ELECTRONIC HEALTH RECORDS (EHR) DATA:

CASE STUDY

John D. Rice, PhD, and
Carter Sevick, MS

ACCORDS
ADULT AND CHILD CONSORTIUM FOR HEALTH OUTCOMES RESEARCH AND DELIVERY SCIENCE
UNIVERSITY OF COLORADO | CHILDREN'S HOSPITAL COLORADO

Example Data

Centers for MEDICARE & MEDICAID Services Data Entrepreneur's Synthetic Public Use File (DE-SynPUF)

- https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/SynPUFs/DE_Syn_PUF
- Synthetic data based on a 5% sample of MEDICARE participants, 2008-2010
- Files divided into 20 separate samples (one participant will only appear in one sample)
  - For this example we are using sample 20
  - Code to import the data and for the basic analysis will be provided

Data Structure

- **8 Comma Separated Value (CSV) files stored in ZIP folders**
  - Beneficiary Summary File
    - 3, one for each year
  - Carrier Claims, 2 files
  - Inpatient Claims, 1 file
  - Outpatient Claims, 1 file
  - Prescription Drugs, 1 file
- **Relational Model**
  - Some data providers will attempt to fit all data into one file, with each row representing a person, or hospital stay
  - In the relational model different data types (entities) will be stored in separate tables
  - The SynPUF data is semi – relational
    - In the Inpatient file there are 10 columns for diagnoses and 6 for procedures.
    - This could be further broken out into separate diagnoses (one row per stay-diagnosis) and procedure tables.

Sources of Date Value Definitions

- **Sometimes data will come with all the needed look up tables….sometimes it will not.**
- **Centers for Medicare and Medicaid Services (www.cms.gov)**
  - ICD 9/10 diagnoses and procedures
    - https://www.cms.gov/Medicare/Coding/ICD9ProviderDiagnosticCodes/codes
    - https://www.cms.gov/Medicare/Coding/ICD10
  - DRGs etc.
- **AHRQ/H-CUP**
  - Provides many excellent grouping algorithms for summarizing health information
  - https://www.hcup-us.ahrq.gov/tools_software.jsp
- **Common Procedural Terminology (CPT)**
  - Used for billing in many settings, such as clinic and outpatient visits.
  - Proprietary to the American Medical Association
    - https://www.ama-assn.org/amaone/cpt-current-procedural-terminology
- **Additionally, many associations exist for the study of different disease states and sometimes provide tools to identify cases in EHR data**
  - Birth defects: National Birth Defects Prevention Network (www.nbdpn.org)

Example Study

WARNING:
This is not a rigorous publish ready analysis.
The intent is to show some of the qualities that abound in EHR data.

- **Characteristics of discharges with spinal fusion and associations with 60 day re-admission**
- **Data sources:**
  - DE-SynPUF Inpatient Claims and Beneficiary Summary files
- **Spinal fusion definition:**
  - ICD-9 procedure codes:

8100	Spinal fusion, not otherwise specified
8101	Atlas-axis spinal fusion
8102	Other cervical fusion of the anterior column, anterior technique
8103	Other cervical fusion of the posterior column, posterior technique
8104	Dorsal and dorsolumbar fusion of the anterior column, anterior technique
8105	Dorsal and dorsolumbar fusion of the posterior column, posterior technique
8106	Lumbar and lumbosacral fusion of the anterior column, anterior technique
8107	Lumbar and lumbosacral fusion of the posterior column, posterior technique
8108	Lumbar and lumbosacral fusion of the anterior column, posterior technique
8109	Other spinal fusion

