

Utilization of SAS® Input Statements for the Barell Injury Diagnosis Matrix to Categorize Senior Sports Related Injury by Body Region and Nature of Injury

Kassandra Lynn McGlonn, M.S, Florida A&M University

ABSTRACT

The Barell Injury Diagnosis Matrix is used by a wide variety of health practitioners and researchers worldwide to identify, analyze, and communicate injury information. In this paper, we will present a thorough investigation of Florida statewide emergency department (ED) data for senior sports injuries using SAS® input statements that categorized injuries by body region and nature of injury per the Barell Injury Diagnosis Matrix for ICD-9-CM codes. This paper will detail practical use, benefits, limitations, and ease of use for Barell Matrix SAS® input statements.

INTRODUCTION

Researchers looking to statistically analyze health care datasets are challenged with identifying how best to organize the data. This can be an intricate and time-consuming task depending on the study variable(s), number of observations, and the way the dataset must be cleaned to utilize it. However, when the research is specific to injuries presented to hospitals, there are tools, such as SAS input statements for the Barell Injury Diagnosis Matrix, available to make your analysis process more efficient.

This paper will discuss how you can utilize the Barell Injury Diagnosis Matrix SAS Input statements available through the CDC's ICE on Injury Statistics to categorize, stratify, and analyze data by body region and nature of injury.

BARELL INJURY DIAGNOSIS MATRIX AND SAS

The Barell Injury Diagnosis Matrix, developed by Vita Barell and Ellen MacKenzie, is a tool available through the Centers for Disease Control's (CDC) National Center for Health Statistics under Injury Data and Resources. As a product of the International Collaborative Effort (ICE) on Injury Statistics, it currently is based on 3-digit ICD-9-CM codes and is valid for years 1998 and beyond. International Classification of Disease (ICD) codes are used by health institutions across country lines to identify and communicate health conditions. With the vast amount of injuries that present annually to hospitals, this matrix is a very efficient tool for communicating injury occurrences. This matrix has become a widely-used standardized method for the identification and billing of injuries presented in health care settings.

The Barell Matrix was chosen to be used for this project because researchers wanted a way to easily extract and analyze injury data from the datasets, which included injury and non-injury reasons individuals reported to emergency departments. This matrix allowed researchers to focus on the ICD-9-CM codes that were exclusive to injuries. Furthermore, it served as a tool to focus on injuries that occurred in lower extremities as the specific body region of interest.

SAS is an international statistical analysis tool that many researchers across different industries regularly employ for personal and professional analysis projects. The development of SAS input statements for the Barell Injury Diagnosis Matrix has made processing injury related data more efficient and accurate for researchers, as SAS input statements allow for faster analysis of different types of injuries according to their assigned ICD-9-CM code. Below is an example of how the ICD-9-CM codes are paired in input statements to specify certain injuries;

```
IF DX13 GE '900' AND DX13 LE '904' THEN ISRCODE=7;
IF DX13 GE '910' AND DX13 LE '924' THEN ISRCODE=8;
IF DX13 GE '925' AND DX13 LE '929' THEN ISRCODE=9;
```

```
IF DX13 GE '940' AND DX13 LE '949' THEN ISRCODE=10;
```

WORKING WITH THE BARELL INJURY DIAGNOSIS MATRIX SAS INPUT STATEMENTS

PREPARING YOUR DATA FOR USE

The Barell Matrix input statements include many lines of code that must be added to your code. To successfully do this, researchers must be familiar with their own code before the two are combined. Datasets can have diagnosis codes written with different character features that may not be compatible with these input statements. In the study conducted related to this paper, emergency department data from the Agency for Health Care Administration (ACHA) was used for years 2010-2014. This data included decimals in their diagnosis codes. Decimals are not recognized in the SAS input statements for the Barell Matrix code and would have resulted in observations not being recognized. To address this, an array was used with compress and tranwrd functions to remove the decimals;

```
data ssip.sri_uni4;
set ssip.sri_uni3;
array newdiag prindiag;
do over newdiag;
newdiag = compress(tranwrd(newdiag, '.', ' '));
end;
run;
```

ADDING SAS INPUT STATEMENTS TO YOUR CODE

The data was categorized to match the Barell Injury Diagnosis Matrix through use of SAS input statements via SAS 9.4. The input statements used were found on the CDC's ICE on Injury Statistics web page as code that can be copied and pasted from the web page directly into the SAS program of use. The statements are compatible with multiple versions of SAS. The input statements specify how to use the code and provides updates of recent changes;

```
/*SAS Input Statements for External Cause of Injury Morbidity Matrix */
/*Following are the SAS statements containing ICD-9 CM codes to be used to
program the external cause of injury morbidity matrix.
```

Please cut and paste the SAS codes provided into your SAS programs.

