Paper 3281-2015

Using SAS® to Create Episodes-of-Hospitalization for Health Services Research

Meric Osman, Jacqueline Quail, Nedeene Hudema, Nianping Hu

Saskatchewan Health Quality Council, Saskatoon, SK, CANADA

ABSTRACT

An essential part of health services research is describing the use and sequencing of a variety of health services. One of the most frequently examined health services is hospitalization. A common problem in describing hospitalizations is that a patient might have multiple hospitalizations to treat the same health issue. Specifically, a hospitalized patient might be (1) sent to and returned from another facility in a single day for testing or a procedure, (2) transferred from one hospital to another, and/or (3) discharged home and re-admitted within 24 hours. In all cases, these hospitalizations are likely treating the same underlying health issue and should be considered as a single hospitalization episode. If examined without regard for the episode, a patient with a series of hospitalizations (an initial hospitalization, a nested hospitalization, a transfer hospitalization, and a readmission hospitalization) would be identified as having 4 hospitalizations when in reality, they had one hospitalization episode spanning multiple facilities. Failing to account for multiple hospitalizations in the same episode has implications for many disciplines including health services research, resource planning, and quality improvement for patient safety. We developed SAS® code to identify episodes of hospitalizations, a count of hospitalizations within each episode, and the overall duration of the episode. The output clearly displays the data in an intuitive and easy-to-understand format. The method we describe and the associated SAS code will be useful to not only health services researchers, but also anyone who works with temporal data that includes nested, overlapping, and subsequent events.

INTRODUCTION

Health services research (HSR) investigates the use, costs, and outcomes of health care services with the goal of identifying the best way to deliver the highest quality of care for the lowest cost. Inherent in this definition is avoiding medical harm and improving patient safety (Agency for Healthcare Research and Quality, 2002).

An essential part of HSR is describing the use and sequencing of a variety of health services including physician visits, hospitalizations, emergency department visits, prescription drug dispensations, and services provided through home care and in long-term care, among others. By far the most costly of these services is hospitalizations.

An *episode-of-care* is a collection of temporally contiguous healthcare services that an individual receives to treat or manage a specific health condition or event (Hornbrook et al, 1985). An episode-of-care can include many different health services but for the purpose of this paper, we focus only on hospital services. A common problem in ascertaining hospitalizations is that a patient can have multiple different hospitalizations that are part of the same health care issue. These multiple hospitalizations can be considered as belonging to a single *episode-of-hospitalization*. Specifically, a person may have a nested, transfer, and/or a readmission hospitalization.

Failing to account for these hospitalizations can lead to fallacious conclusions in numerous areas of study. In health services research, hospitalizations will be counted multiple times leading to an overestimate of the number of hospitalizations a person had. For example, a person can be identified as having 4 hospitalizations when in reality they had one episode-of-hospitalization. This will result in a person appearing to have a higher incidence or rate of hospitalization than is true.

Regarding resource planning for health services, the average time and resources needed to treat a specific health problem may be underestimated. To illustrate, if a patient spends 10 days each in 3

different hospitals in the same episode, the total number of days needed to treat the health problem is 30 days, but each hospital identify only 10, and planned resourcing may be inadequate.

Patient safety is an essential component of health service research. Hospital-acquired infections are a serious concern and a major cause of extended hospital stays, morbidity, and death. As a result, many hospitals have quality improvement programs that monitor the occurrence of infections in order to identify ways to reduce them. If episodes-of-hospitalization are not considered, an infection acquired in a hospital that does not manifest until a patient is transferred to a different hospital will incorrectly be attributed to the receiving hospital. Furthermore, failing to account for the continuity of hospital services may lead to double counting and overestimation of hospital acquired infections.

In sum, a single patient may have multiple related hospitalizations. Failing to account for the interrelationship of these hospitalizations by treating them as independent events may lead to fallacious conclusions. We created a macro to identify episodes-of-hospitalization and the components of each episode that will be useful to health service researchers and other professionals.

