



## Center for Strategic Planning

---

May 18, 2012

Re: MCBS SAS® Claims Datasets

Dear MCBS User:

This notification is for Medicare Current Beneficiary Survey (MCBS) users who currently use or have worked with the **Medicare claims SAS® datasets** sent out with the Access to Care and Cost and Use MCBS research files. If you use MCBS claims data in flat file format you can disregard this notification. (*Note: This memorandum applies to MCBS files purchased and shipped to users prior to April 2012. If you have received data from CMS following this date, this notification does not apply to these data.*)

MCBS staff has been made aware of, and since confirmed, a data anomaly present in .SAS claims datasets for the following years of MCBS data (see Table 1). ***This issue impacts .SAS claims data only, and is not present in the flat file (.dat) claims also released as part of the annual MCBS research file.***

**Table 1. Files and Years Impacted by Data Anomaly**

<u>Access to Care</u>	<u>Cost and Use</u>
1992 - 2009	1992-2000, 2003-2008

### ***About the Data Anomaly***

Missing diagnosis codes have been identified (for .SAS claims datasets only) in six of the seven claims file types (e.g., inpatient, hospice, DME, etc.). **The carrier/physician claims are not impacted.** These codes are captured in the following claims variables **DGNSCD1- DGNSCD9**, and **LINEDGNS** (for DME claims). MCBS staff has attributed the missing diagnosis codes to a failure in our computer code (used to read raw claims data into SAS) to handle variable formatting in the claims. It is important to note that because of the source of this error, MCBS users can still use the flat file (.dat) claims data without experiencing missing data. (See “Solution: Working with Flat File Claims Data” below.)

### ***Scope of Underreporting of Diagnosis Codes***

**Attachments 1** (Access to Care) and **2** (Cost and Use) to this memo present the average percent of underreporting for each claims type across several years of MCBS data. The number of missing diagnoses, or underreporting compared to data that was complete, is consistent across

the years of MCBS files; therefore we report averages rather than each year independently. In general, underreporting is highest for the first diagnosis code (DGNSCD1) and tends to trend down with each subsequent diagnosis.

There are a few claims types that show significantly more underreporting than others, especially Hospice (HSP), Outpatient (OTP) and DME claims. For example, the Cost and Use hospice claims show 79% underreporting for DGNSCD1, followed by 64% for DME and 32% for outpatient. These claims types (i.e. hospice, outpatient and DME) tend to have less secondary diagnosis codes reported on the claim, therefore the impact of this error is more substantial on the first diagnosis reported.<sup>1</sup>

Most other claims types show underreporting for the first diagnosis (DGNSCD1) ranging from 1% to as high as 13%. This underreporting in most cases becomes minimal (less than 5%) by DGNSCD6.

***Solution: Working with Flat File Claims Data***

The most efficient and convenient way for an MCBS user to address this data anomaly is to not use the SAS files sent out and convert the claims flat files (.dat) into SAS. CMS has provided instructions on how to do this in our documentation that is sent with the MCBS files (see *Section 1: File Structure* in documentation). In addition, CMS provides “README” files which include a SAS INPUT statement, a PROC FORMAT (to interpret the coded fields), and LABELS.

When using the claims flat files, MCBS users should be sure to include (as part of the INFILE statements in SAS) the “truncover” option. This option enables SAS to deal with some formatting abnormalities in CMS claims data that caused these diagnosis variables to be reported as missing. For more information on how to incorporate this option into your SAS code, see the following link:

<http://support.sas.com/documentation/cdl/en/basess/58133/HTML/default/viewer.htm#a002645812.htm>).

On behalf of MCBS staff, we apologize for any inconvenience this issue may have caused and thank our users for continuing to provide valuable feedback on the MCBS survey and data products. If you have questions related to this issue please contact CMS’ Research Assistance Data Center (ResDAC) at [resdac@umn.edu](mailto:resdac@umn.edu) or 1-888-9RESDAC.

