Paper 3327-2015

Automated Macros to Extract Data from the National (Nationwide) Inpatient Sample (NIS)

Ravi Gaddameedi, California State University, Eastbay, CA; Usha Kreaden, Intuitive Surgical, Sunnyvale, CA

ABSTRACT

The use of administrative databases for understanding practice patterns in the real world has become increasing apparent. This is essential in the current healthcare environment. The Affordable Healthcare Act has created the need to probe further into the current usage of technology and into the use of different approaches to surgery. This paper describes a method for extracting surgical procedure-specific information from the Healthcare Cost and Utilization Project (HCUP) database, otherwise referred to as the National (Nationwide) Inpatient Sample (NIS). Using this systematic method in SAS® 9.2, you will be able to extract data from NIS using the ARRAY statement for the specific ICD-9 codes, to format the extracted data for analysis, to merge the different NIS databases by procedures, and to use automatic macros to generate a report. These analyses provide a framework for comparing key outcome variables for different modalities of surgery for procedures of interest.

METHODS

Using an NIS database for a single year, you will identify cohorts based on surgical approach. You do this by identifying the ICD-9 codes specific to robotic surgery, laparoscopic surgery, and open surgery. After you identify the appropriate codes using an ARRAY statement, a similar ARRAY is created based on the ICD-9 codes. Any minimally invasive procedure (robotic or laparoscopic) that results in a conversion is flagged as a conversion. Comorbidities are identified by ICD-9 codes representing the severity of each subject and are merged with the NIS core file. Using a FORMAT statement for all diagnosis variables, you create macros that can be regenerated for each type of complication. These created macros are compiled in SAS and are stored in the library that contains the four macros that are called by tables. You call the macros for different macros variables. In addition, you create the frequencies of all cohorts and create the table structure with the title and number of the table.

STEPS FOR EXTRACTING DATA FROM NIS

Step 1

In step 1, you identify the ICD9 codes for Robotic Hysterectomy (1741, 1742, 1744, 1749), Laparoscopic procedures (6851, 5421, 6501, 6531, 6541, 6553, 6563, 6564). Any procedure without the laparoscopic or robotic codes is assumed to be open Hysterectomy using an ARRAY statement:

```
DATA indata(WHERE=(PROC_68384=1));
 SET sasdata.nis_2009_core (in=in_09)
   sasdata.nis_2010_core (in=in_10);
 ARRAY pr $ pr1- pr15;
DO OVER pr;
IF compress(pr) in ('6831','6841','6839','6849') THEN PROC_68384=1; *which
is icd9 68.3,68.4;
IF compress(pr) in
('6851','5421','6501','6531','6541','6553','6563','6564') THEN LAPR=1;
IF compress(pr) in ('1741') THEN OPenRobot=1;
IF compress(pr) in ('1742', '1744') THEN LaprRobot=1;
IF compress(pr) in ('1749') THEN UnSpecRobot=1;
END;
IF OPenRobot or LaprRobot or UnSpecRobot THEN RACS=1;
```

Step 2

In step 2, you use ARRAY statements to create different diagnosis variables on the different ICD9 codes for acute myoinftracton, bladder injury, deep vein thrombosis, and so on, for the procedure Hysterectomy:

```
 ARRAY dx $ dx1- dx25;
Do OVER dx;
IF compress(dx) in ('6262','6266','6268','6270','218x','219x') THEN
Hysterectomy=1; *menorrhagia and leiomyomas;
IF compress(dx) in ('V6441') THEN PROCV=1;* Laparoscopic surgical procedure
converted to open procedure;
IF compress(dx) in ('4534','V1251') THEN Deepvein_thrombosis=1;
IF compress(dx) in ('4151','V1251') THEN pulmonary_embolism=1;
IF compress(dx) in ('410') THEN Acute_Myo_Infraction=1;
IF compress(dx) in ('5698') THEN bowel_infraction=1; * Other specific
disorders of intenstine is not included;
END;
```

