Risk Adjustment

What data are there in administrative files for risk adjustment and how can we code them?

Paul L. Hebert, Ph.D.
Department of Health Services
University of Washington School of Public Health

Basic elements of a Quasi-Experimental CE study

- Basic elements of CER include
  - Rigorous study design
  - Appropriately chosen study sample
  - Appropriately chosen treatment groups
  - Defined outcomes
  - Measures of patient demographics, co-morbidity and severity of illness
  - Statistical analysis to deal with potential confounding

Measuring Covariates

- Demographics
- Socio-demographics
- Comorbidity
- Severity (good luck)
Demographics

- Focus on Medicare data because coding of race on other administrative data sources are idiosyncratic
  - Example: HCUP. Some states do not require the collection of race and so race is coded infrequently or not at all for all hospitals in the state.
- The CMS Beneficiary Summary File (BSF) contains the following demographic information
  - Date of birth (from which age can be calculated)
  - Gender
  - Office of Management and Budget (OMB) and RTI defined race groups
    - White
    - Asian/Pacific Islander
    - African American
    - Native American
    - Hispanic
    - Other /Unknown

Percentage Distribution of Medicare Enrollees by Race, 2008
BSF race code and RTI Hispanic Code

Demographics

- Problems with the coding of race in the beneficiary summary file
  - Hispanic is a mutually exclusive category.
    - There is no such thing as Black-Hispanic, Non-Hispanic Black, or Hispanic Asian in Medicare
    - Based mostly on an historic programming decision, not on policy
  - Asians are a single category
Sociodemographics (SES)

What measures of socioeconomic status (SES) are available in Medicare files?

- **Two Sources of SES information**
  - Medicare buy-in indicator
  - County and ZIP code identifiers to link with Census-based measures of SES

What measures of socioeconomic status (SES) are available in CMS files?

- County and ZIP code identifiers to link with Census-based measures of SES
  

  5-digit ZIP code can be merged to Census based ZIP Code Tabulation Areas (ZCTA)

  Important SES measures to consider:
  - Median household income
  - Percent of population with income below federal poverty limit
  - Percent of population over 18 with < high school education

---

**Sociodemographics**

- For low income beneficiaries, Medicare will pay for Part B benefits and the copayments and deductibles for Medicare allowed expenses.

<table>
<thead>
<tr>
<th>Buy-in</th>
<th>Code in BSF</th>
<th>% of beneficiaries</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>124,006</td>
<td>75.2%</td>
</tr>
<tr>
<td>Part A only</td>
<td>&lt;0.01</td>
<td></td>
</tr>
<tr>
<td>Part B only</td>
<td>3,239</td>
<td>2.0%</td>
</tr>
<tr>
<td>Part A+B</td>
<td>37,678</td>
<td>22.8%</td>
</tr>
<tr>
<td>Total</td>
<td>164,918</td>
<td></td>
</tr>
</tbody>
</table>

Note: this does not necessarily mean the beneficiary is receiving full Medicaid benefits (e.g., Nursing home care)
Sociodemographics (SES)

What measures of socioeconomic status (SES) are available in CMS files?

- Problems with Census-based measures of SES
  ZIP Code Tabulation Areas (ZCTA) are not perfectly matched to ZIP codes (Krieger 2002)
  Only available for Census years
  Between Census years county level data are available from BLS, but counties can be huge and diverse


Measuring Covariates

- Demographics
- Socio-demographics
- Comorbidity
- Severity (good luck)

Comorbidity

- Definition: A comorbidity is a disease or condition that coexists with the primary condition of interest and affects health outcomes
- Measuring comorbidity using CMS data
  Diagnosis based measures
  Cost-validated measures
- How do you incorporate comorbidities into a statistical model in CER?
Measuring Comorbidity

Diagnosis based measures

- Look for diagnoses for conditions on Medicare claims
- Group conditions into a small number of conditions
- Optionally: Give a weight to each of these conditions reflecting its importance to the outcome
- Result is a series of indicators for each condition, and an overall comorbidity score or index
- Widely used comorbidity indexes
  - Charlson comorbidity
  - Elixhauser

