Paper 01: Older adult visits to the emergency department for ambulatory care sensitive conditions.

Summary

This paper is a summary paper that looks at ACSC related hospitalizations and their cost. This paper was selected as the first paper because it performed no statistics, there is no model, this paper is just summary statistics.

Replication notes

So, right away I’m skeptical of the results of this paper. I know it isn’t a huge thing but the numbers in the tables don’t add up. As in, literally, tables with the same cohort don’t have the same number of members.

Data Source

2016 SAF ED based Claims

ED claims will be inpatient and outpatient claims with a revenue center code equal to 0981 or starting with 045. This is shown in Dev01.

This query resulted in 24,952,921 distinct claims.

I validated the results to confirm that all revenue centers were present and there were no duplicated claims. (dev02)

They excluded under age 65

Next they reference claims which include beneficiaries enrolled in Part A and Part B for the full period.

This implies HI\_COVERAGE and SMI\_COVERAGE equal to 12 depending

They don’t specify if they exclude HMO or not, so I will test that. (dev03)

14,680,040 with the age range, SMI and HI, if I remove HMO 14,108,555

So… I wonder if they’re also including every other claim as well… they do not actually state what datasets that they use just that they’re using the 2016 SAF.

Let me try increasing the datasets to reference every dataset. 14,108,899

So, basically, no difference.

So I will be using just inpatient and outpatient (as per dev 03)

For table 1, The variables needed are

Age (age From MBSF)

Race (race\_code from MBSF)

Sex (sex\_code from MBSF)

Region (state\_code from MBSF)

ED to Admission (ADM\_DATE from inp\_claims)

30-day ED revisit (Looking for the next ED visit by re-joining on the same table on desy\_sort\_key and minimum date > current date on the ED table)

Dual Eligibility status (From MBSF)

CCI (From PQI2016 table that uses all DX)

ACSC is from the Primary Diagnosis column.

For Table 2

I also need

CLM\_PMT\_AMT, NCH\_PRMRY\_PYR\_CLM\_PD\_AMT, and CLM\_TOT\_CHRG\_AMT from the claim file.

And for Table 3 I need to know what the revisit diagnosis was. (So when I join to get the number of days to the next visit, I need to also know what the primary diagnosis for that visit was, AND the PQI for that condition)

The first version of this full SQL query is in dev 05.

This dataset will be pulled into python and parsed in dev 06.

The Data Structure

Dataset:

DESY\_SORT\_KEY Unique Patient Identifier (PK)

CLAIM\_NO Unique Claim Identifier (PK)

AGE Age of the Individual

RACE\_CODE Race of the Individual

SEX\_CODE Sex of the Individual

STATE\_CODE Where the claim happened

DUAL\_STUS\_CD\_01-12 Medicare / Medicade Status

CLM\_THRU\_DT When the claim happened

PRNCPAL\_DGNS\_CD Primary Diagnosis

CLM\_PMT\_AMT What CMS Paid

NCH\_PRMRY\_PYR\_CLM\_PD\_AMT What a non CMS Payer Paid

CLM\_TOT\_CHRG\_AMT The Total Charge of the claim

Figure 1

ASCS / No ASCS Calcuclated from PRNCPAL\_DGNS\_CD

Total Sum of Rows

AGE Ranges calculated from AGE column

Race Ranges calculated from RACE\_CODE Column

Sex Ranges calculated from SEX\_CODE Column

Geographic Region Region Calculated from the STATE\_CODE

ED to admission See Discussion Below

30-day ED revisits See Discussion Below

Dual eligibility status See Discussion Below

Charlson Comorbidity Index Calcuclated Seperately

Figure 2

ACSC Filtered by PQI Flags

ED to admission inpAdmission

Average Pay by medicare CLM\_PMT\_AMT

Average Pay by supplement NCH\_PRMRY\_PYR\_CLM\_PD\_AMT

Average Pay by Bene CLM\_TOT\_CHRG\_AMT - NCH\_PRMRY\_PYR\_CLM\_PD\_AMT - CLM\_PMT\_AMT

Total Pay CLM\_TOT\_CHRG\_AMT

Figure 3

ACSC Filtered by PQI Flags

ED Visits Count Rows

30-day ED revisits same Using DESY\_SORT\_KEY, CLM\_THRU\_DT, and PRNCPAL\_DGNS\_CD