Step 3

In step 3, you use an IF statement to create an overall flag for all of the diagnosis variables:

```
IF COMP_BLEED=1 or COMP_INFECT=1 or  COMP_GI=1 or  COMP_MWOUND=1
   or COMP_URINE=1 or COMP_PULMON=1 or COMP_CARDIO=1 or  COMP_SYSTEM=1
   or COMP_PROCE=1   THEN OVERALL_COMP=1;
```

Step 4

In step 4, you create different cohorts based on the condition of whether the procedure is converted or not:

```
DATA add_cohort(WHERE=(flg_delete ne 1));
 SET indata;
  IF RACS=1 THEN COHORT=3;
  ELSE IF LAPR=1 THEN COHORT=2;
  ELSE                COHORT=1;
   IF PROCV=1 and COHORT=1 THEN flg_delete=1;
 RUN;
```

Step 5

In step 5, Using the SQL procedure ,The NIS severity files which contains the information of disease severity measures is combined with the inpatient Core file to identify different levels of obesity, risk mortality, and loss of function:

```
DATA severity;
 SET sasdata.nis_2009_severity
     sasdata.nis_2010_severity;
     ;
RUN;

PROC SQL; CREATE TABLE with_sev AS SELECT a.*, b.CM_OBESE,
b.APRDRG_Risk_Mortality, b.APRDRG_Severity
   ,case WHEN CM_OBESE=1 THEN 2
    ELSE OBESITY
    END AS ALL_OBESITY
   FROM add_cohort AS a LEFT JOIN severity AS b ON a.key=b.key;
quit;
```

Step 6

In step 6, you used the format statement to format all of the diagnosis variables, which are then used in further automated macros to create the report:

```
DATA addd;
 SET with_sev;
  FORMAT COHORT cohort. female sex. APRDRG_Risk_Mortality risk.
         ALL_OBESITY obese.  APRDRG_Severity sev.  MULTI_ICD9 multpro. ;
 RUN;
```

Step 7

In step 7, you create the SAS compiled macros that are stored in the Work library, which contains the 4 macros that are called by tables and are used to call the macros for different macro variables.