Diagnosis based measures

- Charlson comorbidity
  - 18 conditions abstracted from hospital charts used to predict in one-year mortality
  - Adapted for use with administrative data in several iterations. Romano is commonly used (Romano et al, 1993).
  - Software is available here: [http://healthservices.cancer.gov/surveillance/programs/charlson.comorbidity.macro.txt](http://healthservices.cancer.gov/surveillance/programs/charlson.comorbidity.macro.txt)
  - MARSHALL: WE SHOULD PUT HEIDI's ON the RESDAC SITE

- Elixhauser (Elixhauser et al, 1998)
  - 30 conditions developed to predict length of stay, hospital charges, and in-hospital death
  - Software is available here: [http://www.hcup-us.ahrq.gov/toolssoftware/comorbidity/comorbidity.jsp](http://www.hcup-us.ahrq.gov/toolssoftware/comorbidity/comorbidity.jsp)

References:
- Charlson/Romano conditions
- Some important conditions aren't included in Romano/Charlson...
Some important conditions aren’t included in Romano/Charlson... or in Elixhauser.
Measuring Comorbidity

Which should you use?

Luckily you no longer have to choose because Gagne et al (2010), have combined them and the combined measure works better than either individually.

And, software is available here:

http://www.drugepi.org/downloads/

Note: A good percentage of patients have zero or only 1 or two comorbid conditions recorded in a single year. (57% had <=2 comorbid conditions in our sample)
Measuring Comorbidity

- Get diagnosis codes from administrative data to create comorbidity scores
- Many administrative databases are developed from the same basic claim form
- On claims based on HCFA 1500 claims form (e.g., CMS Physician/Supplier) there are 4 diagnosis fields

Measuring Comorbidity

Diagnosis based measures

- Many institutional (e.g., hospital) administrative claims databases are developed from the UB-04 claim form
  CMS MedPAR (hospital), skilled nursing facility, home health, and outpatient hospital claims
  On MedPar claims,
  Admitting diagnosis: initial reason given for hospitalization. Don’t generally use this to create comorbidity scores because it is superseded by the primary diagnosis
  Primary diagnosis: the condition most responsible for the hospitalization.
  Secondary diagnoses: Other important conditions.
  Use primary and secondary dx’s to create comorbidity score
  “Present on admission” indicators exist (and are almost always yes)

Measuring Comorbidity

Diagnosis based measures based on CMS data

- If you are using CMS data, use indicators from Chronic Conditions Warehouse (CCW)
  http://ccwdata.org/
  Included in the Beneficiary Annual Summary File (request this separately from the Beneficiary Summary File)
  Indicators for the presence of 21 chronic conditions
  Yearly indicators going back to 1999.
  Standard definitions: specific criteria for reference time periods, diagnosis and procedure codes, number/type of qualifying claims (e.g., must have 2 Carrier claims during reference time period), and coverage
  Ever date: Date when the standard definition was first met.
Measuring comorbidity
Example of algorithm from Chronic Conditions Warehouse

<table>
<thead>
<tr>
<th>Condition</th>
<th>Diagnosis Code</th>
<th>Comorbidity Code</th>
<th>Comorbidity Description</th>
<th>Time Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypertension</td>
<td>401.0</td>
<td>401.0</td>
<td>Hypertension</td>
<td>2011-2012</td>
</tr>
<tr>
<td>Diabetes</td>
<td>250.0</td>
<td>250.0</td>
<td>Diabetes</td>
<td>2011-2012</td>
</tr>
<tr>
<td>Heart Disease</td>
<td>429.0</td>
<td>429.0</td>
<td>Heart Disease</td>
<td>2011-2012</td>
</tr>
</tbody>
</table>

Comorbidity

- **Definition:** A comorbidity is a disease or condition that coexists with the primary condition of interest and affects health outcomes
- **Measuring comorbidity using CMS data**
  - Outcomes-validated based measures
  - Cost-validated measures
- **How do you incorporate comorbidities into a statistical model in CER?**