Case Definition

- **Subset inpatient claims to those with a procedure indicating spinal fusion, looking for the initial.**
    - Is the first one seen the actual first? Could be a subsequent fusion of different vertebra, or a miss-coded re-fusion.
    - The data begins with eligible patient in 2008, but what happened before that?
    - To allow for history and be able to look at subsequent care, we will take only initial surgeries that occurred in 2009.

```
data inPatWithFlags;
  set library.inpatClaims;

  array pd(*) ICD9_PRCDR_CD_1 - ICD9_PRCDR_CD_6;

  * procedure flags ;
  spinFus = 0;
  label spinFus = 'Spinal fusion';
  do i = 1 to dim(pd);
     if pd(i) in ('8100','8101','8102','8103','8104','8105','8106','8107','8108','8109')
        then spinFus = 1;
  end;

  drop i;

run; * 66,514 records ;

proc sql;

create table firstSurgRec as
  select *
  from inPatWithFlags
  where spinFus=1
  group by DESYNPUF_ID
  having min(CLM_FROM_DT) = CLM_FROM_DT; * 536 records ;

create table firstSurgRec09 as
  select *
  from firstSurgRec
  where year(CLM_FROM_DT) = 2009; * 207 records ;

quit;
```

What Diagnoses Go Along With Spinal Fusion?

- **Top 10:**

diag1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
4019 : HYPERTENSION NOS	92	6.18	92	6.18
53081: ESOPHAGEAL REFLUX (Begin 1993)	45	3.02	137	9.20
72402: SPINAL STENOSIS-LUMBAR	45	3.02	182	12.22
25000: DIABETES UNCOMPL TYPE II	40	2.69	222	14.91
2724 : HYPERLIPIDEMIA NEC/NOS	37	2.48	259	17.39
72252: LUMB/LUMBOSAC DISC DEGEN	35	2.35	294	19.74
7384 : ACQ SPONDYLOLISTHESIS	31	2.08	325	21.83
7213 : LUMBOSACRAL SPONDYLOSIS	30	2.01	355	23.84
7211 : CERV SPONDYL W MYELOPATH	25	1.68	380	25.52
2851 : AC POSTHEMORRHAG ANEMIA-	24	1.61	404	27.13
The first 10 levels are displayed.				

ICD 9 diagnosis labels from https://www.hcup-us.ahrq.gov/db/tools/I9_Formats.TXT

And a Little Further Down the List

- **Individuals will be in the data for all kinds of reasons.**
  Is it desired to evaluate outcomes among those that may have had the surgery due to accidents?

diag1	COUNT	PERCENT
E8859: FALL FROM SLIPPING NEC (Begin 2000)	2	.13432
E8809: FALL ON STAIR/STEP NEC	1	.06716
E8889: FALL NOS (Begin 2001)	1	.06716

ICD 9 diagnosis labels from https://www.hcup-us.ahrq.gov/db/tools/I9_Formats.TXT

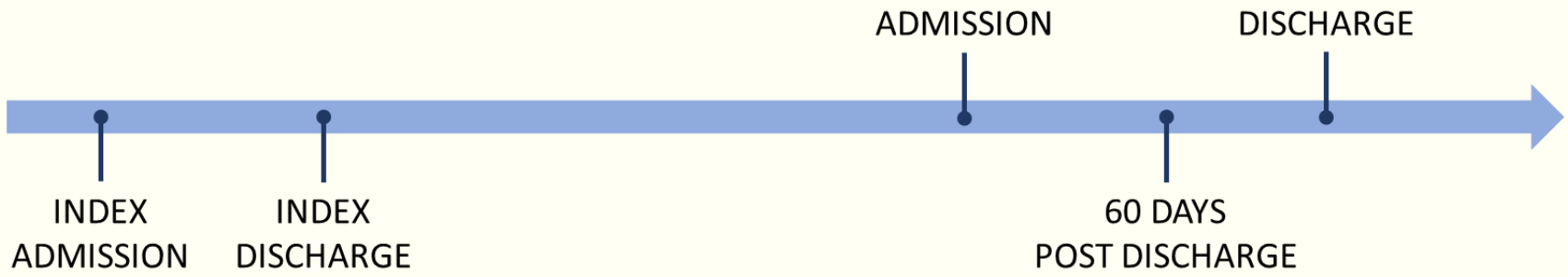
Based on a Review of Patient Diagnoses

- **The PI would like to add two new flags for evaluation in our patients:**
  - Index diagnoses of spinal curvature and falls
  - The following is added to the record classification algorithm:

```
* diagnosis flags ;
 spinCurve   = 0;
 acc_fall    = 0;
 label spinCurve = 'Spinal curvature'
       acc_fall  = 'Falls'
             ;
 do i = 1 to dim(dx);
   if substr(dx(i), 1, 3) = '737'                          then spinCurve = 1;
   if substr(dx(i), 1, 3) in ('E81', 'E82', 'E83','E88') then acc_fall  = 1;
 end;
```