Step 8

In step 8, you call the created macros in order to create the frequencies of all of the cohorts and to create the table structure with the title and number of the table:

```
%MACRO run_tbl (the_if=, tbl_num=, add_title=);

DATA indata;
 SET sasdata.addd
     sasdata.addd(in=for_tot);
     IF for_tot THEN do;
          COHORT=4;
     END;
     &the_if;;
RUN;

%CntCat_COL(indata=indata,uniqueid=key,grpvar=COHORT,nfmt=8,invar=countit,
grpdata=, outdata=countit, OUTNFMT=n);
%CntCat_COL(indata=indata, uniqueid=key,
grpvar=COHORT,nfmt=8,invar=OVERALL_COMP, grpdata=, outdata=OVERALL_COMP);
%StackAll_COL(inlist=countit OVERALL_COMP COMP_BLEED COMP_INFECT  COMP_GI
COMP_MWOUND  OVERALL_COMP9 COMP_BLEED9 COMP_INFECT9  COMP_GI9
COMP_MWOUND9 COMP_URINE9 COMP_PULMON9  COMP_CARDIO9 COMP_SYSTEM9
COMP_PROCE9);

DATA _final;
 SET final;
       IF OUTLABEL=' ' THEN DELETE;

       IF cat not in ( 1,2, 12,13, 23,24) THEN
OUTLABEL="   "||TRIM(OUTLABEL);
       IF cat=12 THEN OUTLABEL=TRIM(OUTLABEL)||",2009";
       IF cat=24 THEN OUTLABEL=TRIM(OUTLABEL)||",2010";

       IF cat=1 THEN OUTLABEL="Procedures in 2009-2010, N";
         IF cat=12 THEN OUTLABEL="Procedures in 2009, N";
           IF cat=23 THEN OUTLABEL="Procedures in 2010, N";


       IF cat<12 THEN pgbk=1;
       ELSE IF cat<23 THEN pgbk=2;
       ELSE pgbk=3;
RUN;
```

```
PROC SORT DATA=_final; BY pgbk cat; RUN;


%N_grp(nvar=COHORT,cvar=,data=indata, unique=key);

TITLE; FOOTNOTE;
OPTIONS NODATE NONUMBER ORIENTATION=portrait;
ODS NORESULTS;
ODS LISTING CLOSE;
ODS ESCAPECHAR = '!';
 %let OUTDIR=C:\N.I.S\Hysterectomy\Tables;
 %let PAGEHDR=9thJune2014;
 %let PAGEFTR=9thJune2014;
 %let PROJECT=hysterectomy;
 %let PGMDIR=&outdir;
 %let PROGRAM=hysterectomy;
 %let RUNDTM=xxx;
ODS RTF FILE="&outdir.\&tbl_num..rtf" STYLE=styles.PORT8FT;


   TITLE1 j=l "&PAGEHDR:&PROJECT" j=r "{Page}\~{\field{\*\fldinst {PAGE
}}}\~{of}\~{\field{\*\fldinst {NUMPAGES }}}{}";
   TITLE2 "&tbl_num.";
   FOOTNOTE j=l "&Line118";
   FOOTNOTE2 j=l "&PAGEFTR &PGMDIR\&PROGRAM..sas &rundtm";*last footer
text;


PROC REPORT DATA=_final MISSING SPLIT='|' CENTER SPACING=3 HEADLINE
HEADSKIP NOWD STYLE(COLUMN)=[just=center];
   column  pgbk cat  OUTLABEL by1grp1 by1grp2 by1grp3 by1grp4;
   DEFINE pgbk/ORDER ORDER=internal noprint;
   DEFINE cat/ORDER  ORDER=internal noprint;
   DEFINE outlabel/DISPLAY     " " FLOW style(column)=[cellwidth=3.0 in
just=left] style(header)=[just=left];
   DEFINE by1grp1/display    "&Group1.|" style(column)=[cellwidth=.9 in];
   DEFINE by1grp2/display    "&Group2.|" style(column)=[cellwidth=.9  in];
   DEFINE by1grp3/display    "&Group3.|" style(column)=[cellwidth=.9 in];
   DEFINE by1grp4/display    "&Group4.|" style(column)=[cellwidth=.9  in];

   COMPUTE BEFORE pgbk;line @1 ' ';ENDCOMP;

RUN;

ODS RTF CLOSE;
ODS LISTING;
QUIT;
%MEND run_tbl;

%run_tbl (the_if=, tbl_num=Table 2, add_title=);
%run_tbl (the_if=if ALL_OBESITY ne ., tbl_num=Table 2.2, add_title=Obesity
Discharges for);
```

RESULTS

	Open	Laparoscopic	Robotic	Overall
Year				
2009	4184 (51.8)	417 (48.7)	91 (33.5)	4692 (50.9)
2010	3900 (48.2)	440 (51.3)	181 (66.5)	4521 (49.1)
Sex				
.	8084 (100.0)	857 (100.0)	272 (100.0)	9213 (100.0)
Male	3838 (47.5)	390 (45.5)	155 (57.0)	4383 (47.6)
Female	4229 (52.3)	466 (54.4)	117 (43.0)	4812 (52.2)
Age (years)				
n	8080	856	272	9208
Mean	62.2	61.5	59.2	62.1
Median	62.0	61.0	59.0	62.0
Std	13.39	13.15	12.61	13.35
Min, Max	12, 97	18, 95	22, 88	12, 97
Obesity				
.	8084 (100.0)	857 (100.0)	272 (100.0)	9213 (100.0)
Over Weight: 278.02	13 (0.2)	2 (0.2)	1 (0.4)	16 (0.2)
Obesity: 278.00, 278.0, 278.03, CM:OBESE	724 (9.0)	72 (8.4)	23 (8.5)	819 (8.9)
Morbid Obesity: 307.50, 278.01	1 (0.0)	0 (0.0)	0 (0.0)	1 (0.0)
APRDRG: Risk Mortality				
.	8084 (100.0)	857 (100.0)	272 (100.0)	9213 (100.0)
Minor likelihood of dying	4264 (52.7)	538 (62.8)	185 (68.0)	4987 (54.1)
Moderate likelihood of dying	2496 (30.9)	209 (24.4)	65 (23.9)	2770 (30.1)
Major likelihood of dying	904 (11.2)	88 (10.3)	14 (5.1)	1006 (10.9)
Extreme likelihood of dying	420 (5.2)	21 (2.5)	8 (2.9)	449 (4.9)
APRDRG: Severity				
.	8084 (100.0)	857 (100.0)	272 (100.0)	9213 (100.0)
Minor loss of function	2670 (33.0)	360 (42.0)	127 (46.7)	3157 (34.3)
Moderate loss of function	3311 (41.0)	334 (39.0)	108 (39.7)	3753 (40.7)
Major loss of function	1536 (19.0)	132 (15.4)	27 (9.9)	1695 (18.4)
Extreme loss of function	567 (7.0)	30 (3.5)	10 (3.7)	607 (6.6)