Comorbidity

Cost-based measures (very briefly)

- Software is available to group patients according to the cost category
  - Diagnostic Cost Groups/Hierarchical Condition Codes (Pope et al, 2004). 70 codes that represent groups of diagnosis codes that are similar in etiology and cost implications
  - Ambulatory Care Groups (Weiner et al, 1991) are 51 categories developed from ambulatory claims
  - Validated for use in cost studies and therefore maybe useful for CER, but rarely used

References:
Comorbidity
Cost-based measures (very briefly)

- Software is available to group patients according to the cost category
  - Diagnostic Cost Groups/ Hierarchical Condition Codes (Pope et al, 2004).
  - 70 codes that represent groups of diagnosis codes that are similar in etiology and cost implications
  - Ambulatory Care Groups (Weiner et al, 1991) are 51 categories developed from ambulatory claims
  - RxRisk (aka Chronic Disease Score) (Fishman et al, 2003): Uses prescription drug claims to group medications into hierarchical groups similar.

Each measure has been validated for use in cost studies and are sometimes used for CER, but usually they perform worse than other measures (e.g., Schneeweiss et al., 2001)

Comorbidity
Definition: A comorbidity is a disease or condition that coexists with the primary condition of interest and affects health outcomes

- Measuring comorbidity using CMS data
  - Outcomes-validated measures
  - Cost-validated measures

How do you incorporate comorbidities into a statistical model in CER

Comorbidities measures are dummy (0/1) variables that indicate whether they have the condition

Charlson/Romano and Gagne can be combined into a summary score based on the sum of the weights provided by the authors.

Two options:
1. Use the summary score only in the statistical model, i.e.:
   \[
   \log\{\text{Odds}\} = \beta_0 + \beta_1 \text{Denom} + \beta_2 \text{SES} + \beta_3 \text{Treatment} + \beta_4 \text{Charlon}
   \]
2. Use each of the dummy variables individually, i.e.,
   \[
   \log\{\text{Odds}\} = \beta_0 + \beta_1 \text{Denom} + \beta_2 \text{SES} + \beta_3 \text{Treatment} + \beta_4 \text{AMI} + \beta_5 \text{CHF} + \beta_6 \text{PVD} + \beta_7 \text{CVD} + \ldots + \beta_n \text{IVF}
   \]
Comorbidity
Incorporating comorbidities measures into a statistical model in CER

- Which should you do?
  - If your sample size is large enough, use each individual measure as an explanatory variables (Schneeweiss et al, 2003)
  - If you are doing a logistic regression, you may have to drop or combine some categories of rare diseases (mild liver disease is often a culprit).
  - If every person who has the condition has the same outcome, the model will spit those observations out.
  - If there are other conditions important to your outcome but not coded or coded imprecisely with existing measure then include them as a separate variables as well.


Measuring Covariates

- Demographics
- Socio-demographics
- Comorbidity
- Severity (good luck)
Severity Measurement

Good luck

- Not as straightforward because severity not directly coded
- Can be a major source of unobserved confounding in claims based analysis
  Two people may have generated claims for heart failure, but one is NYHA Class I and spends his winters skiing, while the other is Class IV and is short of breath at rest. CMS claims cannot distinguish between them
- Can approximate via complications
- Examples:
  - Diabetes complications severity index (Young 2008 AJMC)
  - Complicated versus uncomplicated pneumonia (DRG 89 vs DRG 90)

Summary: Measuring Covariates

- Demographics in administrative data:
  - Institutional and state-specific differences in race information
  - For CMS data race is getting better but still limited.
- Socio-demographics:
  - For CMS data Medicaid buy-in is a good proxy for low income.
  - Otherwise you must use Census-based proxies
- Comorbidity
  - Several good diagnosis- and cost-based options
  - Use the Chronic Conditions Warehouse if you are using CMS data
- Severity (good luck)

References & Resources

References & Resources
