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 & HCPICs

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 (eg. 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

SPINE Volume 32, Number 3, pp 342–347
©2007, Lippincott Williams & Wilkins, Inc.

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‡

SPINE Volume 32, Number 16, pp 1754–1760
©2007, Lippincott Williams & Wilkins, Inc.

Increases in Lumbosacral Injections in the Medicare Population

1994 to 2001

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

The Journal of Bone & Joint Surgery

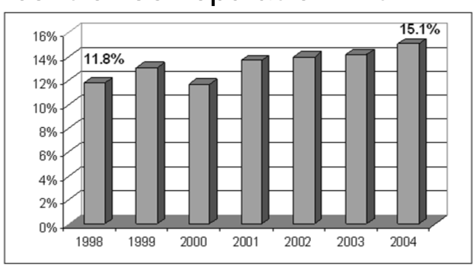
JB&JS

This is an enhanced PDF from The Journal of Bone and Joint Surgery
The PDF of the article you requested follows this cover page.

Geographic Variation in Epidural Steroid Injection Use in Medicare Patients

Janna Friedly, Leighton Chan and Richard Deyo
J Bone Joint Surg Am. 2008;90:1730-1737. doi:10.2106/JBJS.G.00858

Percent of US Population with LBP



Source: National Center for Health Statistics, National Ambulatory Medical Care Survey: 1998-2004

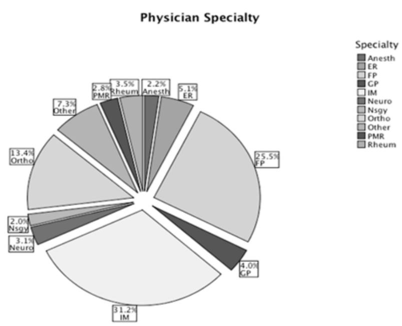
Prevalence of Low Back Pain (>age 65)

1995	15%
2000	20%
2005	28%
2010	32%

*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

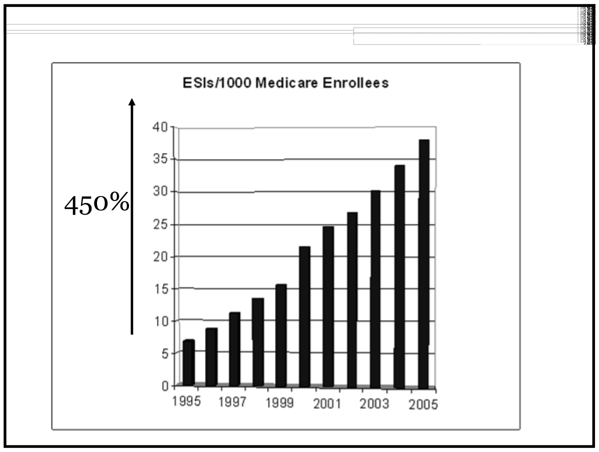
Transforaminal

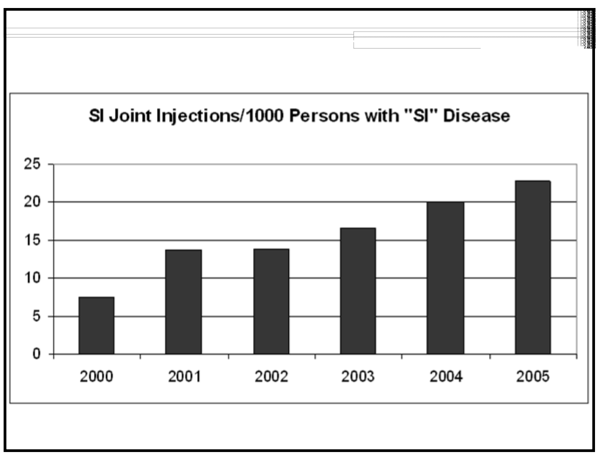
Are ESIs Effective in LBP?

