Use of Large Databases for Cancer Prevention and Control Research

Gregory S. Cooper, MD
Case Western Reserve University
Goals

• Introduction to large database research
• Colon cancer screening
• Screening for Barrett’s esophagus
Existing Cancer Data Sources

• Screening
  – Telephone surveys (BRFSS)
  – Single health plan/institution

• Incidence/Treatment
  – Population-based registries (state, region)
  – Health plan/institution

• Follow Up
  – Health plan/institution
Limitations of Existing Data

• Generalizability
• Screening
  – Recall biases
  – Oversample higher SES
• No national-level registry
• Cancer patients
  – No pre- or post-treatment data
“Breakable bones, a tendency to bleed when cut, vulnerability to germs and viruses. These are all preexisting conditions.”
Advantages of Claims Data

• Nationally Representative
  – Community practice
  – Different treatment locations
  – Efficacy vs. effectiveness
• Large Sample Sizes
• Little Incremental Research Costs
• Turnaround Time Typically Low
Disadvantages

- Not Designed for Research
  - Nonstandardized definitions
  - No data on severity
- Coding Inaccuracies
  - Systematic
  - Random
- Missing Patients (VA, Age Cutoff, HMO, Outpatient)
Sources of Claims Data

- Federal Government
  - Medicare
  - VA
- State Level
  - Medicaid
  - Hospital discharges
- Private Insurers
Types of Medicare Files

- Part A 100% of patients (hospitalizations, skilled nursing short stay)
- Part B > 95% of patients (extra premium, outpatient treatment, physician services)
- All contain demographics, diagnosis (ICD-9) and procedure codes (ICD-9 hospital, CPT-4 outpatient services)
Diagnosis Codes - Example

- Enterostomy Malfunction
- Intestinal Adhesion with Obstruction
- Depressive Disorder
- Surgical Complication GI Tract
- Postoperative Infection
Procedure Codes - Example

- Pericolostomy Hernia Repair
- Peritoneal Adhesiolysis
- Culture-Peritoneum
- Culture and Sensitivity-Lower Resp Tract
- Pulmonary Artery Wedge Monitor
- Insert Endotracheal Tube
- Mechanical Respiratory Assistance
- Small Bowel Series
Example - Narrative

Basic Methodology

- Exclude HMO, patients < 65
  - Incomplete claims
- Linkage with Other Data
- Select Diagnosis or Procedure
- Cohort of Interest
- Exclude Previous Diagnosis (Prevalent cases)
Appropriate Conditions to Study

• Common diagnoses
• Seen in Medicare age population
• Result in hospitalization or encounter
• Coded frequently
  – Impact on reimbursement
Potential Linkages

• Socioeconomic Measures
  – Census data (proxy)

• Tumor Registry
  – SEER-Medicare

• Physician and Provider
  – AMA Master File, AHA
SEER-Medicare Database

- Collaborative effort with NCI and CMS
  - patients ≥ 65 years linked by unique identifiers
- SEER population-based registries (states and metro areas)
  - Covers ~ 25% of US population
- All Medicare files
  - inpatient and outpatient records
storks
babies

pairs of brooding storks

millions of newborn babies

Colon Cancer Screening

What is current status of colon cancer screening?
What interventions protect against cancer development?
Physician Specialty and Endoscopic Procedures

• 5% sample of Medicare claims, 1993
• Procedures examined
  – Colonoscopy
  – EGD
  – Sigmoidoscopy
• Medicare and AMA designated specialties
  Meyer, J Gen Intern Med 2000
Colonoscopy by Specialty

- GI
- GenSurg
- ColSurg
- GIM
- FP/GP
Flexible Sigmoidoscopy Use
Procedure Characteristics

- Specialists more likely to perform therapeutics, cancer or bleeding indications
- Generalists more likely in underserved counties
- Most colonoscopies and EGD’s by PCP’s in counties with at least one GI
Hospital Type and Patient Outcomes

• SEER-Medicare database 1984-93
• Major surgery: Whipple, esophagectomy, pneumonectomy, liver resection (colorectal), pelvic exenteration
• Hospitals classified by total number of procedures
• Adjusted for age, stage and comorbidity

Begg, JAMA 1998
Current Status of Screening

- Use largely determined by telephone surveys
- Differences among population subgroups
- Impact of 1998 legislation
  - Colonoscopy in “high risk” every 2 years (fam hx, hx polyp or cancer, IBD)
  - Yearly FOBT, sigmoidoscopy every 4 years, screening barium enema as substitute
Cohort

- 1997-1999 Medicare Physician/Supplier and Outpatient files
- Claims for FOBT, flex sig, BE, colonoscopy
- Two comparison procedures (UGI and EGD)
- Approximately 25 million beneficiaries/year
Rates of Fecal Occult Blood Test in the Medicare Population (age $\geq 65$)

3-year average, 1997-1999

Proportion with FOBT

- $\leq 12.18$
- 12.19-13.14
- 13.15-16.61
- 16.62-20.90
- $\geq 20.91$
Rates of Colonoscopy in the U.S. Population (age > 65)

