1. Introduction (10 min) – Kern Singh, MD (Rush)
a. Exploding use of national database research
   i. Examples of major databases: NIS, ACS NSQIP, and NTDB
   ii. Increased reliance on these databases for conducting orthopedic research
       1. Researchers are drawn to these sources because of their advantages in providing powerful sample sizes, inclusion of patients from demographics across the country and world, and the ability to easily observe trends throughout time
   iii. Spine is leading the way in this

b. What databases are in use and how are they constructed
Administrative databases: These are typically claims-based datasets. Within the United States, diagnosis and procedural coding is required for billing of government and private insurance. This data is typically in the form of ICD-9 (diagnostic and procedural) and CPT codes (professional services).

1. Claims data is typically used in spine to described temporal trends, geographic variation, disparities, complications, outcomes, and resource utilization.
2. Common examples include Medicare Claims (government), and private insurance claims (UnitedHealth Care, BlueCross BlueShield, Humana, Market Scan, etc)
3. “Discharge” type databases such as the National Inpatient Sample (NIS), also are compiled using claims data.
   a. Only the inpatient stay

Clinical registries: Contain secondary data on patients with a specific diagnosis or procedure. Data is typically used to monitor patient outcomes with the purpose of improving patient safety and optimizing quality. Data is gathered directly from patient charts

1. Registries used in orthopaedic research exist at the regional, national, and international levels. Many were designed to specifically collect outcomes relevant to orthopaedics / spine such as short-term surgical complications, or long-term arthroplasty implant survival.
2. Common examples include American Joint Replacement Registry (AJRR), Kaiser Registry, NSQIP, Force-TJR, CJRR, International Joint Replacement registries.

National Inpatient Sample (NIS)2
1. Most frequently used database for orthopedic research.5
3. Designed to annually sample 20% of U.S. hospitals to collect private and public discharge claims that can be utilized to determine national estimates and changes across time.

American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP)3
1. Collects extensive data in order to track and compare risk-adjusted surgical outcomes in hospitals across the country.

2. National level registry type data
   a. Descendent of VASQIP
   b. 30-h-quality data collection

v. Medicare
   1. Maintained by the Centers for Medicare and Medicaid Services.
   2. Availability of outpatient and inpatient claims.
   3. Designed to allow for long-term follow up.

vi. National Trauma Data Bank (NTDB)¹
   1. Combination administrative and registry
   2. Collected annually and distributed as data quality reports, hospital benchmark statements, and data sets for research.
   3. Includes the injury severity score and how the injury occurred.⁴

vii. Pediatric databases
   1. Kid’s Inpatient Database KID
      a. Pediatric NIS
      b. Comprehensive database of hospital inpatient stays which tracks regional and national trends in hospital utilization, access, charges, and quality.
   2. NSQIP Pediatric

viii. Researchers, journals, and readers need to take into account the relative considerations of the different databases and the conclusions that they allow to be made.

2. What are the advantages and limitations of different data types (10 min)
   – Andrew Pugely, MD (University of Iowa / Wash U St. Lou)
   a. Administrative claims data features
      i. Advantages:
         1. High patient numbers
         2. Large representation of hospitals from across US
         3. Most claims data can be used to follow patients through episodes of care
            a. Typically irrespective of changing hospitals or provider
         4. Reliable for certain “billing related” data points such as patient death
      ii. Disadvantages:
         1. Often poor data granularity—reliance of ICD-9 claims coding
a. All data including patient demographics, comorbidities, outcomes, etc. is coded as ICD-9s
b. ICD-9 codes lack precision, (ie exact fusion levels, approach, laterality)
   i. Very poor different types of surgical site infections
   ii. Multiple codes for thrombo-embolic disorders, cardiac events, etc.
   iii. Anemia example, (Golinvaux, Bohl, Basques, Grauer, Spine 2014) 

c. Input down by providers and coders, non-billable coding less accurate
   i. ie rates of “hospital acquired conditions” such as UTIs, pressure wounds, central line infections, etc.
   ii. Poor or incomplete documentation

d. Issues with up-coding

2. Data is difficult and costly to obtain, manipulate, and analyze
   a. Specially trained statistical analyst often required
   b. Private claims data is not readily available
   c. Medicare claims data costs can exceed $300,000 (and limited to Medicare population)

