

# Analytical Issues: Using Linked MAX/NHANES Data to Study Obesity Costs

August 7, 2012

Presentation to the National Conference on Health Statistics

Allison Hedley Dodd, PhD

**MATHEMATICA**  
Policy Research

# Overview

- **Medicaid Analytic eXtract (MAX) background**
- **Analytical issues for MAX cost data**
  - **Linking NHANES to multiple MAX records**
    - Multi-year Medicaid enrollment
    - Inherent MAX issues
  - **Factors that affect Medicaid cost data**
    - Dual status
    - Benefit status
    - S-CHIP status
    - Managed care enrollment
  - **State variation in Medicaid data**

---

# Background

# MAX Background

---

- **Generated from quarterly Medicaid Statistical Information System (MSIS) files submitted by state**
- **Seven quarters processed to create annual files with adjudicated claims and reconciled enrollment records**
- **Goes through validation process to identify (and possibly) fix data issues**
- **Data issues summarized in anomaly tables (available on CMS website)**

# MAX Background

- **Annual set of files:**
  - **Claims files**
    - Inpatient (IP)
    - Long-term care (LT)
    - Prescription drug (RX)
    - Other (outpatient services/home-health care/medical equipment ) (OT)
  - **Person Summary file (PS)**
    - Enrollment information
    - Summary claims information
  
- **PS file sufficient for most analyses**

# Medicaid Eligibility Data

- **Basis of Eligibility (BOE) category**
  - Child
  - Adult
  - Disabled (should be non-aged)
  - Aged
- **BOE provided by state**
- **Eligibility coded monthly**
- **Eligibility=Enrollment**
- **Eligibility≠Utilization**

# Types of Medicaid Claims

---

- **Fee-for-service data – reports cost of services provided**
- **Capitation data – reports monthly fee paid by states for enrollment in managed care**
- **Encounter data – reports no Medicaid cost but reports utilization of managed care**

---

# Analytical Issues

# Linking to Multiple MAX Records – Part 1

- **NHANES record matches to same Medicaid enrollee in multiple MAX years**
  - Medicaid enrollees are likely to be enrolled in Medicaid for more than one year
  - Records that match to same NHANES participant will be highly correlated
  - Adjust analysis plan
  - Suggestion: Only keep records that match in same year that NHANES data collected

**(Note: NHANES survey year is a restricted variable)**

# Linking to Multiple MAX Records – Part 2

- **NHANES record matches to same Medicaid enrollee within the same MAX year**
  - **Enrolled in Medicaid in more than one state**
    - Likely okay to leave as is
  - **Data issue**
    - Assigned more than one Medicaid ID within the same state
    - May consider combining records

# MAX Classifications That Affect Cost Data

- Including certain enrollees may underestimate average costs:
  - Duals
  - Enrollees with restricted benefits
- S-CHIP only enrollees can muddy cost data
- Managed care enrollees
  - Effect on costs unclear
  - May weaken associations with cost

# Dual Status – Part 1

- **Dual Status – Enrollees who qualify for both Medicaid and Medicare benefits**
- **Medicare first payer for:**
  - Inpatient services
  - Outpatient and physician services
  - Some home health
  - Prescription Drugs (starting in 2006 with Part D)
  - Limited Skilled Nursing Facility (SNF) services
- **Medicaid pays for:**
  - Medicare premiums and cost-sharing
  - Long-term care
  - Optional services: dental, hearing, vision, home- and community-based services

# Dual Status – Part 2

- **Dual enrollees likely to have lower Medicaid costs than non-duals due to cost-sharing**
- **39% disabled enrollees are duals (2008)**
- **92% aged enrollees are duals (2008)**
- **Suggestion: Remove duals from cost analyses**
- **Note: If remove duals, will be difficult to do analysis of ages 65+**

Source: Kaiser Commission on Medicaid and the Uninsured and Urban Institute estimates based on data from FY 2008 MSIS, 2012.

# Restricted Benefit Status

- **Restricted-Benefit – Medicaid enrollees eligible for only limited coverage**
- **Three main categories:**
  - Aliens eligible only for emergency services
  - Duals
  - Enrollees in 1115 waivers that provide only family planning services
- **11% enrollees have restricted benefits (2008)**
- **Suggestion: Remove restricted-benefit enrollees from cost analyses because likely to have lower costs than full-benefit enrollees**

Source: Borck et al. "The Medicaid Analytic Extract 2008 Chartbook." 2012.