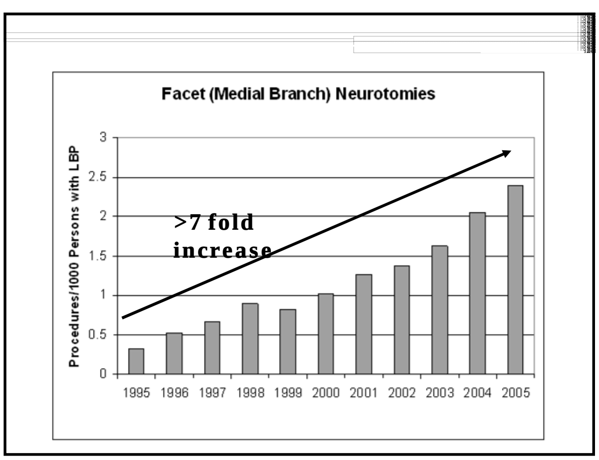
Conflicting data on effectiveness of ESIs
- 18-90% success rates
- Few good quality studies (Freidly, 2014)
- No consistency in methods
  - Patient selection
  - Technique used
  - Definition of success
- Recent incidents confirm that the risks of the procedure are greater than suspected

Methods

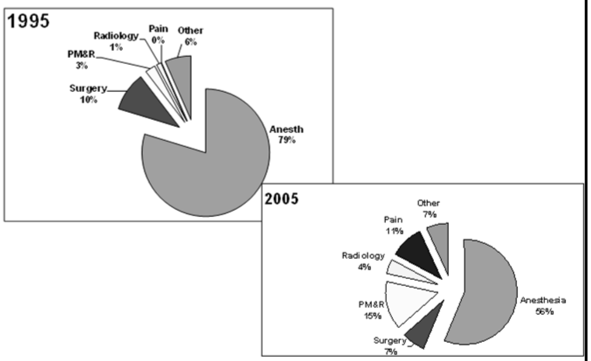
- Retrospective Cohort Study
- Medicare claims data
- 1995-2006, 5%-20% sample of physician bills
- Cohorts defined by CPT and ICD-9 codes
  - 62311 Caudal or interlaminar
  - 64483 Transforaminal
  - 64475 Facet Injection
- LBP dx from ICD-9 codes
- Physician specialty from UPIN


ESIs/1000 Medicare Enrollees
450%
40
35
30
25
20
15
10
5
0
1995
1997
1999
2001
2003
2005

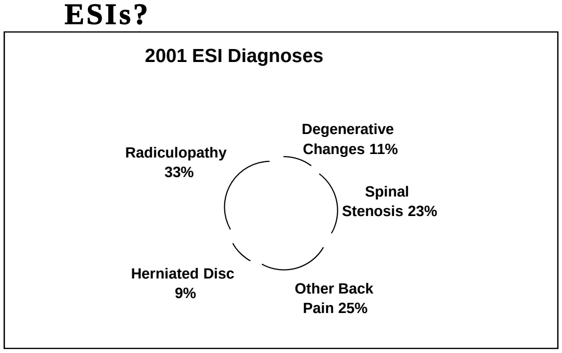

SI Joint Injections/1000 Persons with "SI" Disease
25
20
15
10
5
0
2000
2001
2002
2003
2004
2005


Facet (Medial Branch) Neurotomies
Procedures/1000 Persons with LBP
>7 fold increase
3
2.5
2
1.5
1
0.5
0
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005

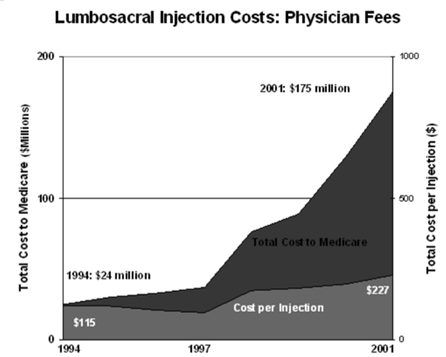
Who Is Performing ESIs?



