheet D and payments from worksheet E. To compute accurate Medicare margins, it is crucial that you properly include the costs to treat Medicare patients and the corresponding Medicare payments associated with reimbursing those costs.

## CATEGORICAL DATA ANALYSIS

Worksheet S-2 contains information about hospital characteristics that can be used to analyze cost, payment, and utilization measures for different categories of hospitals, such as urban/rural, control type (e.g. for-profit/nonprofit), teaching status, and bed size, as identified above. Adding a CLASS statement or BY statement to a SUMMARY, MEANS, or UNIVARIATE procedure produces the statistical measure grouped by the specified categorical data variable.

The example below demonstrates how to compute the average IME payments by teaching status (defined above). The output dataset, IME\_by\_Teach, contains four records of IME payments, one for each teaching status (major teaching, minor teaching, nonteaching) and one for all hospitals.

```
proc means data= payments2017;
    var IME_Pay;
    class Teach;
    output out=IME_by_Teach mean=;
run;
```

The example below shows how to compute average salary per FTE by control type (defined above). The output dataset, SumSalaryFTE, contains four records of total salaries and total FTEs, one for each control type (for profit, not for profit, government) and one for all hospitals. The output dataset, Salary\_per\_FTE, contains four records of average salary per FTE one for each control type (for-profit, nonprofit, government) and one for all hospitals.

```
proc summary data= mcr2016;
    var Salary Facility_FTEs;
    class CtrlStat;
    output out=SumSalaryFTE sum=;
run;

data SalarybyControl_Wgt;
    set SumSalaryFTE;
    Salary_per_FTE = Salary/Facility_FTEs;
run;
```

## TOTAL FACILITY AND MEDICARE COSTS

The MCR collects total facility costs (after overhead allocation) by cost center in Worksheet B. A share of the total facility costs are apportioned to the Medicare program in Worksheet D. Costs for routine services are apportioned using **total cost per day** times **Medicare days** and costs for ancillary services are apportioned using **total cost to charge ratio** times **Medicare charges**.

## AUDITS AND EDITS

There are standard edits applied to MCRs at submission to validate certain cost report fields:

- Verify that provider number, date, and categorical fields have valid values.
- Make sure the detail data correctly adds to the totals.

In addition to standard edits, a minimal amount of cost reports and fields are audited. The MCR variable, report status code (rpt\_stus\_cd) indicates whether or not a cost report has been audited. As shown by the table below, only a small percentage (approximately 2%) of cost reports are audited:

**Table 7 Report Status Code Distribution**

Report Status Code Values	2014 (%)	2015 (%)	2016 (%)
1 – As Submitted	23	20	75
2 – Settled without Audit	58	49	10
3 – Settled with Audit	2	2	.08
4 – Reopened	1	2	.08
5 – Amended	15	27	15

Of the small percent of audited cost reports, only cost-reimbursed fields (e.g. organ acquisition, nurse anesthetist, and critical access hospital fields) are actually reviewed. As expected the number of cost reports audited, reopened and amended increases as the cost reports mature.

### **OUTLIER TRIMMING**

Since the MCR auditing and HCRIS editing procedures do not eliminate all unreasonable MCR data, it is a good practice to apply your own edits specific to the data needed for each project. This is particularly helpful if you are analyzing small groups of hospitals (e.g. a state with a few hospitals) where a provider's unreasonable data has a higher probability of skewing your end results. Possible reasonableness edits include setting threshold ranges for variables of interest or trimming high and low outliers based on selected criteria.

The code below demonstrates how to apply an outlier trim, excluding the top and bottom 5 percent outliers based on a key field used in computing a specific measure:

```
data temp;
    set payments2017;
    if totrev GT 0 AND totexp GT 0;
        tottrim=LOG(totrev/totexp);
        alike = 1;
run;
proc univariate data=temp;
    var tottrim;
    output out=trim_limits_2 P5=totlow P95=tothigh;
run;

data trim_limits_2;
    set trim_limits_2;
    alike=1;
run;

proc sort data=temp; by alike; run;
proc sort data=trim_limits_2; by alike; run;
```



```

data trimmed_file_2;
merge temp trim_limits_2;
by alike;
if tottrim > totlow AND tottrim < tothigh;
    totmarg=SUM(totrev,-totexp)/totrev;
run;

```

Since the cost report data set is fairly robust, analysts find that computing averages on a national level, or across larger groups of providers, yields reasonable results. In addition, using relationship metrics (i.e. profit margins, compensation costs as a percent of total costs) instead of data levels, is a more effective way to analyze MCR data.