Sincerely,

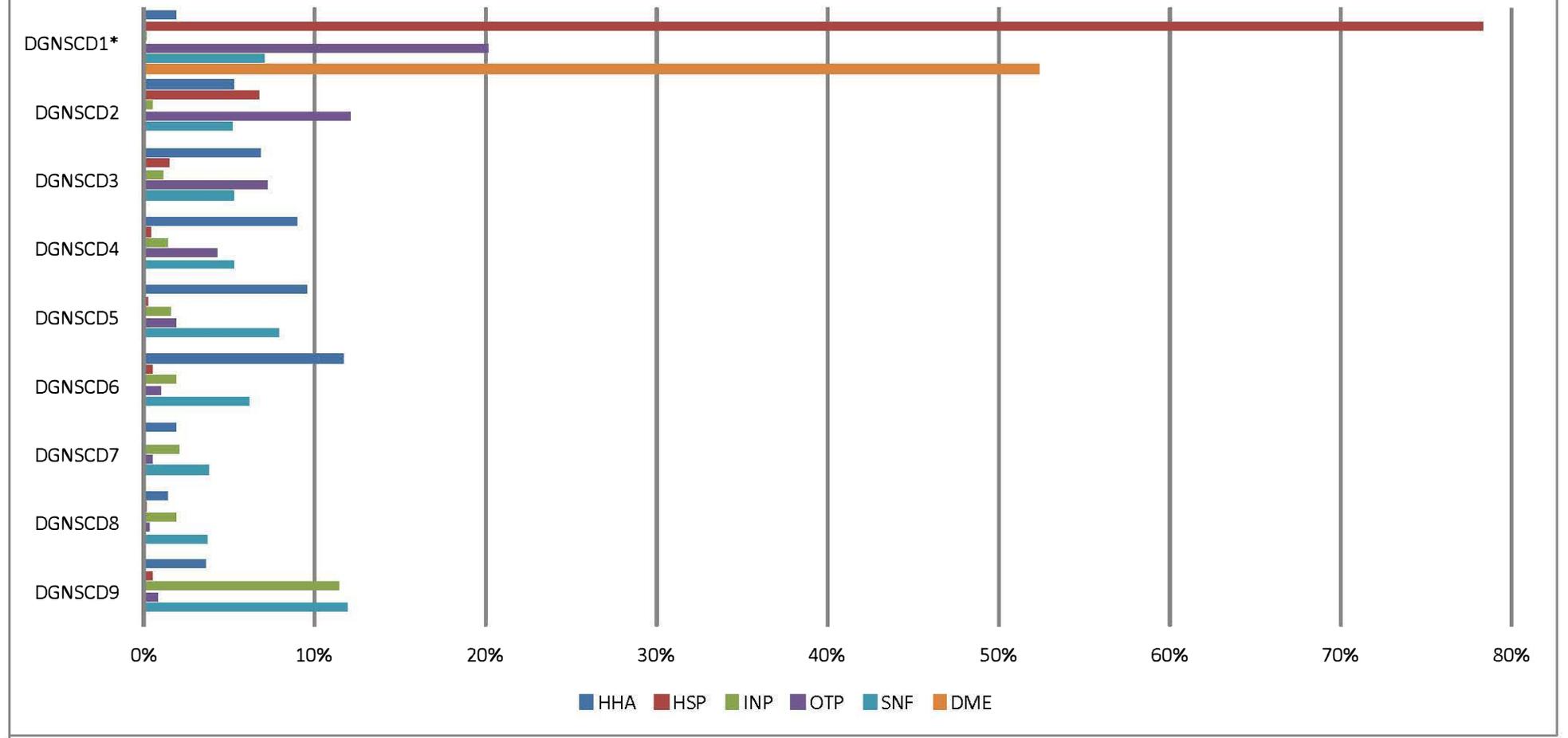
Jason Petroski, Ph.D., MPA  
Director, Division of Survey Management and Analysis

---

<sup>1</sup> Unlike other claims types, DME claims only have one diagnosis code variable in MCBS claims files (LINEDGNS).

