Overview of the SEER-Medicare Data

Rachel Ballard-Barbash, MD, MPH
Martin Brown, Ph.D.
Joan Warren, Ph.D.

Applied Research Program
DCCPS

BSA Meeting
November 2, 2009
What are the SEER - Medicare data?

The SEER Program:
- NCI has contracted with universities and state health departments since 1973 to operate population-based cancer registries
- SEER collects data on incident cancer cases
- Since 1992, SEER has included 12 geographic areas, 14.5% of U.S. population; expanded in 2001 to include 26% of U.S. population

SEER-Medicare data:
- NCI matches people in the SEER data to the Medicare’s master enrollment file
- For SEER cases who have Medicare, NCI obtains all claims for Medicare covered health services
- There are currently over 1.5 million persons age 65+ in the files
- Years of data currently available:
  - SEER cases from 1973-2005
  - Medicare claims from 1991-2007
Persons included in the SEER-Medicare Data

- 100% of patients in the SEER data who are found to be Medicare eligible

- 5% random sample of persons residing in the SEER areas who have not been diagnosed with cancer
  - Comparison groups for assessing screening, diagnostic testing and treatment practices in the 65+ population
What is included in the SEER-Medicare Data?

• **SEER Data includes:**
  – Incidence, anatomic site, stage, initial treatment, demographics and vital status, cause of death

• **Medicare claims for:**
  – Short stay hospitals
  – Physician and lab services
  – Hospital outpatient claims
  – Home health and hospice bills

• Recurrences/progression not reported from either data source
Why Link the SEER-Medicare Data?

Linked data can address questions across the cancer continuum

<table>
<thead>
<tr>
<th>Screening / Detection</th>
<th>Diagnosis / Treatment</th>
<th>Survivorship → Second Cancer</th>
<th>Death / Terminal Care</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of PSA testing, mammography sigmoidoscopy/colonoscopy</td>
<td>Patterns of care</td>
<td>Late effects of treatment</td>
<td>Rates of second primaries</td>
</tr>
<tr>
<td>Impact of new technology or practice patterns on cancer detection/incidence</td>
<td>Peri-operative complications</td>
<td>Post-diagnostic testing</td>
<td>Relationship of second events to initial treatment and ongoing surveillance</td>
</tr>
<tr>
<td></td>
<td>Volume outcomes studies</td>
<td>Treatment of prevalent cancers</td>
<td>Use of hospice</td>
</tr>
<tr>
<td></td>
<td>Extent of staging</td>
<td></td>
<td>End-of-life patterns of care</td>
</tr>
<tr>
<td></td>
<td>Comorbidities</td>
<td></td>
<td>Survival</td>
</tr>
</tbody>
</table>

Health disparities, quality of care and cost of treatment
Growth of SEER-Medicare for Research

The SEER and Medicare data were first linked in 1992 for the purpose of a single research project to assess the cost of care. Since then, research uses and users have increased significantly.
Advantages of Using SEER-Medicare Data

- **SEER-Medicare data:**
  - include large numbers of cases
  - are longitudinal - from the time of Medicare coverage until death
  - span most clinical areas where health care is delivered
  - represent a diversity of geographic areas across the U.S.
  - are population-based and thus reflect “real world” practice
  - include data on multiple disease conditions so can adjust for complexity of disease and care
Key Limitations of the SEER-Medicare Data

- Observational data, thus selection bias
- Non-covered services excluded: prescription drugs (until 7/2006), long-term care, free screenings
- Reasons for & results of tests/procedures not known
- Limited population
  - Does not include claims for persons in HMOs (~ 22% in SEER areas)
  - Under 65 population includes only the disabled/ESRD
- Lag of 4 years to obtain linked data (not “rapid”)
More Details on the SEER-Medicare data

SEER-Medicare WEB site
appliedresearch.cancer.gov/seermedicare

The WEB site has information on
• Publications using the data (>400)
• How to obtain the files
• Some technical support
Comparative Effectiveness Research Using SEER-Medicare Data

Nancy L. Keating, MD, MPH
Associate Professor of Medicine & Health Care Policy
Harvard Medical School
Brigham and Women’s Hospital
Overview

• Rationale
• Examples
  – Adjuvant chemotherapy for colon cancer
  – Primary treatment for prostate cancer
    • Evaluation of a new technology
    • Survival after primary prostate cancer treatment
  – Comparisons of systems of care
• Methodology for observational data
Rationale for Using SEER-Medicare Data for Comparative Effectiveness Research

• Limitations of randomized controlled trials
  – Not feasible for many interventions
  – Certain populations underrepresented (elderly, sick, minorities, low SES)