TYPES OF HOSPITALIZATIONS

There are three main types of hospitalizations that may comprise an episode-of-hospitalization: nested, transfer, and readmission hospitalizations. See Figure 1 and Table 1 for a graphical and tabular example of each of the types of hospitalizations.

A **nested** hospitalization occurs when a patient is temporarily sent to another facility, often for testing or a procedure that cannot be conducted at the current facility, and the admission and discharge dates occur **within** the bounds of the admission and discharge dates of the initial hospitalization.

A **transfer** (or sequential) hospitalization occurs when a patient is transferred directly from one hospital to another and there is no intent to return to the first hospital. The admission date of a transfer hospitalization occurs on the discharge date of the patient from the first hospital.

A **readmission** (or subsequent) hospitalization occurs when a patient is discharged from the hospital but returns to any hospital within 24 hours. The short time between discharge and readmission indicates that the patient was discharged prematurely and that the underlying reason for their hospitalization was not fully resolved or they are experiencing a complication of the health issue that was responsible for their most recent hospitalization. Thus, the readmission hospitalization can be considered as an extension of the most recent hospitalization. The admission date of the readmission hospitalization will be either on the discharge date of the most recent hospitalization or the day after.

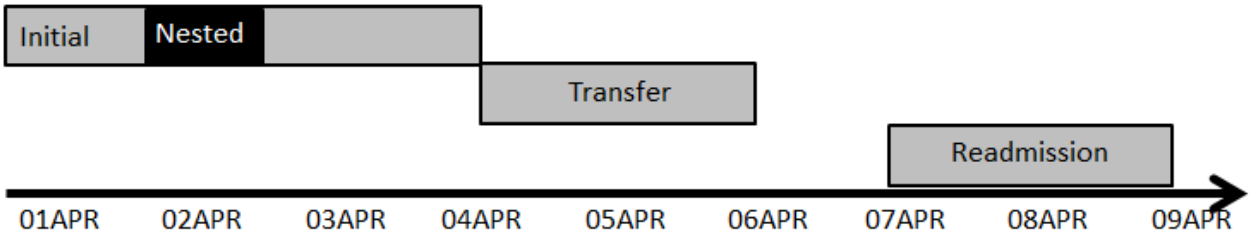

Figure 1: Graphical representation of an episode-of-hospitalization.

Patient ID	Admit date	Discharge date	HOSP TYPE
001	01APR2014	04APR2014	Initial
001	02APR2014	02APR2014	Nested
001	04APR2014	06APR2014	Transfer
001	07APR2014	09APR2014	Readmission

Table 1. Database representation of the episode-of-hospitalization shown in Figure 1.

It has to be noted that it is not possible to differentiate between transfer and readmission hospitalizations without additional information on discharge destination and from where a person is admitted. For simplicity, the macro we present in this paper does not use 'admission from' and 'discharge destination' data and so all transfer and readmission hospitalizations are identified as transfer hospitalizations in the macro output.

SAS CODE TO IDENTIFY EPISODES OF HOSPITALIZATION

Our macro requires only three variables to be input: patient identification number, hospital admission date, and hospital discharge date.

Our SAS macro outputs 7 variables to thoroughly describe episodes of hospitalizations. See Table 2 for an example of each. Please note that the HOSP TYPE variable is not part of the output but is included for descriptive purposes only.

(1) numbering of each hospitalization fed to the macro (TOT_HOSP) – a count of each hospitalization without regard for episode-of-hospitalization;

(2) the episode-of-hospitalization (EOH) – indicates to which episode each hospitalization belongs and can include any number of nested and transfer hospitalizations;

(3) the sequence and count of hospitalizations (COUNT) - the number of hospitalizations in the temporal order in which they occurred regardless of their type within each episode;

(4) a flag to identify nested hospitalizations (NEST);

(5) a flag to identify transfer and readmission hospitalizations (TRANS);

(6) the length of stay (in days) of each hospitalization (HOSP_LOS); and

(7) the length of stay (in days) for the entire episode (EOH_LOS).

