

MCDB Data Release Policy Workgroup

February 25, 2014



Overview of Workgroup Goals

- Defining data products
 - Stakeholder concerns
 - Sensitive files
 - Protected Health Information (PHI)
 - Proprietary information
 - Privacy level of data products
- Transparency of the review process
 - Standardized review process for scientific merit, qualifications of applicants, data management plans, and data use agreements
 - IRB

Stakeholder Concerns

- Consumers
 - Privacy and general concerns about release of data
 - PHI and HIPAA
- Payors
 - Release of negotiated rates and potential for anticompetitive activities
 - Identification of payor
- Providers, State partners, and users of data
 - Easy-to-use and understandable data products
 - Transparency of application review process

Strategies for Protection of PHI

- Strategies depend on the needs of analysis and the level of privacy protection desired:
 - Most secure / least information: Do not release any PHI (e.g. no individual-level identifiable information including demographics, diagnosis, payment information, etc.). Data products provide summaries.
 - Modestly secure / substantial information: Data products have individual-level information, but key fields are scrambled or limited (e.g. scrambled patient ID, limit geographic information, age bands, etc.)
 - Less secure / detailed information: Data products provide detailed, individual-level information.
- Transparent and systematic decisions regarding the types of data products available and the process used to review applications are essential.

Strategies for Protection of Proprietary Information

- Strategies depend on the needs of analysis and the level of protection desired. Approaches similar to protection of PHI may be used.
- Most research questions do not require specific provider or negotiated rate information. NPI and payor identification may be scrambled or removed.
- Proprietary information, such as add-on metrics and groupers, may only be released if licenses permit it

Differences between CMS Research Identifiable Files and Limited Data Sets

Variable	File	Research Identifiable File	Limited Data Set
Beneficiary identifier	Claims & Enrollment files	Encrypted identifier	Encrypted identifier
Health Insurance Claim (HIC) or Social Security Number (SSN)	Claims & Enrollment files	Not included in file	Not included in file
Dates (MM/DD/YYYY)	Claims files	Included	Included as of CY2010
Beneficiary Zip Code	Enrollment file	Included	Not included, only County or State
Beneficiary Date of Birth	Enrollment file	Included	Not included, either Age year or 5-yr age range
Date of Death	Enrollment file	Included	Included, only for validated dates of death
UPIN/NPI of performing provider	Claims files	Included	Encrypted
Institutional provider number	Claims files	Included	Included

Source: <http://www.resdac.org/resconnect/articles/148>

Environmental Scan of Claims Data Products Released by Other APCD States

DRP Elements	Colorado	Maine	Massachusetts	New Hampshire	Oregon	Vermont
<i>Privacy Levels</i>						
Research Identifiable Files		X	X	X		X
Limited Data Sets	X	X	X	X	X	
Public Use Files				X	X	
Summary Reports	X		X			

Proposed Field Inclusion/Exclusion by File Privacy Level for MCDB Claims Files

Variable	Research Identifiable File*	Limited Data Set*	Public Use File*
Beneficiary identifier	Encrypted identifier	Encrypted identifier	Not included
Dates (MM/DD/YYYY)	Included	Included	Not included
Beneficiary Zip Code	Included	3-digit Zip, County or State	Zip, County or State (summary counts)
Beneficiary Date of Birth	Included	Not included, 5-yr age range	Not included, 5-yr age range
Diagnosis	Included	Included, unless rare	Select diagnoses included (summary counts)
Procedures	Included	Included, unless rare	Select procedures included (summary counts)
NPI	Included	Encrypted	Not included
Encounter Price	Included	Included	Not included, summary costs
Payor	Included	Not included	Not included
Product Type (e.g. HMO vs. PPO)	Included	Included	Included (summary counts)
Coverage Type (e.g. Individual vs. Small Group)	Included	Included	Included (summary counts)
Proprietary Information (e.g. MULTUM)	Not included	Not included	Not included

Research Identifiable File and Limited Data Set are person-level claims files. The Public Use File is an aggregate file, summarized by geographic location, usually zip code, county, or State.