30-day ED revisits all Using DESY\_SORT\_KEY and CLM\_THRU\_DT

Figure 1 Expanded

Age ranges

65-74

75-84

85+

Race

1 = White

2 = Black

4 = Asian

5 = Hispanic

6 = North American Native

3 = Other

0 = Unknown

Sex

0 = Drop

2 = Female

1 = Male

Geographic region https://www.census.gov/programs-surveys/economic-census/guidance-geographies/levels.html

Northeast

07 Connecticut

20 Maine

22 Massachusetts

30 New Hampshire

41 Rhode Island

47 Vermont

31 New Jersey

33 New York

39 Pennsylvania

Midwest Region

14 Illinois

15 Indiana

23 Michigan

36 Ohio

52 Wisconsin

16 Iowa

17 Kansas

24 Minnesota

26 Missouri

28 Nebraska

35 North Dakota

43 South Dakota

South

08 Delaware

09 District of Columbia

10 Florida

11 Georgia

21 Maryland

34 North Carolina

42 South Carolina

49 Virginia

51 West Virginia

01 Alabama

18 Kentucky

25 Mississippi

44 Tennessee

04 Arkansas

19 Louisiana

37 Oklahoma

45 Texas

West

03 Arizona

06 Colorado

13 Idaho

27 Montana

29 Nevada

32 New Mexico

46 Utah

53 Wyoming

02 Alaska

05 California

12 Hawaii

38 Oregon

50 Washington

Else? Maybe Some of these are dropped?

40 Puerto Rico

48 Virgin Islands

54 Africa

55 Asia

56 Canada

57 Central America and West Indies

58 Europe

59 Mexico

60 Oceania

61 Philippines

62 South America

63 U.S. Possessions

97 Saipan - MP

98 Guam

99 American Samoa

Talk more about inp admission.

Initially, when I thought about how to determine how I wanted to determine inpatient admission I made it much harder than it needed to be. I looked for visits where there was a revenue charge to a bed made. Or I tried to look at the facility type of where the visit was, or how someone was discharged from the ED. The Solution I ended up using was that if someone had an admission date… they were admitted. And if they did not have an admission date, they were not admitted.

Talk more about next admission

So the readmission variable has a few complexities, When you explain it it’s simple, It’s the next ED or Inpatient Admission for that person; however, there are a few complexities.

So, lets answer some questions

1: Are there cases where someone has multiple ED visits on the same day?

Dev 07

14,108,555 Rows 🡪14,110,826 Rows

Yes, there are ~2,000 cases where there are multiple matches I know that they resolved this by just having an N that grows. So I will also do that.

2: Are there cases where someone has multiple readmissions on the same day?

Dev08

Talk more about dual coverage 2016

What % of people have any dual coverage Dev09

(No column name) DUAL\_STUS

52,553,678 0

7,195,286 1

People 65+ with full A and B coverage

(No column name) DUAL\_STUS

37,417,750 0

3,531,311 1

And without HMO Coverage

(No column name) DUAL\_STUS

23,545,160 0

2,026,457 1

What % of people have full dual coverage Dev10

(No column name) DUAL\_STUS

54,562,638 0

5,186,326 1

65+ with A and B

(No column name) DUAL\_STUS

38,083,291 0

2,865,770 1

65+ with A and B and no HMO

(No column name) DUAL\_STUS

23,904,615 0

1,667,002 1

I guess the short of it is… there is a difference in these populations depending on how you classify things.

With the Queries Established, I’m going to try to math my results to the papers results now, as closely as I can

Table 1, total N

16,908,378

I am looking at 14,680,040 with the age range, SMI and HI,

if I remove HMO 14,108,555

I believe I should remove HMO because they use price data, so 14,108,555 is the number.

I don’t have any decisions with age, race, sex, region, or ED to admission or 30-day revisit.

For Dual Status they have 4,637,340:12,271,038 or 1:2.64

This is hugely over-represented… which I suppose is to be expected. When I look into this for my ED…

So I think the best way to answer these questions are in the statistical software. So I am building a version of this (Dev11) that just has multiple variables to investigate.