Patient ID	Admit date	Discharge date	HOSP TYPE	TOT_ HOSP	EOH	COUNT	NEST	TRANS	HOSP_ LOS	EOH_ LOS
001	01APR2014	04APR2014	Initial	1	1	1	0	0	4	9
001	02APR2014	02APR2014	Nested	2	1	2	1	0	1	9
001	04APR2014	06APR2014	Transfer	3	1	3	0	1	3	9
001	07APR2014	09APR2014	Transfer	4	1	4	0	1	3	9
001	01SEP2014	05SEP2014	Initial	5	2	1	0	0	5	5
001	01DEC2014	03DEC2014	Initial	6	3	1	0	0	3	6
001	03DEC2014	03DEC2014	Nested	7	3	2	1	0	1	6
001	03DEC2014	06DEC2014	Transfer	8	3	3	0	1	4	6

Table 2. Example of the admission and discharge dates of individual, nested, and transfer hospitalizations and associated macro output detailing episodes-of-hospitalization

Notice that in each episode-of-hospitalization the sum of HOSP_LOS (calculated as discharge date – admission date + 1) does not always equal the EOH_LOS. This is because nested hospitalizations, same-day transfers, and same-day readmissions occur on the same day of an existing hospitalization. To include them in EOH_LOS would count the same day twice. Consequently these days are subtracted from the EOH_LOS.

OTHER METHODOLOGICAL CONSIDERATIONS

As seen so far, our macro only utilizes hospital admission and discharge dates to create patient level episodes and descriptive variables. However, researchers might have access to other information that can impact the episode and/or hospital type assignments. Level of hospital care (i.e., day surgery vs.

inpatient stay), admission reason (i.e., diagnosis), 'admission from' and 'discharge to' categories, and the time between hospitalizations are examples of such information.

Level of care has a potential to impact if a hospitalization is identified as a nested or transfer hospitalization. If the second hospitalization in an episode is an inpatient stay, then it can be considered a transfer. Alternatively, if it is for a day surgery, it is most likely a nested hospitalization and the person will return to the initial facility. Interested researchers can connect with the main author to request the full version of SAS code that incorporates the level of care into episode creation.

Admission reason (i.e., most responsible diagnosis) can impact the number of hospitalizations within each episode. Even if a person has two or more close hospitalizations, the hospitalizations may be for different reasons and thus belong to different episodes.

Information on where a patient was 'admitted from' or 'discharged to' can be used to differentiate between transfer and readmission hospitalizations. Finally, for different jurisdictions, allowing for only a single day between hospitalizations for episodes may not be adequate and other time frames might be a better choice for episode assignment.

For simplicity, our macro does not take the above considerations into account. However, it can serve as foundational code to be adapted and expanded by other researchers depending upon the data sources available to them.

LIMITATIONS

A limitation of our macro is that it is not able to distinguish between nested and single-day transfer hospitalizations when they occur on the same day as the start or end of any hospitalization within the episode-of-hospitalization. The following tables demonstrate two possible courses of an episode-of-hospitalization with the same dates but different types of hospitalizations.

Patient ID	Admit date	Discharge date	HOSP TYPE
003	31OCT2012	11NOV2012	Initial
003	31OCT2012	31OCT2012	Nested
003	11NOV2012	25DEC2012	Transfer
003	25DEC2012	25DEC2012	Nested

Table 3a. Example of an episode-of hospitalization with nested hospitalizations

Patient ID	Admit date	Discharge date	HOSP TYPE
003	31OCT2012	31OCT2012	Initial
003	31OCT2012	11NOV2012	Nested
003	11NOV2012	25DEC2012	Transfer
003	25DEC2012	25DEC2012	Transfer

Table 3b. Example of an episode-of-hospitalization with one-day transfer hospitalizations

For each patient, our macro sorts rows first by the date of admission and then the date of discharge, reordering the components of the episode-of-hospitalization. The following table displays the results of the course of an episode-of-hospitalization when processed by our macro.