What are the indication for ESIs?



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

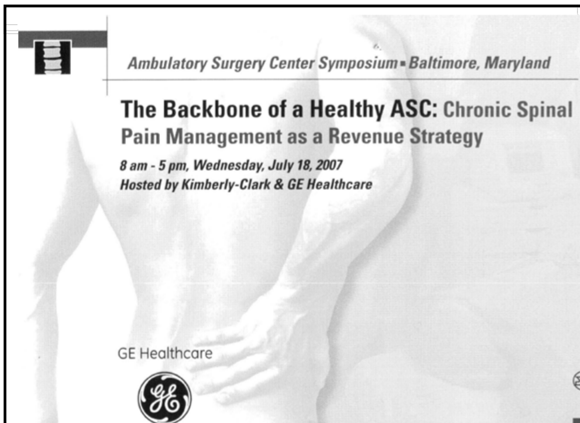
Ambulatory Surgery Centers (ASCs)

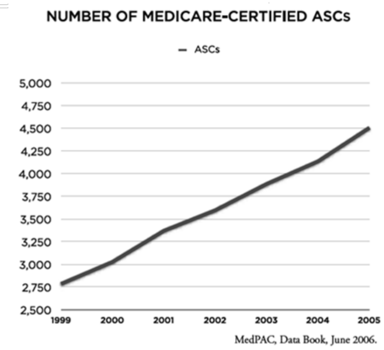
Most Medicare certified ASCs are: (n=5,000)

- 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





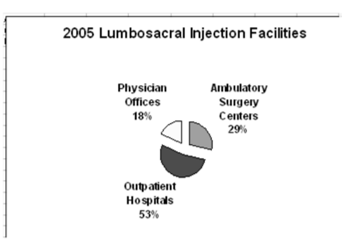
ASC Advantages

- 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



Results

ASC ESI Facility Payment*

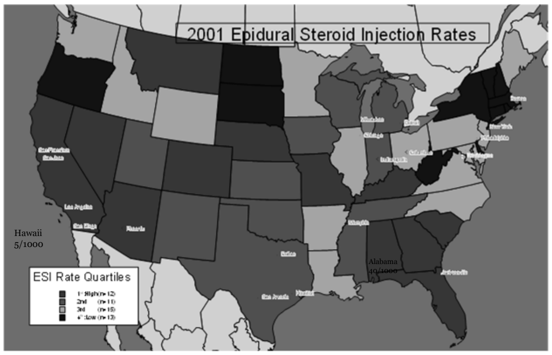
1995: $7.5 million
2005: $101 million
>1200% increase

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

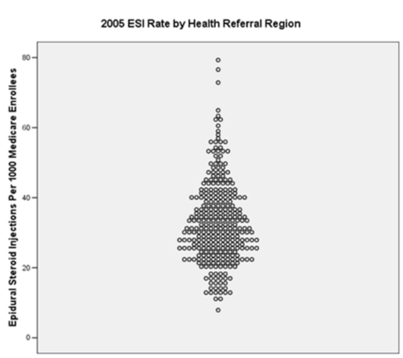
- Tom Wennberg & Alan Gittelsohn
- Examine procedure rates in different geographic areas
- If rates differ this suggests inequity or inefficiency in practice



