Population-based Health Services Research in the Era of Big Data

Leighton Chan, MD, MPH
Chief, Rehabilitation Medicine
Department
Clinical Center
NIH
Disclosures

• No financial disclosures
A “new” era: BIG DATA

- Advances in computing power - IBM Watson
- Increasing data capture - EMR
- We can use these data to:
  - Provide “decision support” to providers
  - Inform policy
Types of Data

• **Structured**: highly organized data containing pre-defined elements with standardized relationships to one another (e.g., data in a database)

• **Unstructured**: data that is not structured in a pre-defined manner
  - Clinical notes, Faxed documents
  - *Anything written in narrative*

• **Majority of medical data is unstructured**
Where Does Data Come From?

- **Primary Data** – generated for research purposes, including national surveys and disease registries
- **Secondary Data** – Secondary data is administrative/billing/encounter data
  - Often generated with utilization in mind
  - Enough data to make meaningful population-base conclusions
Common Sources of Secondary Data

- Populations:
  - Medicare (54 million)
  - Integrated health systems (e.g., Kaiser, 10 million)
  - Medicaid (55 million)
  - Pharmacies

- Billing /encounter/administrative data
  - Electronic Medical Record (RMR)
Why Study Medicare Patients?

- Largest purchaser of health care in the world
  - 54 million enrollees
  - $613 billion in expenditures in 2014
- The percentages
  - Almost 16% of U.S. budget outlays
  - 22% of all health care dollars in US
  - 26% of hospital spending
  - 22% of nursing home spending
  - 22% of physician billings
- Single data system
Who are Medicare patients?

- 83% age > 65
- <1% ESRD
- 16% disabled
- 20% in an HMO
Medicare

- FFS beneficiaries health care costs ~$8,000/yr
- Costs/beneficiary rise ~7 percent/yr
- Part A trust fund depletion projection: 2030
Organization of Medicare Data: Files

- Hospital stays
- Physician visits (inpt & outpt)
  - specialty
  - experience
- SNF, DME, & Home Health
- Hospital level data: size, non-profit status, staffing, etc.
Organization of Medicare Data: Variables

- Patient demographics:
  - age, sex, race, zip code
- Primary Diagnosis (ICD-9)
- Associated co-morbidities
- Procedure codes - CPT-4 & HCPCS
Organization of Medicare Data: Variables

• Hospitalization and Rehospitalization
• Sentinel Events
• Length of Stay
• Disposition
• Mortality
• Pharmacy use
• Costs
Benefits of Medicare Data

- Pre-existing data
  - less expensive
  - less time
- large numbers of cases
  - Generalizbility
- Links to other data
  - Zip codes, SSN
- Accurate measure of resource use
- Can measure “effectiveness”
Limitations of Medicare data

• Lots of limitations...
  ▫ Limited data on severity of illness
  ▫ Not generalizable to the US working population
  ▫ Coding and billing errors/bias
  ▫ Limited outcome measures of interest
    • No QOL, Patient Satisfaction, functional assessment, illness severity
Limitations

• Studies limited to non-experimental design (observational studies)
• Difficult to avoid selection bias
• Impossible to control for all possible confounders (e.g. severity of illness & functional status)
• HMO patients are excluded
• Cost of obtaining the data
• Administrative overhead
How can I get data access?

• Medicare administrative/billing/encounter data
  ▫ http://www.resdac.org/about-resdac/our-services
  ▫ Available through CCW Data Enclave
Types of Studies using Medicare data

- Monitoring secular trends
- Measuring disparities
  - Race, ethnicity, SES, geographic variation
- Supporting the evaluation of specific conditions, treatments or procedures
Monitoring Secular Trends

- Examine changes in health care over time
- Take advantage of “Natural Experiments”
  - Examine the impact of policy changes
  - e.g., Epidural Steroid Injections for Back Pain
    - 14,000 recently exposed to contaminated steroids
Complications and Mortality Associated With Cervical Spine Surgery for Degenerative Disease in the United States

Marjorie C. Wang, MD, MPH,* Leighton Chan, MD, MPH,† Dennis J. Maiman, MD, PhD,*
William Kreuter, MPA,† and Richard A. Deyo, MD, MPH†

Increases in Lumbosacral Injections in the Medicare Population
1994 to 2001

Janna Friedly, MD,* Leighton Chan, MD, MPH,* and Richard Deyo, MD, MPH†

Geographic Variation in Epidural Steroid Injection Use in Medicare Patients
Janna Friedly, Leighton Chan and Richard Deyo
Percent of US Population with LBP

