

Complicated Data? SAS User Formats Will Simplify Your Code

April 8 – 11 | Denver, CO **#SASGF**

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.

SAS[®] GLOBAL FORUM 2018 USERS PROGRAM



Complicated Data? SAS User Formats Will Simplify Your Code Margaret Burgess and Stephanie Nelson Optum

ABSTRACT

The International Classification of Diseases (ICD) is used for medical coding in the U.S. It includes medical diagnosis coding and inpatient hospital procedure coding. On October 1, 2015, health care professionals changed from using the 9th revision (ICD-9) to the 10th revision (ICD-10). When analyzing administrative claims data for a research study that includes with both ICD-9 and ICD-10 codes and a time period that spans over October 1, 2015, we have identified three main areas that present challenges. Below are examples of these three challenges and our suggested solutions to overcome them.

CHALLENGE 1

The same diagnosis code value can appear in ICD-9 and ICD-10 with different meanings. An example is shown in Table 1 below.

TABLE 1

Code	Version	Definition
E8981	ICD-9	ACCIDENT CAUSED BY OTHER BURNING MATERIAL
E8981	ICD-10	POSTPROCEDURAL HEMORRHAGE OF AN ENDOCF
		FOLLOWING A PROCEDURE
V293	ICD-9	OBSERVATION FOR SUSPECTED GENETIC OR META
V293	ICD-10	MOTORCYCLE RIDER (DRIVER) (PASSENGER) INJUR

CHALLENGE 2

Specifying a range of ICD procedure codes can return inaccurate results. Procedure codes, which are character values, are sorted by SAS as character values. Consequently the ICD-9 range '0015' – '0016' shown in Table 2, would now include a long list of ICD-10 procedure codes, some of which are displayed in Table 3.

CRINE SYSTEM ORGAN OR STRUCTURE

ABOLIC CONDITION **RED IN UNSPECIFIED NONTRAFFIC ACCIDENT**

Code	Version	De
0015	ICD-9	HIC
0016	ICD-9	PR SU
		30
0017	ICD-9	INF

Version	Def
ICD-10	BYF
	PEF
ICD-10	BYF
	PEF
ICD-10	BYF
	SUE
	ICD-10 ICD-10

A claims analysis study that includes both ICD-9 and ICD-10 codes can lead to overly complicated programming logic. This complicated logic includes "if-then-else" analysis along with the interrogation of code versions, to insure proper interpretation of the claims data.

If diag_ver_cd_1='9' then do.....; end; Else if diag_ver_cd_1='10' then do....; end;

CHALLENGE 2 (continued)

TABLE 2

efinition

GH-DOSE INFUSION INTERLEUKIN-2 [IL-2]

ESSURIZED TREATMENT OF VENOUS BYPASS GRAFT [CONDUIT] WITH PHARMACEUTICAL BSTANCE

FUSION OF VASOPRESSOR AGENT

TABLE 3

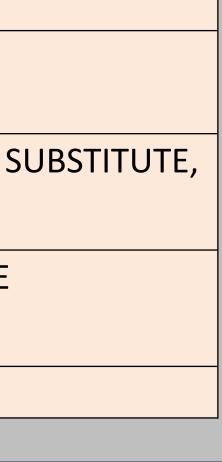
finition

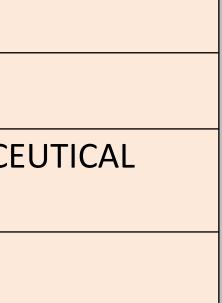
PASS CEREBRAL VENTRICLE TO NASOPHARYNX WITH SYNTHETIC SUBSTITUTE, **RCUTANEOUS ENDOSCOPIC APPROACH**

PASS CEREBRAL VENTRICLE TO PERITONEAL CAVITY WITH AUTOLOGOUS TISSUE SUBSTITUTE, **RCUTANEOUS APPROACH**

PASS CEREBRAL VENTRICLE TO URINARY TRACT WITH NONAUTOLOGOUS TISSUE BSTITUTE, OPEN APPROACH

CHALLENGE 3





Complicated Data? SAS User Formats Will Simplify Your Code Margaret Burgess and Stephanie Nelson Optum

METHODS

Our first suggestion is to derive a new variable and add it to your analytic SAS dataset. This is illustrated below in Table 4 with a diagnosis code. This new variable value holds the original diagnosis code, prefixed with a version identifier.

TABLE 4

Version	Original Diagnosis Code	New Variable – Diagn
ICD-9	E8981	D09_E8981
ICD-10	E8981	D10_E8981

Creating this new variable now allows the programmer to create one SAS user Format, and bypass the more complicated logic required using the correct SAS user format with the corresponding diagnosis code version.

```
PROC FORMAT;
VALUE $MYFMT
'D09_E8981' = 'ACCIDENT CAUSED BY OTHER BURNING MATERIALS'
'D10_E8981' = 'POSTPROCEDURAL HEMORRHAGE OF AN ENDOCRINE SYSTEM ORGAN OR
STRUCTURE FOLLOWING A PROCEDURE'
'D09_V293' = 'OBSERVATION FOR SUSPECTED GENETIC OR METABOLIC CONDITION'
'D10 V293' = 'MOTORCYCLE RIDER (DRIVER) (PASSENGER) INJURED IN UNSPECIFIED NONTRAFFIC
ACCIDENT
'P09_0015'= 'HIGH-DOSE INFUSION INTERLEUKIN-2 [IL-2]'
'R_0499' = 'AMBULATORY SURGICAL CARE-OTHER AMBULATORY SURGICAL CARE'
'H_S0187' = 'TAMOXIFEN CITRATE, ORAL'
```

nosis Code with a Version Prefix

In other words, this:

```
if code_type = "9" then do;
category = put(diag, $fmtnine.);
end;
else if code_type = "10" then do;
category = put(diag, $fmtten.);
end;
```

Becomes this simpler logic without the need for if –then –else logic:

category = put(diag, \$myfmt.);

SAS and all other SAS Institute Inc. product or service names are registered 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.

METHODS (continued)



Complicated Data? SAS User Formats Will Simplify Your Code Margaret Burgess and Stephanie Nelson Optum

<u>RESULTS</u>

1. Using a combination of a new variable (code value with prefixed version) along with a SAS user format removes any ambiguity in interpreting claims data containing a mix of ICD-9 and ICD-10 codes.

2. Analyzing claims data using ranges of alpha numeric values across versions can still be used without the risk of including unwanted values in your specified ranges.

3. Incorporating the version along with the code value, in a new variable, reduces the complexity of SAS coding required.

SAS user formats can be used to:

- Simplify programming logic

CONCLUSION

• Remove ambiguity in code sets that vary across versions

• Generally make writing SAS programs and understanding SAS programs used to analyze claims data contain a mix of codes and code versions, much easier.





SAS and all other SAS Institute Inc. product or service names are registered trademarks of SAS Institute Inc. in the USA and other countries. [®] indicates USA registration. Other brand and product nes are trademarks of their respective cor



SAS[®] GLOBAL FORUM 2018

April 8 – 11 | Denver, CO Colorado Convention Center