3. Smaller datasets are samples, requiring extrapolation (e.g. Medicare 5%, NIS 20%)

4. Lack spine specific outcomes,
   a. no patient reported outcomes or health related quality of life measures

b. Clinical registry data features
   i. Contain data in heterogeneous forms, but typically database specific variables, some ICD-9, and CPT codes.
   ii. Advantages:
      1. Registry is designed for prospective data collection
      2. Variables collected are typically more precise.
         a. Specific definitions for patient demographics, comorbidities, outcomes.
      3. Data typically collected by person specifically trained personal.
         a. data may be audited for accuracy
         b. NSQIP example
      4. Data may include more relevant orthopaedic or spine related variables.
      5. Some contain patient reported outcomes
iii. Disadvantages:

1. Lower sample sizes
   a. Varies tremendously by registry type
   b. No universal spine / orthopaedic registry in the US
2. Variable follow up time frame
   a. 30 days, 90-days, to indefinite,
3. Many registries still lack “spine-specific” variables
4. Expensive to setup, maintain, and operate
5. Data not always available to researchers outside the system (e.g. Kaiser)

c. Conclusions

i. Summary
ii. Reconcile differences of similar cohort with different datasets.
   1. Grauer, Bohl et al. examples.3
iii. Implications of ICD-10
   1. Potential for improvement in variable precision, but there will be a big learning curve, and years of inaccurate data.

---

**TABLE II Comparison of Database Characteristics**

<table>
<thead>
<tr>
<th>Database</th>
<th>Patients Included</th>
<th>Payers</th>
<th>Coding Scheme</th>
<th>Comorbidities</th>
<th>Laboratory Results</th>
<th>Operative Variables</th>
<th>Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medicare</td>
<td>All Medicare beneficiaries</td>
<td>CMS</td>
<td>ICD-9</td>
<td>No</td>
<td>Continuous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Part A</td>
<td>100%, or samples</td>
<td>CMS</td>
<td>ICD-9</td>
<td>No</td>
<td>Continuous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Part B</td>
<td>100%, or samples</td>
<td>CMS</td>
<td>ICD-9 and CPT</td>
<td>No</td>
<td>Continuous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Summary</td>
<td>All beneficiaries</td>
<td>CMS</td>
<td>ICD-9 and CPT</td>
<td>No</td>
<td>Continuous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other insurance</td>
<td>All beneficiaries</td>
<td>Variable</td>
<td>ICD-9</td>
<td>No</td>
<td>Continuous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PearlDiver</td>
<td>Sample</td>
<td>Variable</td>
<td>ICD-9 and CPT</td>
<td>No</td>
<td>Continuous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MarketScan</td>
<td>Sample (large companies)</td>
<td>Variable</td>
<td>ICD-9 and CPT</td>
<td>No</td>
<td>Continuous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>United Healthcare and BCBS</td>
<td>Sample or all</td>
<td>Variable</td>
<td>ICD-9 and CPT</td>
<td>No</td>
<td>Continuous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Premier</td>
<td>Variable</td>
<td>Variable</td>
<td>ICD-9, CPT, and database specific variables</td>
<td>No</td>
<td>Continuous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discharge databases</td>
<td></td>
<td>Definitions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NIS</td>
<td>20% sample</td>
<td>All</td>
<td>ICD-9</td>
<td>No</td>
<td>No</td>
<td>Inpatient</td>
<td></td>
</tr>
<tr>
<td>KID</td>
<td>20% sample</td>
<td>All</td>
<td>ICD-9</td>
<td>No</td>
<td>No</td>
<td>Inpatient</td>
<td></td>
</tr>
<tr>
<td>SID</td>
<td>Variable</td>
<td>All</td>
<td>ICD-9</td>
<td>No</td>
<td>No</td>
<td>Variable</td>
<td></td>
</tr>
</tbody>
</table>

*All databases were based on claims data, and all databases had data collection done by coders. CMS = Centers for Medicare & Medicaid Services, ICD-9 = International Classification of Diseases, Ninth Revision, CPT = Current Procedural Terminology Code, BCBS = Blue Cross Blue Shield, NIS = Nationwide Inpatient Sample, KID = Kids' Inpatient Database, and SID = State Inpatient Databases.