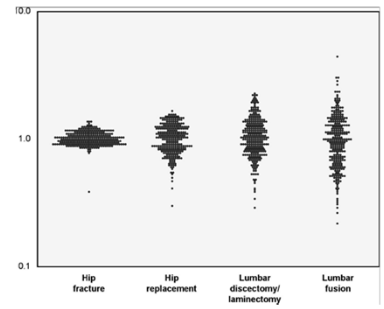
Geographic Variations

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

Geographic Variation



Comparative magnitude of variability of orthopedic procedures



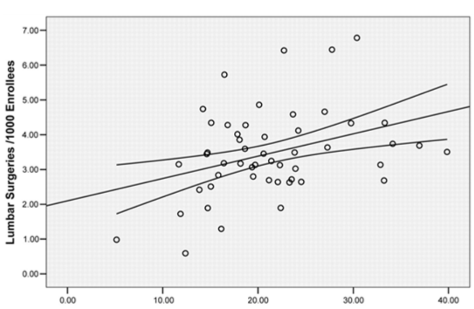
From Weinstein, et. al., Spine, 2006

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

J Bone Joint Surg Am. 2008;90(8):1730-1737. doi:10.2106/JBJS.G.00858



Date of download: 12/10/2012

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

ORIGINAL CONTRIBUTION

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

Context 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 sex 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 ($n=1381$; 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

IN THE UNITED STATES, MOST ADULTS are overweight or obese,[1] and obesity is soon to become the leading cause of death. Bariatric surgical procedures are the only interventions[2] that consistently help patients achieve significant and sustained weight loss and improvements in comorbid medical conditions.[3-5] 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.[6] 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.[7] Sev-

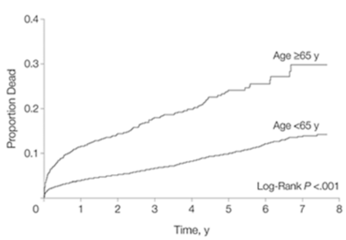
Bariatric Surgical Procedures Mortality Study

- 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



No. at Risk Age, y								
≥65	1517	1350	1126	708	453	282	150	52
<65	14 638	14 062	12 486	8647	5866	3842	2174	759

Flum, D. R. et al. JAMA 2005;294:1903-1908.

JAMA

Mortality Rate After Bariatric Surgery, by Age and Sex

Table 2. Mortality Rate After Bariatric Surgery, by Age and Sex

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

Flum, D. R. et al. JAMA 2005;294:1903-1908.

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

Table 4. Rate of Early Mortality After Bariatric Surgery, Stratified by Surgeon Volume*

		Mortality Rate, %		
Annual Surgeon Volume†	**No.**	**30 Days**	**90 Days**	**1 Year**
Patients aged <65 y				
<15	3200	2.2	3.0	5.0
15-35	3191	1.7	2.2	3.5
36-70	3295	1.7	2.3	4.2
71-268	3205	1.2	1.8	3.1
Total	12 891	1.7	2.3	4.0
Patients aged ≥65 y				
<15	480	9.0	13.8	21.0
15-35	282	3.2	4.6	6.4
36-70	284	1.8	2.1	4.2
71-268	274	1.1	1.1	3.6
Total	1320	4.5	6.7	10.7

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

Flum, D. R. et al. JAMA 2005;294:1903-1908.

JAMA

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

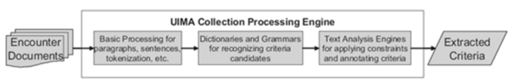
- Unstructured data
  - There is a lot!
  - Medical history
  - Large numbers of patients
  - Machine learning methods: identify patterns, trends, and long-term changes
- Need input from clinician and a linguist
- Successful pre-processing critical

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

Source: R.J. Byrd, et al., Automatic identification of heart failure diagnostic criteria, using text analysis of clinical notes from electronic health records, Int. J. Med. Inform. (2013), http://dx.doi.org/10.1016/j.ijmedinf.2012.12.005

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

Source: R.J. Byrd, et al., Automatic identification of heart failure diagnostic criteria, using text analysis of clinical notes from electronic health records, Int. J. Med. Inform. (2013), http://dx.doi.org/10.1016/j.ijmedinf.2012.12.005

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

Source: R.J. Byrd, et al., Automatic identification of heart failure diagnostic criteria, using text analysis of clinical notes from electronic health records, Int. J. Med. Inform. (2013), http://dx.doi.org/10.1016/j.ijmedinf.2012.12.005