• Population-based observational data better than single institution studies

• Longitudinal data with large N’s from various regions across US

• Statistical methods available to address nonrandom assignment
Example 1: Adjuvant Chemotherapy for Stage III Colon Cancer

• Adjuvant chemotherapy improves survival for stage III colon cancer
  – Older patients underrepresented in trials

• Low rates of adjuvant chemotherapy for older patients
  – Is this appropriate patient selection or underuse of effective care?
Example 1: Adjuvant Chemotherapy for Stage III Colon Cancer

Propensity-Adjusted Kaplan Meier Survival Curve

Iwashyna and Lamont, JCO 2002
Remaining Questions

• Were treatment groups similar?
  – Or were treated patients healthier?

• Since results are consistent with RCT, does this prove effectiveness in this population?

• Would we believe results if they contradicted RCT findings?
Example 2: Minimally-Invasive vs. Open Radical Prostatectomy

- Minimally-invasive radical prostatectomy (MIRP) has diffused rapidly in recent years
- Few data about benefits over open surgery
- High costs to adopt technology
Example 2: Minimally-Invasive vs. Open Radical Prostatectomy

- Overall complications:
  - MIRP: 22.2%
  - ORP: 23.2%
  - P = 0.58

- Respiratory:
  - MIRP: 4.3%
  - ORP: 6.6%
  - P = 0.04

- Genitourinary:
  - MIRP: 4.7%
  - ORP: 2.1%
  - P = 0.01

- Anastomotic stricture:
  - MIRP: 14%
  - ORP: 5.8%
  - P < 0.001

Hu, Keating et al, JAMA 2009
Example 2: Minimally-Invasive vs. Open Radical Prostatectomy

Short-term Outcomes

<table>
<thead>
<tr>
<th></th>
<th>Rate per 100 person years</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aditional cancer therapy</td>
<td>8.2</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>6.9</td>
<td></td>
</tr>
<tr>
<td>Incontinence</td>
<td>15.9</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>12.2</td>
<td></td>
</tr>
<tr>
<td>Erectile dysfunction</td>
<td>28.6</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>19.2</td>
<td></td>
</tr>
</tbody>
</table>

Hu, Keating et al, JAMA 2009
Remaining Questions

• Are men in both groups similar?
  – Or are men who choose MIRP more likely to complain about incontinence and erectile dysfunction post operatively?

• Are urologists in SEER-Medicare cohort representative of surgeons elsewhere
  – Steep learning curve, outcomes related to experience
  – SEER areas include 2 very high volume areas

• Would patient-reported outcomes differ?
Example 3: Survival Following Primary Treatment of Localized Prostate Cancer

• Uncertainty about benefits of treatments due to lack of clinical trials

• Most men with prostate cancer will not die of their cancer

• What can we learn about long term outcomes after prostate cancer treatment?
Example 3: Overall Survival Following Primary Treatment for Prostate Cancer

Wong et al. JAMA 2006
But Treated Patients have BETTER Survival than Controls without Cancer

Giordano et al, Cancer 2008
Challenges in Examining Associations of Treatment and Survival

• Selection to different treatments not random
  – Health status of patient an important unmeasured confounder

• Difficult to account for follow up care, additional treatments, new comorbidities
Comparisons of Systems of Care

• Improvements to care delivery a priority

• SEER-Medicare data can be used for comparisons if other similar data available
  – Veterans Health Administration (VHA)
    • Cancer registry data linked with administrative data on visits, medications, labs
Example 4: Cancer Care in the VHA vs. SEER-Medicare for Older Men

Adjusted Rates of Adjuvant Therapy for Colorectal Cancer

- Adjuvant Chemo for Colon
  - VHA: 65.1%
  - SEER-Medicare: 63.8%
  - P = 0.66

- Adjuvant Chemo & XRT for Rectal
  - VHA: 58.7%
  - SEER-Medicare: 53.3%
  - P = 0.31

Keating et al, unpublished
Challenges to Comparisons of Care Systems

- Patients may differ
  - Veterans typically have more comorbid illness and are of lower socioeconomic status

- Data may differ
  - Administrative data in VHA not used for billing purposes
Methods for Observational Data

- Standard regression methods typically insufficient for observational data analyses
- Propensity score methods and instrumental variables methods can help to address unobserved confounding
  - But may not overcome all biases
- Development and application of new methods for rigorous observational studies greatly needed
  - This will be an active area of research under Comparative Effectiveness Research initiatives
Future Plans for SEER-Linked Data