Pugely et al. JBJS 2015.2
### TABLE III Strengths and Weaknesses of National Databases *

<table>
<thead>
<tr>
<th>Database</th>
<th>Time Trends</th>
<th>Geographic Variation</th>
<th>Patient Comorbidities</th>
<th>Inpatient Complications</th>
<th>Short-Term Complications</th>
<th>Long-Term Complications</th>
<th>Financial Analysis</th>
<th>Accessibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medicare</td>
<td>++</td>
<td>++</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>--</td>
</tr>
<tr>
<td>Part A</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>--</td>
</tr>
<tr>
<td>Part B</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>--</td>
</tr>
<tr>
<td>Summary</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>--</td>
</tr>
<tr>
<td>Other insurance</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>--</td>
</tr>
</tbody>
</table>
| PeerDiv
MarketScan    | +           | +                    | +                     | +                      | +                        | +                       | +                   | +            |
| United Healthcare | +           | +                    | +                     | +                      | +                        | +                       | +                   | --           |
| and SCDBS         | +           | +                    | +                     | +                      | +                        | +                       | +                   | --           |
| Premier           | +           | +                    | +                     | +                      | +                        | +                       | +                   | +            |
| Discharge databases |           |                      |                       |                        |                          |                         |                     |              |
| NE            | +           | +                    | +                     | +                      | +                        | +                       | +                   | +            |
| KID              | +           | +                    | +                     | +                      | +                        | +                       | +                   | +            |
| SID              | +           | +                    | +                     | +                      | +                        | +                       | +                   | +            |

*Values within the table correspond to how appropriately each database can evaluate various study questions. -- = unable to answer question (strong weakness); - = poorly suited to answer question, but can in limited fashion (minor weakness); + = able to answer question, but better database exists (minor strength); and ++ = best suited to answer question (major strength). SCDBS = Blue Cross Blue Shield, NS = National Inpatient Sample, KID = Kids’ Inpatient Database, and SID = State Inpatient Databases.

**Pugely et al. JBJS 2015.²**

### TABLE II Comparison of Database Characteristics

<table>
<thead>
<tr>
<th>Database*</th>
<th>Type</th>
<th>Patients Included</th>
<th>Payers*</th>
<th>Coding Scheme</th>
<th>Data Collector</th>
<th>Comorbidities*</th>
<th>Laboratory Results</th>
<th>Operative Variables</th>
<th>Follow-up</th>
</tr>
</thead>
</table>
| Regional clinical registries
CJR
MARCOI
SCOP
National clinical registries
AJR
FORCE-TJR
Kaiser
NTDB
NSQIP
ACS NSQIP
VAN-SQIP National certification registry
ABCS International registries
<table>
<thead>
<tr>
<th>Registry</th>
<th>All (select hospitals)</th>
<th>All</th>
<th>Definitions</th>
<th>Trained data abstractor</th>
<th>Definitions</th>
<th>Some</th>
<th>Yes</th>
<th>Continuous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registry</td>
<td>All (select hospitals)</td>
<td>All</td>
<td>Definitions</td>
<td>Trained data abstractor</td>
<td>Definitions</td>
<td>Some</td>
<td>Yes</td>
<td>Continuous</td>
</tr>
<tr>
<td>Registry</td>
<td>All (select hospitals)</td>
<td>All</td>
<td>Definitions</td>
<td>Trained data abstractor</td>
<td>Definitions</td>
<td>Some</td>
<td>Yes</td>
<td>Continuous</td>
</tr>
<tr>
<td>Registry</td>
<td>All (select hospitals)</td>
<td>All</td>
<td>Definitions</td>
<td>Trained data abstractor</td>
<td>Definitions</td>
<td>Yes</td>
<td>Yes</td>
<td>Continuous</td>
</tr>
<tr>
<td>Registry</td>
<td>All (select hospitals)</td>
<td>All</td>
<td>Definitions</td>
<td>Trained data abstractor</td>
<td>Definitions</td>
<td>Yes</td>
<td>Yes</td>
<td>Continuous</td>
</tr>
<tr>
<td>Registry</td>
<td>Sample (select hospitals)</td>
<td>All</td>
<td>Definitions</td>
<td>Trained data abstractor</td>
<td>Definitions</td>
<td>Yes</td>
<td>Yes</td>
<td>30 day</td>
</tr>
<tr>
<td>Registry</td>
<td>Sample</td>
<td>VA</td>
<td>Definitions</td>
<td>Trained data abstractor</td>
<td>Definitions</td>
<td>Yes</td>
<td>Yes</td>
<td>30 day</td>
</tr>
<tr>
<td>Registry</td>
<td>Part II, Board collection period</td>
<td>All</td>
<td>Definitions, less strict</td>
<td>Surgeon</td>
<td>Definitions</td>
<td>Yes</td>
<td>Yes</td>
<td>6 mo</td>
</tr>
<tr>
<td>Registry</td>
<td>All</td>
<td>All</td>
<td>Variable: surgeon and/or technical</td>
<td>Definitions</td>
<td>Variable</td>
<td>Yes</td>
<td>Continuous</td>
<td></td>
</tr>
</tbody>
</table>