Table 1 shows the resulting output containing key outcome variables from the NIS dataset stratified by method of surgery (Open, Laparoscopic, and Robotic).

	Open	Laparoscopic	Robotic	Overall
Year				
2009	4184 (51.8)	417 (48.7)	91 (33.5)	4692 (50.9)

	Open	Laparoscopic	Robotic	Overall
2010	3900 (48.2)	440 (51.3)	181 (66.5)	4521 (49.1)
Sex				
.	8084 (100.0)	857 (100.0)	272 (100.0)	9213 (100.0)
Male	3838 (47.5)	390 (45.5)	155 (57.0)	4383 (47.6)
Female	4229 (52.3)	466 (54.4)	117 (43.0)	4812 (52.2)
Age (years)				
n	8080	856	272	9208
Mean	62.2	61.5	59.2	62.1
Median	62.0	61.0	59.0	62.0
Std	13.39	13.15	12.61	13.35
Min, Max	12, 97	18, 95	22, 88	12, 97
Obesity				
.	8084 (100.0)	857 (100.0)	272 (100.0)	9213 (100.0)
Over Weight: 278.02	13 (0.2)	2 (0.2)	1 (0.4)	16 (0.2)
Obesity: 278.00, 278.0, 278.03, CM:OBESE	724 (9.0)	72 (8.4)	23 (8.5)	819 (8.9)
Morbid Obesity: 307.50, 278.01	1 (0.0)	0 (0.0)	0 (0.0)	1 (0.0)
APRDRG: Risk Mortality				
.	8084 (100.0)	857 (100.0)	272 (100.0)	9213 (100.0)
Minor likelihood of dying	4264 (52.7)	538 (62.8)	185 (68.0)	4987 (54.1)
Moderate likelihood of dying	2496 (30.9)	209 (24.4)	65 (23.9)	2770 (30.1)
Major likelihood of dying	904 (11.2)	88 (10.3)	14 (5.1)	1006 (10.9)
Extreme likelihood of dying	420 (5.2)	21 (2.5)	8 (2.9)	449 (4.9)
APRDRG: Severity				
.	8084 (100.0)	857 (100.0)	272 (100.0)	9213 (100.0)
Minor loss of function	2670 (33.0)	360 (42.0)	127 (46.7)	3157 (34.3)
Moderate loss of function	3311 (41.0)	334 (39.0)	108 (39.7)	3753 (40.7)
Major loss of function	1536 (19.0)	132 (15.4)	27 (9.9)	1695 (18.4)
Extreme loss of function	567 (7.0)	30 (3.5)	10 (3.7)	607 (6.6)

Table 1. Key Outcome Variables from NIH Dataset Stratified by Method of Surgery

CONCLUSION

This research provides a simple way to summarize data from the NIS Database using different ICD9 codes specific to individual surgical procedures. It allows for the comparison of multiple surgical approaches in order to make statistical inferences.

REFERENCES

[1]. National Inpatient Sample. Available at: https://www.hcup-us.ahrq.gov/nisoverview.jsp

[2]. Burlew, Michele M. (1998), SAS Macro Programming Made Easy, SAS Institute Inc., Cary, NC, USA.

ACKNOWLEDGMENTS

The author would like to thanks April Hebert and James Lavin for their valuable suggestions and contribution to this project.

CONTACT INFORMATION

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

Ravi Gaddameedi
California State University
rgaddameedi@horizon.csueastbay.edu