Source: National Center for Health Statistics, National Ambulatory Medical Care Survey: 1998-2004
## Prevalence of Low Back Pain (>age 65)

<table>
<thead>
<tr>
<th>Year</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995</td>
<td>15%</td>
</tr>
<tr>
<td>2000</td>
<td>20%</td>
</tr>
<tr>
<td>2005</td>
<td>28%</td>
</tr>
<tr>
<td>2010</td>
<td>32%</td>
</tr>
</tbody>
</table>

*One or more MD encounter for LBP*
Medicare patients 2002-2006

- Nonspecific backache 60.3%
- Degenerative changes 14.7%
- Sciatica 11.8%
- Spinal Stenosis 7.3%
LBP Costs

- Deyo, MEPS
- Costs are very high ($86B)
- Mean adjusted costs
  - 1997 $4,695 (95% CI, $4,181-$5,209)
  - 2005 $6,096 (95% CI, $5,670-$6,522)
    - 30% increase in costs
  - Self-reported measures of mental health, physical functioning, work or school limitations, and social limitations among adults with spine problems were worse in 2005 than in 1997.
Definitions

Epidural Steroid Injections

Transforaminal
Are ESIs Effective in LBP?

Conflicting data on effectiveness of ESIs

- 18-90% success rates
- Few good quality studies (Freidly, 2014)
- No consistency in methods
  - Patient selection
  - Technique used
  - Definition of success
- Recent incidents confirm that the risks of the procedure are greater than suspected
Methods

• Retrospective Cohort Study
• Medicare claims data
• 1995-2006, 5%-20% sample of physician bills
• Cohorts defined by CPT and ICD-9 codes
  • 62311 Caudal or interlaminar
  • 64483 Transforaminal
  • 64475 Facet Injection
• LBP dx from ICD-9 codes
• Physician specialty from UPIN
ESIs/1000 Medicare Enrollees

450%
SI Joint Injections/1000 Persons with "SI" Disease
Facet (Medial Branch) Neurotomies

Procedures/1000 Persons with LBP

>7 fold increase
Who Is Performing ESIs?

1995
- Anesthesia: 79%
- Surgery: 10%
- PM&R: 3%
- Radiology: 1%
- Pain: 0%
- Other: 6%

2005
- Anesthesia: 56%
- Surgery: 15%
- PM&R: 7%
- Radiology: 4%
- Pain: 11%
- Other: 7%
What are the indication for ESIs?

2001 ESI Diagnoses

- Radiculopathy: 33%
- Herniated Disc: 9%
- Degenerative Changes: 11%
- Spinal Stenosis: 23%
- Other Back Pain: 25%
How much has this increase cost Medicare?
Inflation adjusted to 2001
Total Estimated Costs

Physician Professional Fees $175 million
+ Facility Fees $275 million

Total Cost to Medicare = $450 million
Where ESIs are Performed?

- Outpatient Hospital Clinics
- Physician Offices
- Ambulatory Surgical Centers
Ambulatory Surgery Centers (ASCs)

Most Medicare certified ASCs are: (n=5,000)

- privately owned
- for profit
- urban locations
Ambulatory Surgery Centers (ASCs)

- Supposed to reduce costs by avoiding hospital overhead
- Majority owned by local physician investors
- The Stark self-referral law (1989 Social Security Act) does not apply to ASCs
  - MDs can invest in ASCs and increase revenue by receiving ASC facility payments
Ambulatory Surgery Center Symposium - Baltimore, Maryland

The Backbone of a Healthy ASC: Chronic Spinal Pain Management as a Revenue Strategy

8 am - 5 pm, Wednesday, July 18, 2007
Hosted by Kimberly-Clark & GE Healthcare

GE Healthcare
NUMBER OF MEDICARE-CERTIFIED ASCs

ASC Advantages

- More convenient locations, shorter wait times
- Medicare coinsurance is lower than in hospitals
  - ($9 difference in 2004)
- Customized environments, specialized staffing
- Customer friendly
ESIs at ASCs: 1995-2005

1995: 13% of ESIs performed at ASCs

2005: 29% of ESIs performed at ASCs

2005 Lumbosacral Injection Facilities

- Physician Offices: 18%
- Ambulatory Surgery Centers: 29%
- Outpatient Hospitals: 53%
Results