3. Statistical overview (10 min) – Daniel Bohl, MD, MPH (Rush)
   a. Back to the basics
      i. Univariate analysis
         1. Why we do it
         2. What it tells us
         3. Type 1 error
         4. Type 2 error
         5. P-value is a probability score, not an importance score
   b. Dealing with confounding
      i. Multivariate analysis
         1. Standard methods (linear, logistic, poisson regression)
            a. Historical perspective
            b. How multivariate adjustment works
            c. Limitations
         2. Other methods: Propensity score adjustment and matching
            a. Probably does not add much over standard methods
c. Measures of association
   i. Linear regression reporting coefficients (interpretation)
   ii. Logistic regression reporting odds ratios (interpretation and limitation)
   iii. Poisson regression reporting relative risk (interpretation and reasons to use this over logistic regression)

d. Multiple comparisons
   i. The predicament
   ii. The options
      1. Do nothing: What is often done
      2. Bonferroni: Conservative option
      3. Other options
   iii. My recommendation: Define primary and secondary hypothesis
      1. Test primary hypothesis at p<0.05 level
      2. Test secondary hypotheses at Bonferroni-corrected level

e. Role of power analyses in database studies
   i. Pre-hoc power analyses are always worthwhile when truly pre-hoc – but they rarely are
   ii. Post-hoc power analyses are of limited utility

f. Making database analyses useful
   i. Predictive modeling: Calculators for risk for the patient sitting directly in front of you
      1. Example 1: NSQIP risk calculator- with limitations

4. Examples of several contrasting studies (10min) – Andre Samuel, BBA (Yale)
   a. NIS example: “National trends in the management of central cord syndrome: an analysis of 16,134 patients” (Brodell et al.)
      i. Summary of analysis
         1. National trends in surgical management
         2. Analysis of predictors for surgery
         3. Analysis of predictors for inpatient mortality
      ii. Advantages of NIS
         1. Nationally representative data
         2. Large sample size
      iii. Disadvantages of NIS
         1. Poor detail in ICD-9 codes
         2. Poor sensitivity of ICD-9 codes, especially procedure codes
      iv. Clinical impact
   b. NSQIP example: “Perioperative outcomes after cervical laminoplasty versus posterior decompression and fusion: Analysis of 779 patients in the ACS-NSQIP database” (Varthi et al.)
      i. Summary of analysis
CSRS ICL Symposium on National Database Research

1. Comparison of length of stay
2. Comparison of adverse event rates
3. Comparison of readmission rates

ii. Advantages of NSQIP
   1. 30-day outcomes
   2. Reliable chart-abstracted data elements

iii. Disadvantages of NSQIP
   1. Smaller patient samples
   2. No trauma patients

   c. NTDB example: “Delayed Surgery After Acute Traumatic Central Cord Syndrome Is Associated With Reduced Mortality” (Samuel et al.)
      i. Summary of analysis
         1. Analysis of time to surgery
         2. Surgical timing and minor adverse events
         3. Surgical timing and serious adverse events
         4. Surgical timing and inpatient mortality
      ii. Advantages of NTDB
         1. Prehospital data
         2. Procedural timing data (to the hour)
         3. Trauma populations
      iii. Disadvantages of NTDB
         1. No outpatient or long term follow-up data
         2. No spine specific data elements
         3. Low sensitivity of ICD-9 data elements

5. Closing remarks (5 min) – Jonathan Grauer, MD
   a. Why have these types of studies caught on?
   b. The bar is elevating and these studies are getting more rigorous
   c. Just because they are large patent numbers does not mean that they are clinically meaningful
   d. Expected evolution of databases themselves as well as the studies based on them

6. Discussion (15min)
References:

Singh section

Pugley section
3. Bohl, DD; Russo, GS; Basques, BA; Golinvaux, NS; Fu, MC; Long, WD, 3rd; and Grauer, JN: Variations in data collection methods between national databases affect study results: a comparison of the nationwide inpatient sample and national surgical quality improvement program databases for lumbar spine fusion procedures. J Bone Joint Surg Am, 96(23): e193, 2014.

Bohl section

Samuel section
2. Varthi AG, Basques BA, Bohl DD, Golinvaux NS, Grauer JN. “Perioperative Outcomes after Cervical Laminoplasting versus Posterior Decompression and