• Enhancing Data Resources

• Facilitating Sophisticated Research Uses

• Communicating Research Results
Enhancing Data Resources: Potential for Augmenting SEER-Medicare for CER

• Addition of Part D medication data to SEER-Medicare
  – In data validation phase – not yet approved for public release

• Expansion to non-SEER area registries with Medicare data
  – Only 19 of the 65 NCI cancer centers are in SEER areas
  – CDC exploring potential to link Medicare data to some state registries
  – Could enhance ability to study effect of health care systems on care

• Link SEER to claims data sources other than Medicare
  – Medicaid: enrollment data would be helpful; claims data of uncertain quality and timeliness.
  – Private insurance claims, i.e. BCSC (Health Core), however, proprietary nature of data have limited release for research
Enhancing Data Resources: Other SEER Linkages

- **SEER–Medical Health Outcomes Survey (SEER-MHOS)**
  
  
  - Includes HRQOL data from Medicare beneficiaries in HMOs
  - Data set spans from 1998 – 2004
  - 40,000+ cancer patients and survivors; 200,000+ respondents never diagnosed with cancer (controls)

- **SEER–Consumer Assessment in Healthcare Providers and Systems (SEER-CAHPS) in development**
  
  - CAHPS items include reports about care (e.g., how well doctors communicate) and ratings of care (e.g., health plan, primary or specialist care)
  - One CAHPS survey includes Medicare beneficiaries enrolled in Medicare managed care plans and includes over 2 million beneficiaries for the period 1998-2007
  - The second CAHPS survey includes 1.5 million Medicare beneficiaries enrolled in Medicare fee-for-service from 2000-2007
Facilitating Sophisticated Research Uses: Training

• Issue: Increased use of SEER-Medicare data, has resulted in an increase in inexperienced users

• NCI Response: Training and Technical Advice
  – SEER-Medicare training offered by NCI bi-annually
  
  – Contracts awarded to provide more support throughout the process, including understanding data as investigators design research studies
  
  – SEER-Medicare WEB site has been enhanced to include technical advice and a question function which allows investigators to submit and receive timely responses to questions
Facilitating Sophisticated Research Uses: Methods and Validation Research

• Issue: The expanded use of SEER-Medicare data has resulted in projects that propose questions for which SEER-Medicare data may not be appropriate

• NCI Response: Support validation research to examine if SEER-Medicare data can address key questions
  – Can SEER-Medicare data be used to identify:
    • persons with disease recurrence/progression?
    • specific chemotherapy regimens and dosage?
  – CER: Assessing treatment effectiveness using observational data is complicated because of non-random assignment. Can new methods control for differences between groups?
Facilitating Sophisticated Research Uses: Tools

• SEER*Stat software allows investigators to obtain rapid answers using variables collected by SEER

• A comparable “SMART” software system using variables from the SEER-Medicare data is under development. This will be able to provide rapid answers for questions such as:
  – Chemotherapy use
  – Costs of care
  – Comorbidities
Example of the SMART system: Use of chemotherapy among Medicare patients with Stage III colon cancer by age group and race

**Statistics**

**Statistic:** Frequencies

**Selection**

Case: {Patient Characteristics & Diagnosis.Modified AJCC stage 3rd ed (1988+)} = '10', '20', '30'
AND {Patient Characteristics & Diagnosis.Site rec with Kaposi and mesothelioma} = 'Colon excluding Rectum'
AND {Year-Specific Variables.Chemotherapy flag} = 'Yes chemotherapy given'
AND {Year-Specific Variables.Follow-up year} = 'Year 1 (dx to dx+11)'
AND {Patient Characteristics & Diagnosis.Year of diagnosis} = '2001', '2002'
AND {Year-Specific Variables.Complete and continuous entitlement flag} = 'Complete entitlement'

**Table**

Row: Race (White, Black, Other) [Race]
Column: Age at diagnosis - 10 year groups [Age at diagnosis recode]

<table>
<thead>
<tr>
<th></th>
<th>65-74 years</th>
<th>75-84 years</th>
<th>85+ years</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>1,768</td>
<td>1,325</td>
<td>154</td>
</tr>
<tr>
<td>Black</td>
<td>169</td>
<td>97</td>
<td>9</td>
</tr>
<tr>
<td>Other</td>
<td>138</td>
<td>88</td>
<td>4</td>
</tr>
</tbody>
</table>
Communicating Research Results

- Researchers
  - Publications and meetings
- Methods
  - Journal supplements
  - Web
- Clinicians
  - Work with DCTD/DCP
- Policy
- Public
CER related SEER-Medicare publications