ASC ESI Facility Payment*

1995: $7.5 million
2005: $101 million
>1200% increase
Conclusions

• Lots of growth in ESI
• Growth associated with shift in “injectionists”
• Growth associated ASC growth
• Significant cost increases for Medicare
• Do patients benefit?
Measuring geographic variation

- Tom Wennberg & Alan Gittelsohn
- Examine procedure rates in different geographic areas
- If rates differ this suggests inequity or inefficiency in practice
2001 Epidural Steroid Injection Rates

ESI Rate Quartiles

1st: High (n=12)
2nd (n=11)
3rd (n=15)
4th: Low (n=13)

Alabama 40/1000

Hawaii 5/1000
Geographic Variations

Health Referral Regions (HRR):
• Smaller geographic regions
• Defined by Dartmouth’s Atlas for Health Care (http://www.dartmouthatlas.org/)
• 306 HRRs across the country
• Defined by where most of the cardiovascular and neurosurgery is performed
Comparative magnitude of variability of orthopedic procedures

From Weinstein, et. al., Spine, 2006
2005 Geographic Variations: Health Referral Regions:

- 9-fold difference in ESIs/1000 patients
  - 7.9/1000 in Honolulu, HI
  - 103.6/1000 in Palm Springs, CA
From: Geographic Variation in Epidural Steroid Injection Use in Medicare Patients

Geographic Variations

Summary

1. Large geographic variations in ESI use

2. High ESI rates are not associated with lower surgery rates

3. High ESI rates are moderately associated with “injectionist” supply
Limitations

- Only study those over age 65 in Medicare
  - No young active workers
  - No HMO patients
- Possible errors in diagnosis/billing codes
- Which ESI rate is right?
  - No data on pain relief
  - No data on return to work
  - No data on functional improvement
Unanswered Questions

• Are ESIs effective?
  ▫ How do we select the ideal patients for ESIs?
  ▫ How many should we be doing?
  ▫ How often should we be doing them?
  ▫ Should we be doing them with other treatments? (i.e. multidisciplinary approach)
Analyzing a specific procedure

- Bariatric surgery for obesity
Early Mortality Among Medicare Beneficiaries Undergoing Bariatric Surgical Procedures

David R. Flum, MD, MPH
Leon Salem, MD
Jo Ann Broeckel Elrod, PhD
E. Patchen Dellinger, MD
Allen Cheadle, PhD
Leighton Chan, MD, MPH

In the United States, most adults are overweight or obese, and obesity is soon to become the leading cause of death. Bariatric surgical procedures are the only interventions that consistently help patients achieve significant and sustained weight loss and improvements in comorbid medical conditions. As a result, there has been dramatic growth in bariatric surgery over the last decade, with interest in applying it to those at high risk based on associated medical conditions and the growing population of older, obese patients. Balanced against these beneficial effects, however, are the risks of perioperative death and short-term adverse outcomes. These risks have been poorly defined in the community at large, with the expected rates largely derived from case series. Several case series demonstrate that bariatric surgery can be performed with a low rate of perioperative mortality (0.5%), but the rate among high-risk patients and the community at large is unknown.

Objectives To evaluate the risk of early mortality among Medicare beneficiaries and to determine the relative risk of death among older patients.

Design Retrospective cohort study.

Setting and Patients All fee-for-service Medicare beneficiaries, 1997-2002.

Main Outcome Measures Thirty-day, 90-day, and 1-year postsurgical all-cause mortality among patients undergoing bariatric procedures.

Results A total of 16,155 patients underwent bariatric procedures (mean age, 47.7 years [SD, 11.3 years]; 75.8% women). The rates of 30-day, 90-day, and 1-year mortality were 2.0%, 2.8%, and 4.6%, respectively. Men had higher rates of early death than women (3.7% vs 1.5%, 4.8% vs 2.1%, and 7.5% vs 3.7% at 30 days, 90 days, and 1 year, respectively; P<.001). Mortality rates were greater for those aged 65 years or older compared with younger patients (4.8% vs 1.7% at 30 days, 6.9% vs 2.3% at 90 days, and 11.1% vs 3.9% at 1 year; P<.001). After adjustment for age and comorbidity index, the odds of death within 90 days were 5-fold greater for older Medicare beneficiaries (aged ≥75 years; n=136) than for those aged 65 to 74 years (n=1,381; odds ratio, 5.0; 95% confidence interval, 3.1-8.0). The odds of death at 90 days were 1.6 times higher (95% confidence interval, 1.3-2.0) for patients of surgeons with less than the median surgical volume of bariatric procedures (among Medicare beneficiaries during the study period) after adjusting for age, sex, and comorbidity index.