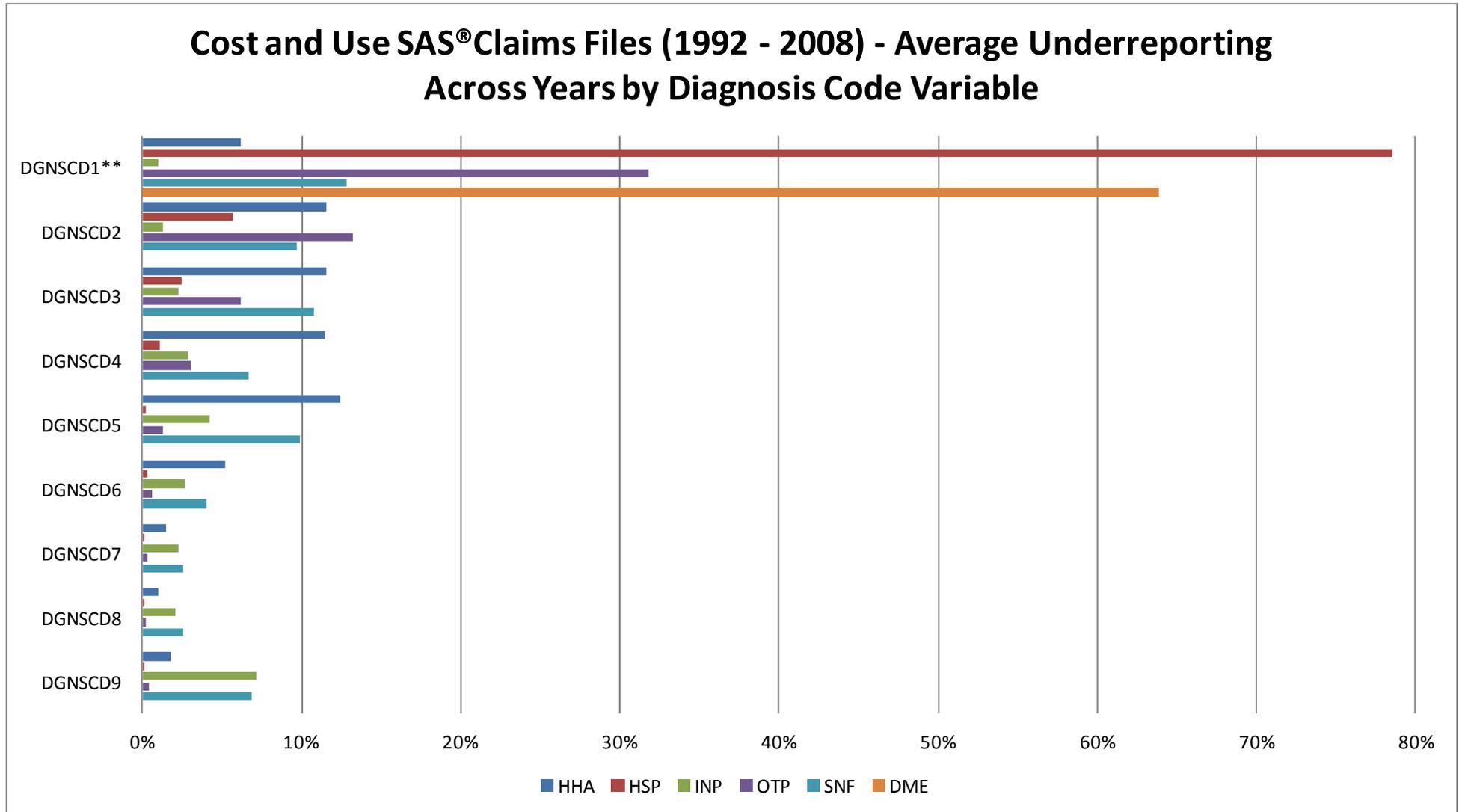
**Attachment 1**

## Access to Care SAS® Claims Files (2004-2009) - Average Underreporting Across Years by Diagnosis Code Variable



**Notes:**  
 \*LINEDGNS variable for DME claims type captured graphically under DGNSCD1 line.

**Attachment 2**



**Notes:**

\*\*LINEDGNS variable for DME claims type captured graphically under DGNSCD1 line