60 Day Re-admission

- Of course there are many ways to achieve these results, but the way the analysis is being structured this will require the interaction of two different tables:
  - Index admissions and the original inpatient claims




```
create table firstReAdmit60 as
  select next.*
  from firstSurgRec09 as index inner join inPatWithFlags as next
    on index.DESYNPUF_ID = next.DESYNPUF_ID
  where next.CLM_FROM_DT between index.CLM_THRU_DT+1 and index.CLM_THRU_DT+60
  group by next.DESYNPUF_ID
  having min(next.CLM_FROM_DT) = next.CLM_FROM_DT;
```

Top 10 Diagnoses Associated with a Re-admission

diag1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
4019 : HYPERTENSION NOS	6	3.90	6	3.90
40390: HYP REN NOS W/O REN FAIL (Begin 1989)	4	2.60	10	6.49
2720 : PURE HYPERCHOLESTEROLEM	3	1.95	13	8.44
2724 : HYPERLIPIDEMIA NEC/NOS	3	1.95	16	10.39
40391: HYP RENAL NOS W REN FAIL (Begin 1989)	3	1.95	19	12.34
4280 : CONGESTIVE HEART FAILURE	3	1.95	22	14.29
5849 : ACUTE RENAL FAILURE NOS	3	1.95	25	16.23
5990 : URIN TRACT INFECTION NOS	3	1.95	28	18.18
25000: DIABETES UNCOMPL TYPE II	2	1.30	30	19.48
2639 : PROTEIN-CAL MALNUTR NOS	2	1.30	32	20.78
The first 10 levels are displayed.				

ICD 9 diagnosis labels from https://www.hcup-us.ahrq.gov/db/tools/I9_Formats.TXT

Bringing Everything Together

```
create table studyDat as
  select a.DESYNPUF_ID,
         a.BENE_SEX_IDENT_CD                              as sex
                                                             format=$sex.,
         a.BENE_RACE_CD                                   as race
                                                             format=$race.,
         b.CLM_FROM_DT                                    as index_date,

          1 + b.CLM_THRU_DT - b.CLM_FROM_DT                as los
                                                             label = 'Length of stay (days)',
        (b.CLM_FROM_DT - a.BENE_BIRTH_DT)/365.25          as ageAtAdmit
                                                             label = 'Age at admission',
          b.spinCurve,
          b.acc_fall,

          case when c.DESYNPUF_ID is not missing
                then 1 else 0
           end                                             as read60
                                                             label = 'Re-admitted within 60 days'
                                                             format=yesno.,

         c.CLM_FROM_DT - b.CLM_THRU_DT                    as daysToRead
                                                             label = 'Days to re-admission'

  from library.beneFile as a inner join
       firstSurgRec09   as b on a.DESYNPUF_ID = b.DESYNPUF_ID left join
       firstReAdmit60   as c on a.DESYNPUF_ID = c.DESYNPUF_ID
  where a.year = 2009
  ;
```

Summarize and Publish!

Patient characteristics (N=207)		Re-admitted within 60 days	
		No	Yes
All	N	188	19
	%	91	9
DESYNPUF: Sex			
Male	Row %	88	12
	Col %	40	53
Female	Row %	93	7
	Col %	60	47
DESYNPUF: Beneficiary Race Code			
White	Row %	90	10
	Col %	83	95
Black	Row %	95	5
	Col %	11	5
Other	Row %	100	.
	Col %	3	.
Hispanic	Row %	100	.
	Col %	3	.
Age at admission	Mean	73	74
	StdDev	11.7	13.5
Length of stay (days)	Mean	5	6
	StdDev	4.7	4.7
Days to re-admission	Mean	.	28
	StdDev	.	17.4
Spinal curvature	Mean	8%	11%
Falls	Mean	2%	0%