Conclusions Among Medicare beneficiaries, the risk of early death after bariatric surgery is considerably higher than previously suggested and associated with advancing age, male sex, and lower surgeon volume of bariatric procedures. Patients aged 65 years or older had a substantially higher risk of death within the early postoperative period than younger patients.

JAMA. 2005;294:1903-1908

www.jama.com
Bariatric Surgical Procedures Mortality Study

• Dramatic growth in bariatric surgery in obese adults
  ▫ the number between 1998 and 2004 from 13,386 to 136,000
• No national coverage decision or consensus regarding efficacy and safety in older adults
Bariatric Surgical Procedures Mortality Study

- **Objective:**
  - Evaluate the risk factors of early mortality among Medicare beneficiaries (age, gender, surgeon experience)
  - Determine relative risk of death among older patients
- **Retrospective cohort design using Medicare physician bills, (1996-2002), 16,155 cases**
Survival After Bariatric Surgery by Age Group

# Mortality Rate After Bariatric Surgery, by Age and Sex

<table>
<thead>
<tr>
<th>Age Category (y) and Sex</th>
<th>No.</th>
<th>30 Days</th>
<th>90 Days</th>
<th>1 Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>150</td>
<td>0.7</td>
<td>1.3</td>
<td>2.0</td>
</tr>
<tr>
<td>Men</td>
<td>53</td>
<td>0.0</td>
<td>1.9</td>
<td>1.9</td>
</tr>
<tr>
<td>Subtotal</td>
<td>203</td>
<td>0.7</td>
<td>1.5</td>
<td>2.0</td>
</tr>
<tr>
<td>25-34</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>1341</td>
<td>0.8</td>
<td>1.3</td>
<td>2.5</td>
</tr>
<tr>
<td>Men</td>
<td>486</td>
<td>2.1</td>
<td>3.3</td>
<td>4.3</td>
</tr>
<tr>
<td>Subtotal</td>
<td>1827</td>
<td>1.1</td>
<td>1.8</td>
<td>3.0</td>
</tr>
<tr>
<td>35-44</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>3288</td>
<td>1.0</td>
<td>1.5</td>
<td>2.7</td>
</tr>
<tr>
<td>Men</td>
<td>1121</td>
<td>3.2</td>
<td>3.7</td>
<td>5.6</td>
</tr>
<tr>
<td>Subtotal</td>
<td>4409</td>
<td>1.5</td>
<td>2.0</td>
<td>3.4</td>
</tr>
<tr>
<td>45-54</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>4214</td>
<td>1.1</td>
<td>1.8</td>
<td>3.1</td>
</tr>
<tr>
<td>Men</td>
<td>1191</td>
<td>4.5</td>
<td>5.4</td>
<td>7.7</td>
</tr>
<tr>
<td>Subtotal</td>
<td>5405</td>
<td>1.9</td>
<td>2.6</td>
<td>4.1</td>
</tr>
<tr>
<td>55-64</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>2126</td>
<td>2.0</td>
<td>2.5</td>
<td>4.7</td>
</tr>
<tr>
<td>Men</td>
<td>668</td>
<td>2.1</td>
<td>3.1</td>
<td>6.9</td>
</tr>
<tr>
<td>Subtotal</td>
<td>2794</td>
<td>2.0</td>
<td>2.7</td>
<td>5.2</td>
</tr>
<tr>
<td>65-74</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>1039</td>
<td>2.6</td>
<td>3.4</td>
<td>6.2</td>
</tr>
<tr>
<td>Men</td>
<td>342</td>
<td>5.8</td>
<td>8.2</td>
<td>12.9</td>
</tr>
<tr>
<td>Subtotal</td>
<td>1381</td>
<td>3.4</td>
<td>4.6</td>
<td>7.8</td>
</tr>
<tr>
<td>≥75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>85</td>
<td>18.8</td>
<td>28.2</td>
<td>40.0</td>
</tr>
<tr>
<td>Men</td>
<td>51</td>
<td>19.6</td>
<td>35.3</td>
<td>51.0</td>
</tr>
<tr>
<td>Subtotal</td>
<td>136</td>
<td>19.1</td>
<td>30.9</td>
<td>44.1</td>
</tr>
<tr>
<td>Total</td>
<td>16 155</td>
<td>2.0</td>
<td>2.8</td>
<td>4.6</td>
</tr>
</tbody>
</table>

