# Multi-Payer Claims Database (MPCD) for Comparative Effectiveness Research

Andre Chappel, PhD

June 16, 2011

NCVHS Full Committee Meeting



# Catalyst

- \$1.1 billion investment in CER enterprise funded by American Recovery and Reinvestment Act of 2009
- \$400 million allocated to the Office of the Secretary (OS)
- Federal Coordinating Council (FCC) for CER advised Secretary on how to invest allocated funds
- Due to immediate need, FCC recommended OS focus on CER infrastructure development
- HHS identified value of creating a database that would combine claims data from a range of public and private payers

# Objectives

- Goal: Build a MPCD to support CER, on a foundation of public and private payer claims data
  - High priority: Privacy and protection of patients
  - Build a comprehensive database to enable research on priority populations, interventions, and conditions
  - CER broadly defined to include both clinical & delivery system research
  - Increase access and usability of the data
  - Engage private sector in CER infrastructure development and research
  - Offer analytic tools for greater functionality
  - Lay the foundation for future enhancements with clinical data



# Value-Added

- Incorporating public and private data into one source will enhance the value of claims data for CER
  - Easier to find and obtain relevant data
  - Greater geographic coverage and density
  - Ability to study less common conditions
  - Increased demographic and clinical representativeness
  - Potential linking across payers and time

# **Project Flow**

- Phase 1: Strategic and technical design
  - Strategic plan completed April 29, 2010 by Avalere Health
  - Technical design pilot testing nearing completion
    - Tools to combine data from different sources Vexcel/Microsoft
    - User interface to create customized data extracts Thomson Reuters
- Phase 2: Implementation of MPCD Awarded to Ingenix,
   Start date: January 2011

### Phase 2 Overview

- Collaborative effort between ASPE, CMS, Ingenix and other healthcare industry and research leaders
- Database components
  - Hybrid design
    - Centralized warehouse
    - Distributed/federated data network
  - CMS Chronic Condition Warehouse (CCW)
  - Ingenix Normative Health Information Database (NHI)
  - Two additional data sources.



# **Advisory Structure**

#### Governance Board

- Members recruited by AcademyHealth
- Consists of researchers, private payers, providers, consumers, states, and the Federal Government
- Guide MPCD's implementation and future sustainability

#### HHS Leadership Council

- Policy leaders and data experts across multiple agencies within HHS
- Gain insight into potential implementation issues
- Facilitate collaboration with similar initiatives



### **Governance Board**

**Voting Members** 

W. David Helms, PhD (Chair) Alan Garber, MD, PhD

Jeff Allen, PhD Mark Gaskill, MFT

Patrick Baier, DPhil Bruce Hamory, MD

Melinda Beeuwkes Buntin, PhD Craig Jones, MD

Bob Berenson, MD Patrick Miller, MPH

Rachel Behrman, MD James Rohack, MD

Ken Buetow, PhD Lucia Savage, JD

Robert Dubois, MD, PhD Jim Scanlon

Irene Fraser, PhD Tony Trenkle

**Non-Voting Members** 

Denise Hynes, PhD, MPH, RN Amol Navathe, MD, PhD

Larry Kessler, ScD Steve Rokes, MIS

Denise Love, MBA, RN Lisa Simpson, MD, BCh, MPH

Mark McClellan, MD, PhD, MPA Lu Zawistowich, ScD

# **HHS Leadership Council**

Sherry Glied, PhD (Chair)	ASPE
Anne Elixhauser, PhD	AHRQ
Charles Friedman, PhD	ONC
Rick Gilfillan, MD MBA	CMS
Rick Kronick, PhD	ASPE
Farzad Mostashari, MD ScM	ONC
Todd Park	IOS
Tom Reilly, PhD	CMS
Melissa Robb	FDA
Jean Slutsky, PA MSPH	AHRQ

# **Advisory Structure**

- Data Stewardship Council (DSC)
  - Will include all data contributors
  - Provide recommendations on technical design and data access
  - Monitor processing of data in MPCD
  - Review and approve data requests
- HHS Expert Panel
  - Researchers within HHS familiar with claims data and CER
  - Utilize and test the system beginning in February 2012
  - Provide feedback on how the system can be improved



# Privacy and Security

- Emphasis on privacy and security controls
  - DUAs to address security concerns of data partners
  - Encryption of PII for matching across partners, de-identified data
  - Compliance with FISMA, HIPPA, ARRA, and CMS IT policies
  - Masking of small cells and dates of service
  - Statistician review of de-identification
  - Audit reports for contributors to track data processing
  - Testing data enclave models



# Technical Design

- Leverage insights from CMS' CER project on de-identification
- Distributed vs. centralized datasets
  - Advances in database structure, high-speed computing, and health IT are enabling ways to link across data sources
- Distinction from other federated data network designs such as FDA Sentinel:
  - Analytic queries vs. queries for data extracts
    - Health services researchers have expressed a need to work with data
    - Statistical power



# **Data Access Model**

- Web-based interface
- Three tiers of data access
  - Tier 1: Public usage files (PUFs)
    - Aggregate data
  - Tier 2: Extracts from standard analytic files (SAFs)
    - Limited data set
    - Requests subject to approval of DSC
  - Tier 3: Extracts from full claims files
    - Obtain richer data than what is available in tier 2
    - Request subject to approval of DSC and relevant data contributors



#### **Data Sources**

- CMS CCW
  - Medicare FFS and Medicaid
- Ingenix NHI
  - Geographically and demographically distributed commercial claims
- Two additional sources in Stage 1
  - States with all-payer claims databases (APCDs)
  - Commercial data contributors
    - Health plans, data aggregators



#### **Data Sources**

- Additional sources beyond Stage 1
  - Pursue additional commercial and public payer data sources
  - Conduct periodic data analysis to determine gaps in coverage



### Data Partner Evaluation Criteria

- Restrictions on use of data
- Quantity and coverage of data
- Quality and format of data
- Cost to MPCD for their participation
- Overlap with other partners

### **Incentives**

- Subsidize cost of participation
  - Provide hardware for distributed partners
  - Cover cost of work to prep data
  - Royalties/revenue sharing
  - Limited number of free data extracts
    - Subject to standard approval process
  - Market awareness of contributor data
  - Representation on DSC
  - Public recognition



# **Key Partners**

- AcademyHealth
  - Coordinating Governance Board
  - Supporting outreach to the research community
- National Association of Health Data Organizations (NAHDO)
  - Facilitating state data acquisition
- Buccaneer Computer Systems and Services
  - Hosting MPCD platform
  - CCW data extraction and integration
- University of Washington
  - Participating in database design to ensure value for CER research
  - Will conduct demonstration of CER power and usability of MPCD



### **Outreach Activities**

- Build awareness and demonstrate MPCD utility through:
  - Presentations at selected annual research conferences (e.g., AcademyHealth ARM, NAHDO)
  - Presentations to other agencies (e.g. NCVHS meeting)
  - Publish articles on demonstration analyses and policy applications in selected peer reviewed journals

# **Future Plans**

- Include value-added analytics
  - Episode groupers
  - Risk adjustment
  - Quality metrics
- Incorporate non-claims data
  - Lab results
  - Clinical records
  - Electronic medical records



#### **ASPE Project Team**

Andre Chappel, PhD
Kate Goodrich, MD MHS
David Knutson, MS

Andre.Chappel@hhs.gov 202-205-4064

www.aspe.hhs.gov