Patient ID	Admit date	Discharge date	HOSP TYPE	TOT_ HOSP	EOH	COUNT	NEST	TRANS	HOSP_ LOS	EOH_ LOS
003	31OCT2012	31OCT2012	Nested	1	1	1	1	0	1	56
003	31OCT2012	11NOV2012	Transfer	2	1	2	0	1	12	56
003	11NOV2012	25DEC2012	Transfer	3	1	3	0	1	45	56
003	25DEC2012	25DEC2012	Transfer	4	1	4	0	1	1	56

Table 3c: Episode-of-hospitalization results

This limitation does not affect the correct identification of the start and end dates for the episode-of-hospitalization. It is possible to differentiate between initial versus nested versus transfer hospitalizations using an identification number that is unique to health care facilities, but to do so is beyond the scope of this paper.

CONCLUSION

We created a macro to identify hospitalizations that are nested, overlapping, and subsequent in a single episode-of-hospitalization. The results of the episode-of-hospitalization are presented in a format that is easy to understand and that also allows researchers to separate or aggregate hospitalizations as desired.

Reporting in an episode-of-hospitalization format, rather than accounting for each hospitalization singly, will prevent misinterpretation of results and provide more accurate information that can be used to plan and improve health care services. This macro will be useful to all health services researchers, resource planning officials, and hospital quality improvement specialists.

REFERENCES

Agency for Healthcare Research and Quality, 2002. “What is Health Services Research?” Accessed December 30, 2014. http://www.academyhealth.org/Training/content.cfm?ItemNumber=1003&navItemNumber=2384.

Hornbrook MC, Hurtado AV, Johnson RE. 1985. Health care episodes: definition, measurement and use. *Medical Care Review*, 42 (2):163-218.

CONTACT INFORMATION

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

Meric Osman
Saskatchewan Health Quality Council
306-668-8810 ext. 146
mosman@hqc.sk.ca
www.hqc.sk.ca