The following 2 modifications were made on March 15, 2007: 1) E979.6 and E979.7 were moved to the Poisoning, assault category and removed from

"Other specifiable and classifiable" to be consistent with the placement of similar terrorism codes used in ICD-10 external cause mortality matrix, and 2) E999 was separated into E999.0 and E999.1 with E999.1, "late effects of injury due to terrorism" moved to the assault category where the other terrorism codes, E979.0-.9, are located from the "Other" category.

On Jan 29, 2009, the SAS code has been checked against the December 1, 2008 updated Matrix. (accessed on line at <http://www.cdc.gov/ncipc/osp/matrix2.htm>) The following modification was made: E928.6 "Environmental exposure to harmful algae and toxins" was added to other specified and classifiable. Variable names used in the SAS statements

CAUSEDT: 3 DIGIT ICD CODE fields for morbidity data

CAUSEDT4: 4 DIGIT ICD CODE field for morbidity data

PERSINJ: 4TH DIGIT OF THE ICD CODE */

After the SAS input statements identify all injuries found in the dataset, now you can continue with using the tool. It is recommended you next clean the dataset to allow for only the injuries of interest that occurred within the population of interest to remain under analysis. In this study's case, the only observations that remained in the dataset occurred among seniors aged 65 years and older that presented to an emergency department within the state of Florida that were diagnosed with a sports-related injury to lower extremities.

There are many options for epidemiological and clinical analysis by using the matrix's input statements as a research tool. For example, you can use these input statements to stratify data by body region the injury occurred in and nature of the injury; comparisons can be made to identify how likely certain types of injuries are to occur between different demographic variables. Data can be collapsed to include multiple injuries associated with specific activities. Furthermore, the implications of this research allow health professionals to better develop interventions, plan and report on related financial issues, and identify emerging trends in injury. During this study, trends were studied over a 5-year period for various demographic variables. Figure 1 is an example of statewide injuries reported to emergency departments for the variable of sex.

Frequency Percent Row Pct Col Pct	Table of SEX by YEAR						
	SEX(SEX)	YEAR(YEAR)					Total
		2014	2013	2012	2011	2010	
M		42295	42427	40546	39023	38922	203213
		15.60	15.65	14.96	14.40	14.36	74.97
		20.81	20.88	19.95	19.20	19.15	
		73.74	74.67	75.05	75.52	76.03	
F		15064	14392	13476	12647	12273	67852
		5.56	5.31	4.97	4.67	4.53	25.03
		22.20	21.21	19.86	18.64	18.09	
		26.26	25.33	24.95	24.48	23.97	
Total		57359	56819	54022	51670	51195	271065
		21.16	20.96	19.93	19.06	18.89	100.00

Figure 1. Demographic Trend Analysis Example

CONCLUSION

SAS input statements for the Barell Injury Diagnosis Matrix should continue to be a well-known tool for any researcher working with injury diagnoses-related health care data. It allows for more efficient and accurate categorization, stratification, and analysis of diagnosed injuries. These SAS input statements are readily accessible at no cost and includes clear instructions that allow for easy use. This may arguably become the gold-standard method for analyzing such data in SAS.

REFERENCES

Agency for Health Care Administration, Data Summaries and Reports. (2018, June 12). Retrieved June 12, 2018, from <http://www.floridahealthfinder.gov/researchers/QuickStat/quickstat.aspx>

The Barell Injury Diagnosis Matrix, Classification by Body Region and Nature of the Injury. (2010, June 24). Retrieved June 12, 2018, from https://www.cdc.gov/nchs/injury/ice/barell_matrix.htm

CONTACT INFORMATION

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

Kassandra Lynn McGlonn, M.S.
Florida A&M University, Florida Department of Health
Kassandra.mcglonn@flhealth.gov
Kassandra1.mcglonn@famu.edu

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.