Rate of Early Mortality After Bariatric Surgery, Stratified by Surgeon Volume*

<table>
<thead>
<tr>
<th>Annual Surgeon Volume†</th>
<th>No.</th>
<th>30 Days</th>
<th>90 Days</th>
<th>1 Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients aged &lt;65 y</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;15</td>
<td>3200</td>
<td>2.2</td>
<td>3.0</td>
<td>5.0</td>
</tr>
<tr>
<td>15-35</td>
<td>3191</td>
<td>1.7</td>
<td>2.2</td>
<td>3.5</td>
</tr>
<tr>
<td>36-70</td>
<td>3295</td>
<td>1.7</td>
<td>2.3</td>
<td>4.2</td>
</tr>
<tr>
<td>71-268</td>
<td>3205</td>
<td>1.2</td>
<td>1.8</td>
<td>3.1</td>
</tr>
<tr>
<td>Total</td>
<td>12891</td>
<td>1.7</td>
<td>2.3</td>
<td>4.0</td>
</tr>
<tr>
<td>Patients aged ≥65 y</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;15</td>
<td>480</td>
<td>9.0</td>
<td>13.8</td>
<td>21.0</td>
</tr>
<tr>
<td>15-35</td>
<td>282</td>
<td>3.2</td>
<td>4.6</td>
<td>6.4</td>
</tr>
<tr>
<td>36-70</td>
<td>284</td>
<td>1.8</td>
<td>2.1</td>
<td>4.2</td>
</tr>
<tr>
<td>71-268</td>
<td>274</td>
<td>1.1</td>
<td>1.1</td>
<td>3.6</td>
</tr>
<tr>
<td>Total</td>
<td>1320</td>
<td>4.5</td>
<td>6.7</td>
<td>10.7</td>
</tr>
</tbody>
</table>

*Numbers are adjusted to reflect complete-case-only analysis.
†Number of open bariatric procedures in Medicare beneficiaries, 1997-2002.

Limitations

- CPT codes not precise enough to determine exact procedure
- Analysis of missing data?
- Only those over age 65
- Surgeon volume related to Medicare pts only
Bariatric Surgery Mortality Study

• Conclusions:
  ▫ Medicare beneficiaries >=65 nearly 3 fold increase in risk of early mortality
  ▫ Post-operative mortality rates associated with:
    • Advancing age
    • Male sex
    • Lower surgeon volume bariatric procedures
  ▫ Medicare allowed those over 65 to have surgery, but only at those sites that did more than 125/year and by surgeons doing more than 50/yr
  ▫ 2010 follow-up study showed a reduction in LOS and complications after the NCD
Conclusions

- Medicare Claims data is widely used in outcomes research
- The data has significant advantages and disadvantages
- The importance of this data will increase as the US demographics change and Medicare enrollment accelerates
CCW Data Enclave

- CMS developed virtual data access for investigators through a new data enclave
- Cost determined by the number of licenses (seats) in the enclave, and not by the amount of data requested
- Users are assigned a dedicated workspace within the CCW Virtual Data Enclave where they can directly access approved CMS data and run analyses in SAS
- Users may:
  - Upload external files to their Data Enclave workspace for use with CMS data
  - Download aggregate, statistical files to their workstations
NLP: The Next Big Thing

• Unstructured data
  ▫ There is a lot!
  ▫ Medical history
  ▫ Large numbers of patients
  ▫ Machine learning methods: identify patterns, trends, and long-term changes

• Need input from clinician and a linguist
• Successful pre-processing critical
NLP Application to HF

- Identification of Framingham HF criteria in PCP notes
- Based on Unstructured Information Management Architecture (UIMA) framework
- Partnership between IBM T.J. Watson Research Center, Geisinger Medical Center, and Sutter Health

**NLP Application to HF**

- Iterative process: cardiologist, linguist, and coders in partnership
  - Cardiologist and linguist jointly review case files for key words and linguistics
  - Linguist builds NLP tools
  - Joint review of outcomes; extraction improvement
  - Coders create a “gold standard” for comparison

Promising Results

- High accuracy in identification of Framingham HF criteria
  - Few false negatives: successfully identified 90% of true positives
  - Few false positives: >92% of cases labeled positive were true positives
- Demonstrates PC notes can be successfully extracted
- Shows potential for early identification methods