APPENDIX

SAS® macro to calculate episodes of hospitalizations.

```
%macro hospital_eoh(dataset_in=,dataset_out=);

/* Step 1 - Number the sorted data for EoH */

data data_1;
       set &dataset_in;
       by patient_id adm_date dschg_date;
       retain n;
       if first.patient_id then n=0;
       n=n+1;
run;


/* Step 2 - Transpose the variables in adm_date and dschg_date for EoH */

%macro eoh(inA=,inB=,outA=,from=,to=);
proc transpose data=&inA out=&outA (drop=_name_) prefix=&to;
       by patient_id;
       var &from;
run;
%mend;
%eoh(inA=data_1,outA=admit,from=adm_date,to=adm_date)
%eoh(inA=data_1,outA=dschg,from=dschg_date,to=dschg_date)


/* Step 3 - Combine the transposed admission & discharged date & main files */

%macro join(inA=, inB=, outA=, b=, xvar=);
data &outA;
       merge &inA (in=x) &inB (in=y drop=&xvar);
       by &b;
       if x and y; run;
%mend;
%join(inA=admit, inB=dschg, outA=Admit_dschg, b=patient_id)
%join(inA=data_1, inB=Admit_dschg, outA=hospital_adm_dschg, b=patient_id)


proc sql;
    select left(trim(put(MAX(count),5.))) into: Idy from
    (select count(*) as count
     from hospital_adm_dschg
     group by patient_id);
QUIT;
%PUT &Idy;

proc sort data=hospital_adm_dschg;
       by patient_id adm_date dschg_date;
run;

/* Step 4 - Compute and assign EoH & Hospitalization Types */

data data_2 (drop= i diff done eoh2 seq2 adm_date1-adm_date&Idy dschg_date1-
dschg_date&Idy);
       set hospital_adm_dschg;
       by patient_id adm_date dschg_date;
    where n LE &Idy;
```

```
	los = dschg_date - adm_date + 1;
	array adm (*) adm_date1-adm_date&Idy;
	array dcg (*) dschg_date1-dschg_date&Idy;
	retain eoh2 seq2; done=0;
		if first.patient_id then do;
			eoh1=0; eoh2=eoh1; seq=0; seq2=seq; nest=0; trans=0;
		end;

	if n = 1 then do;
		eoh1=1; eoh2=eoh1; seq=1; seq2=seq; nest=0; trans=0;

	 if adm(2) ^= . AND adm(1) >= adm(2) and dcg(1) < dcg(2) then do;
			eoh1=eoh2; seq=seq2; eoh2=eoh1; seq2=seq; nest=1; trans=0;
		end;
	end;
	else if n = 2 then
	do;
		if 0 <= abs(adm(2)-dcg(1)) <= 1 and (dcg(2) > dcg(1) or
			(dcg(2) = dcg(1))) then do;
				eoh1=eoh2; seq=seq2+1; eoh2=eoh1; seq2=seq; nest=0;
trans=1;
		end;
		else if adm(2) >= adm(1) and (dcg(2) < dcg(1) or
				(dcg(2) = dcg(1))) then do;
			eoh1=eoh2; seq=seq2; eoh1=eoh2; seq=seq2; nest=1; trans=0;
		end;
		else if adm(3) ^= . AND adm(2) >= adm(3) and dcg(2) < dcg(3) then do;
			eoh1=eoh2; seq=seq2; eoh2=eoh1; seq2=seq; nest=1; trans=0;
		end;
		else do;
			eoh1=eoh2+1; seq=seq2; eoh2=eoh1; seq2=seq; nest=0; trans=0;
		end;
	end;
	else if n > 2 then do;
		do i=1 to n-1;
			diff=abs(adm(n)-dcg(i));
			if 0 <= abs(adm(n)-dcg(i)) <= 1 and
				((dcg(n) > dcg(i) or dcg(n) = dcg(i))) then do;
				eoh1=eoh2; seq=seq2+1; eoh2=eoh1; seq2=seq; nest=0;
trans=1; done=1;
				leave;
			end;
		end;
		do i=1 to n-1;
			if not(done) and adm(n) >= adm(i) and (dcg(n) < dcg(i) or
				(dcg(n)=dcg(i))) then do;
				eoh1=eoh2; seq=seq2; eoh1=eoh2; seq=seq2; nest=1; trans=0;
done=1;
				leave;
			end;
		end;
		do i=1 to n-1;
			if not(done) and abs(adm(n) - dcg(i)) > 1 then do;
				done=1; seq=1; seq2=seq; eoh1=eoh2+1; eoh2=eoh1; nest=0;
trans=0;
				leave;
			end;
		end;
		if (n + 1) LE &Idy AND adm(n+1) ^= . and adm(n) >= adm(n+1) and dcg(n) <
dcg(n+1) then do;
			eoh1=eoh2; seq=seq2; eoh2=eoh1; seq2=seq; nest=1; trans=0; done=1;
			END;
	end;
```

```
run;

proc sort data=data_2;
       by patient_id eoh1 adm_date dschg_date;
run;

/* Step 5 - Compute other variables (i.e. LOS) & prepare the final output */

DATA data_3; retain patient_id adm_date dschg_date n eoh1 count nest trans los
los_eoh; SET data_2;
  by patient_id eoh1 adm_date dschg_date;
  RETAIN count; count + 1;
  IF FIRST.eoh1 THEN count = 1;
rename eoh1 = eoh;
rename n      = total_hosp;
rename los    = hosp_los;
drop seq;
run;

data data_nonest;
set data_3;
  IF nest ^= 1;
run;


DATA data_4;
  SET data_nonest;
  BY patient_id EOH;
  IF FIRST.EOH OR LAST.EOH;
  dschg_date_first = lag(dschg_date);
  adm_date_first   = lag(adm_date);
  patient_id_first    = lag(patient_id);
  EOH_first        = lag(EOH);
  IF patient_id = patient_id_first AND EOH = EOH_first THEN adm_date = adm_date_first;
  IF LAST.eoh;
  eoh_los = dschg_date - adm_date + 1;
  format adm_date dschg_date date9.;
  KEEP patient_id EOH eoh_los count;
RUN;


PROC SQL;
  CREATE TABLE &dataset_out AS
  SELECT a.*, b.eoh_los
  FROM data_3 a LEFT JOIN data_4 b
  ON a.patient_id = b.patient_id AND a.eoh = b.eoh
  ORDER BY patient_id, EOH, adm_date, dschg_date;
QUIT;


%mend hospital_eoh;
```