### ***HANDLING MULTIPLE REPORTS FOR A PROVIDER***

Although hospitals generally submit one cost report per year, occasionally they submit multiple cost reports in a single year, typically under one of two scenarios: (1) each cost report covers part of the year or (2) the second submission replaces the first submission. Both scenarios result in records with duplicate provider numbers. Options for addressing such cases are discussed below:

- The first scenario is due to hospitals submitting their data in partial year increments through separate report submissions in the same year. To identify such “short-period” cases, compute the report period using the variables for fiscal year begin date (FY\_BGN\_DT) and fiscal year end date (FY\_END\_DT). Two options for handling this scenario follow:
  - Combine the reports, summing the appropriate fields, such as costs and revenue, to obtain a full year of data. Be cautious not to add certain numeric fields such as beds and FTEs since they are not meant to be accumulated.
  - Because very few providers submit partial years, simplify the process, only keeping self-defined full year” cost reports. For our purposes, we define a “full year” as having ten to fourteen months. (Note: Reports covering more than twelve months of data are rare, but possible.)
- The second scenario is often a result of a hospital updating or correcting the data initially reported. When this happens, it is common to use the most recently submitted (and likely most up-to-date) report and drop the old report to avoid double counting. The code below subsets your file only keeping reports with a “full year” of data (self-defined as 10-14 months of data), using the fiscal year begin date and end date variables:

```

data MCR2017;
  set prds_hosp10_yr2017;
  report_days = sum(FY_End_Dt,-FY_Bgn_Dt);
  if 300 < report_days < 420;
run;

```

SAS dataset, MCR2017 will contain all “full-year” cost reports with 10-14 months of data. You can keep the most recent cost report using this code:

```

proc sort data=hosp2017; /* Sorts by provider number and report record number.
  by prvdr_num DESCENDING rpt_rec_num;
run;

proc sort data=hosp2009 nodupkey;
  by prvdr_num; /* Keeps the record with the highest report record number. */
run;

```

## OUTPUTTING YOUR RESULTS TO EXCEL

You can use the Output Delivery System (ODS) to send your results such as procedure output and reports to Excel. The results of the code that you place between the two ODS statements is output to the Excel file identified in the FILENAME statement. The code below uses the default formatting, but there are many formatting options available to customize your output if you edit the ODS templates:

```
filename MCR_out "C:\UserFiles\MCR.xls";
```

```
ods html file=MCR_out;
```

```

proc print data= payments2017 NOOBS SPLIT = ' ';
  var DRG_Pay Mngd_Care_Pay Outlier_Pay;
  label DRG_Pay= 'Diagnosis Related Group Payments'
        Mngd_Care_Pay = 'Medicare Costs'
        Outlier_Pay = 'Outlier Payments';
  title 'Hospital Payments';
  footnote 'Source: Medicare Hospital Cost Reports' ;
run;

```

```
ods html close;
```

## CONCLUSION

Medicare cost reports contain a wealth of information used by policymakers, healthcare analysts, and researchers. The MCRs are a comprehensive data source for provider characteristics, utilization (i.e. days, discharges, FTEs), detailed expenses (i.e., adults and pediatrics, skilled nursing facility) and revenues (inpatient and outpatient), making it possible to compute a myriad of measures on a national level; by state; by hospital categories (such as urban/rural, control type, state/region, teaching status, chain status, and bed size); and for individual hospitals.

## REFERENCES

Andrews, K. and Andrews, R. 2012. "SAS® Analytics for the Healthcare Cost Reporting Information System (GH-08)." Alexandria, VA: Southeast SAS Users Group.

Rural Health Information Hub. "Critical Access Hospitals (CAHs)." Accessed September 1, 2018. <https://www.ruralhealthinfo.org/topics/critical-access-hospitals>

Organization for Economic Cooperation and Development (OECD). "Length of Hospital Stay." Accessed September 1, 2018. <https://data.oecd.org/healthcare/length-of-hospital-stay.htm>

Statista. "Hospital occupancy rate in the U.S. from 1975 to 2014." Accessed September 1, 2018. <https://www.statista.com/statistics/185904/hospital-occupancy-rate-in-the-us-since-2001/>

## ACKNOWLEDGEMENTS

The author wishes to thank Rick Andrews and Mollie Knight for providing helpful comments on this paper.

## RECOMMENDED READING

- *Medicare Payment Advisory Commission (MEDPAC) March 2018 Report to Congress, Chapter 3. "Hospital Inpatient and Outpatient Services."* [www.medpac.gov](http://www.medpac.gov)

## CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Kimberly Andrews

Centers for Medicare and Medicaid Services

(410) 786-0516

[Kimberly.andrews@cms.hhs.gov](mailto:Kimberly.andrews@cms.hhs.gov)

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.