3-year average, 1997-1999
Procedure Use by Quarter-Males

![Graph showing procedure use by quarter for males. The x-axis represents quarters from 1997 to 1999, and the y-axis represents volume per 100,000. The graph includes lines for FOBT, FS, CY, and BE, each with distinct markers and colors.]
Procedure Use in Men
Three Year Average

- FOBT: 59%
- Sig: 61%
- Colon: 14%
- BE: 18%
- EGD: 10%
- UGI: 16%
Procedure Use in Women
Three Year Average

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Annual %</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOBT</td>
<td>47%</td>
</tr>
<tr>
<td>Sig</td>
<td>33%</td>
</tr>
<tr>
<td>Colon</td>
<td>0%</td>
</tr>
<tr>
<td>BE</td>
<td>15%</td>
</tr>
<tr>
<td>EGD</td>
<td>18%</td>
</tr>
<tr>
<td>UGI</td>
<td>16%</td>
</tr>
</tbody>
</table>

Cau: [Bar Graph with data]
AfAm: [Bar Graph with data]
Racial Differences in Indications for FOBT

Annualized %

21% 13% 81% 67%

Cau AfAm

DIAGNOSTIC SCREENING
Racial Differences in Indication for Colonoscopy

Annualized %

<table>
<thead>
<tr>
<th>Category</th>
<th>M</th>
<th>W</th>
<th>M</th>
<th>W</th>
<th>M</th>
<th>W</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIAG</td>
<td>20%</td>
<td>17%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SURV</td>
<td></td>
<td></td>
<td>117%</td>
<td>63%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCREEN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>30%</td>
<td>9%</td>
</tr>
</tbody>
</table>

Cau — Cau; AfAm — AfAm
Summary

- Use of screening procedures lower than targets
- Little impact of reimbursement changes
- Use of colonoscopy increased despite reimbursement issues
- Racial disparities persist
Barrett’s Screening

Does endoscopic screening improve outcome?
Prior EGD for Esophageal Adenocarcinoma

- Screening for and Surveillance of Barrett’s Esophagus Recommended
- Goal to Detect Presymptomatic High Grade Dysplasia and Adenocarcinoma
- Actual Use and Potential Impact of Screening Not Well Defined
Cohort

- SEER-Medicare Database
- Patients with Adenocarcinoma of the Esophagus (n=777) or Cardia (n=856)
- Diagnosed 1993-6
- Age ≥ 70
- Non-HMO enrollees
- Staged
Measures

- EGD Prior to Diagnosis
  - > 6 months
  - > 1 year
- Prior Diagnosis of Barrett’s Esophagus
- Stage at Diagnosis
- Survival - unadjusted, adjusted
Rate of Barrett’s Diagnosis & Prior EGD

% Patients

<table>
<thead>
<tr>
<th></th>
<th>Esophagus</th>
<th>Cardia</th>
</tr>
</thead>
<tbody>
<tr>
<td>BE 6 mos</td>
<td>7.5</td>
<td>1.3</td>
</tr>
<tr>
<td>EGD 6 mos</td>
<td>14.7</td>
<td>8.6</td>
</tr>
<tr>
<td>EGD 1 yr</td>
<td>13</td>
<td>6.8</td>
</tr>
</tbody>
</table>
Staging Comparisons: 1 Year
EGD & Esophageal Cancer

% of Patients

<table>
<thead>
<tr>
<th></th>
<th>In Situ</th>
<th>Local</th>
<th>Regional</th>
<th>Distant</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGD</td>
<td>12.2</td>
<td>34.1</td>
<td>21.6</td>
<td>16.2</td>
</tr>
<tr>
<td>No EGD</td>
<td>1.4</td>
<td>50.0</td>
<td>20.7</td>
<td>23.9</td>
</tr>
</tbody>
</table>

p<0.001
Staging Comparisons: 1 Year
EGD & Cardia Adenocarcinoma

% of Patients

- In Situ: 1.3 vs. 0.0
- Local: 26% vs. 48.9%
- Regional: 28.9% vs. 35.1%
- Distant: 37.6% vs. 22.2%

p<0.01
Outcomes of Patients With and Without Prior EGD

- Median Survival Time
  - Cardia: 8 months EGD vs. 6 months others
  - Esophagus: 7 months EGD vs. 5 months others

- Adjusted Survival
  - Cardia: Hazard Ratio 0.87 (0.63-1.20)
  - Esophagus: Hazard Ratio 0.73 (0.57-0.93)
Conclusions

- EGD Performed > 1 Year Prior to Diagnosis of Esophageal or Cardia Adenocarcinoma Associated with Earlier Stage at Diagnosis and Improved Survival
- Most Patients at Risk Appear Undiagnosed
Summary - 1

• Claims data provide a unique opportunity to study:
  – Epidemiological measures
  – Patterns of care/practice patterns
  – Treatment/outcome differences
  – Provider performance

• Data are linkable with registry, population demographics, provider characteristics
Summary - 2

• Complexity of files
  – inpatient alone simplest but also most limited
  – outpatient and pharmacy data complex

• Limitations of data
  – adequate severity adjustment
  – incident vs. prevalent cases
  – miscoding (random vs. systematic)