# S-CHIP Only

---

- **States cover children in Separate Children's Health Insurance Program (S-CHIP) or CHIP**
- **CHIP data included in MSIS/MAX**
- **States not required to submit S-CHIP data but some do**
- **Suggestion: Exclude S-CHIP only enrollees because enrollment and services are not fully reported**

# Types of Managed Care – Part 1

- **Primary Care Case Management (PCCM)**
  - Capitation data for case management only (typically small fees)
  - Cost for services captured in FFS data
  - Vast majority of costs in FFS data
  - Suggestion: Keep PCCM enrollees in analysis
- **Comprehensive Managed Care (HMO/HIO/PACE) (CMC)**
  - Most care provided through program
  - Costs captured in capitation data (not reflective of service use)

# Types of Managed Care – Part 2

- **CMC cont'd**
  - **Possible for CMC enrollees to have FFS claims**
    - Partial-year enrollment
    - Carve out services
  - **Majority of costs in capitation data**
  - **Not clear whether costs lower than FFS enrollees**
  - **Difficult to detect cost association because capitation payments do not vary by service use**
  - **Suggestion: Remove CMC enrollees from analysis if looking for association**

# Types of Managed Care – Part 3

- **Prepaid Health Plan (PHP)**
  - **Specific Services**
    - Behavioral Health
    - Dental
    - Long-term Care
    - Other
  - **Costs captured in capitation data (not reflective of service use)**
  - **Used in conjunction with FFS or CMC**
  - **In some states, all enrollees in PHP**
  - **Typically only covers carve out services**
  - **Suggestion: Leave PHP enrollees in analysis**

# Defining Eligibility/Enrollment

- **Analysis plan needs to define eligibility/enrollment for study population:**
  - **Last month status**
    - Typically used for BOE category
  - **Status occurred for all months in year**
    - Yields smallest but cleanest data set (rarely used)
  - **Status occurred for all months eligible**
    - Very restrictive for CMC enrollment
  - **Status occurred for at least one month**
    - Typically used for CMC enrollment

# Medicaid Data are State Data

---

- **Variation inherent in state-submitted data**
  - Programs differ
  - Classification of programs may differ
    - PHP vs CMC
  - Quality of data
    - Maine 2005-present: only enrollment and RX data
- **Most Medicaid research is state-based but NCHS hesitant to reveal respondent's state even in RDC**

# Controlling for State Variation

---

- **MAX variables can control for some differences but not all**
- **Suggestion: Use masked state variable to control for state differences without actually knowing the state**
- **Issue: Cannot interpret data from MAX-NCHS merged data sets by state**

# MAX Cost Analyses

- **Adjust costs to same year**
- **Analytical decisions**
  - **Think about whether total costs are being affected by/masking important information**
    - Long-term care population
    - Particular service categories
  - **Annual costs or per member per month (PMPM)?**
  - **All enrollees or only service recipients?**

# Obesity Analysis Example

- **Analysis of enrollees 20+ years**
  - 3,227 records - NHANES 1999-2004 and feasibility file
  - 9,809 records matched in NHANES and MAX 1999-2004
  - Of those, 3,847 records matched with MAX year in NHANES data release years
  - Of those, 1,061 records matched who were full-benefit, non-duals who were not in S-CHIP only and never in CMC
    - Restricted Benefits – 439 (11.4% of 3,847 records)
    - Duals – 1,818 (47.3%)
    - S-CHIP Only – 6 (0.2%)
    - CMC enrollees – 1,128 (29.3%)

# Obesity Analysis Example cont'd

- **Analysis cont'd**
  - After implementing MAX cost restrictions, only 50 elderly records (ages 65+)
  - Final sample size of enrollees ages 20-64 years: 603 records
    - Implemented NHANES restrictions
      - Not pregnant
      - Measured BMI data available
    - Limited to 1 match within NHANES data release years (used survey year) (n=329)
    - Resolved duplicates (n=4)

# Conclusions

- **NCHS feasibility files useful but can overestimate sample size**
  - Feasibility file: 3,227 records
  - Final sample: 603 records
    - Restricted matches to same year
    - Implemented MAX cost restrictions
    - Removed elderly
    - Resolved duplicates
- **MAX/NHANES cost analyses may be limited**
- **MAX/NHANES enrollment analyses would not be affected as greatly due to fewer MAX restrictions**

# Questions?

---

- **Allison Hedley Dodd**
  - [adodd@mathematica-mpr.com](mailto:adodd@mathematica-mpr.com)
- **MAX data**
  - <http://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MedicaidDataSourcesGenInfo/MAXGeneralInformation.html>