



*Agency for Healthcare Research and Quality*

*Advancing Excellence in Health Care • [www.ahrq.gov](http://www.ahrq.gov)*

## **Integrated Health Care Survey Designs: Analytical Enhancements Achieved Through Linkage of Surveys and Administrative Data**

---

National Committee  
on Vital and Health Statistics  
September 18, 2006  
**Steven B. Cohen, Ph.D.**

# Purpose of Discussion

- Integrated survey design features
- Related enhancements to data quality and analytical capacity
- Capacity to reduce bias attributable to survey nonresponse
- Applications to AHRQ Data Portfolio and Research Initiatives to inform **health outcomes**
- Limitations
- Future model for consideration

# Health Outcomes Focus Part of AHRQ's Mission

*AHRQ Mission: To improve the  
quality, safety, efficiency, and  
effectiveness of health care for all  
Americans*

# Integrated survey design features

- Direct linkage between sample members in core survey with larger host survey; administrative records; or follow-up surveys
- Use of secondary data (e.g. aggregate data at the county/state level) as core component of survey
- Prior survey record of call data informs data collection strategies
- Informs sample design, nonresponse and poststratification adjustments, imputation and data supplement for item nonresponse
- Need for greater attention to ensuring confidentiality: limitations in public use data

# Capacity to reduce bias attributable to survey nonresponse

## Adjustments for unit nonresponse

- Detailed information available on demographic/socio-economic characteristics of both respondents/and nonrespondents from sample frame of host survey administrative records
- Incorporation of secondary data

## Adjustments for item nonresponse

- Data replacement
- Cold deck imputation



# MEPS Medical Expenditure Panel Survey

Search

## About MEPS

- [What is MEPS?](#)
- [Data & Statistics](#)
- [Publications](#)
- [MEPSNet Query Tool](#)
- [Survey Instruments](#)
- [What's New](#)
- [Data Workshop](#)
- [Data Center](#)
- [Mailing List](#)
- [Participant's Corner](#)
- [Contact Us](#)

This new look is a preview of our redesigned site, which will be launched in a few months. With the exception of this new home page, nothing else has changed on the current site.

## Welcome to MEPS

The Medical Expenditure Panel Survey (MEPS) is a set of large-scale surveys of families and individuals, their medical providers, and employers across the United States. This Web site lets you access the results of these surveys. Here you will find detailed information on health care utilization and expenditures, health insurance, and health status, as well as a variety of demographic, social, and economic characteristics of a representative sample of Americans. You can review and download public use microdata files, access MEPS data presented in a tabular format, analyze MEPS data using online tools, and read and download a variety of analytic publications.

## What's New

### New data files

- [MEPS HC-080](#): MEPS Panel 7 Longitudinal Weight File
- [MEPS HC-079](#): 2003 Full Year Consolidated Data File

### New publications

- [Research Findings #24](#): Utilization and Expenditures for Children with Special Health Care Needs
- [Statistical Brief #113](#): Children's Dental Care: Periodicity of Checkups and Access to Care, 2003
- [Statistical Brief #112](#): Dental Expenditures in the 10 Largest States, 2003
- [Statistical Brief #111](#): Expenses for a Hospital Emergency Room Visit, 2003
- [Statistical Brief #110](#): The Top Five Therapeutic Classes of Outpatient Prescription Drugs Ranked by Total Expenses for Adults Age 18 and Older in the U.S. Civilian Noninstitutionalized Population, 2003

# Medical Expenditure Panel Survey (MEPS)

## Annual Survey of 15,000 households:

provides national estimates of health care use, expenditures, insurance coverage, sources of payment, access to care and health care quality

## Permits studies of:

- Distribution of expenditures and sources of payment
- Role of demographics, family structure, insurance
- Measurement of expenditures in managed care
- Expenditures for specific conditions
- Trends over time

# Key Features of MEPS-HC

- Survey of U.S. civilian noninstitutionalized population
- Sub-sample of respondents to the National Health Interview Survey (NHIS)
- Oversample of minorities and other target groups
- Panel Survey – new panel introduced each year
  - Continuous data collection over 2 ½ year period
  - 5 in-person interviews (CAPI)
  - Data from 1st year of new panel combined with data from 2nd year of previous panel

# MEPS Household Component Sample Design

## Oversampling of policy relevant domains

1996	Minorities (Blacks & Hispanics)
1997	Minorities Low income Children with activity limitations Adults with functional limitations Predicted high expenditure cases Elderly
1998-2001	Minorities
2002+	Minorities, Asians, Low Income

# MEPS Components

- **Household Component (HC)** - 15, 000 households, 37, 000 individuals
- **Medical Provider Component (MPC)** - designed to supplement /replace household reported expenditure data
- **Insurance Component (IC)** - 30,000 establishments; elicits insurance availability, premium contribution, and benefit provision information; can be used to generate estimates at the state level

IC sample linked to HC designed to supplement or replace household reported health coverage data

# MEPS - Integrated Survey Design Features

- National Health Interview Survey serves as sample frame for HC
- Census Bureau Business Register serves as IC sample frame
- Secondary data on health care measures supplement surveys  
Linked survey of medical providers
- Linked survey of employers
- Distinct data sources linked for longitudinal analyses

# Detailed Information for Unit Nonresponse: NHIS to MEPS

Factors used in nonresponse adjustments

- Size of dwelling unit
- Family income
- Employment classification
- MSA classification
- Dwelling unit level personal help measure
- Phone number availability

# Adjustment factor

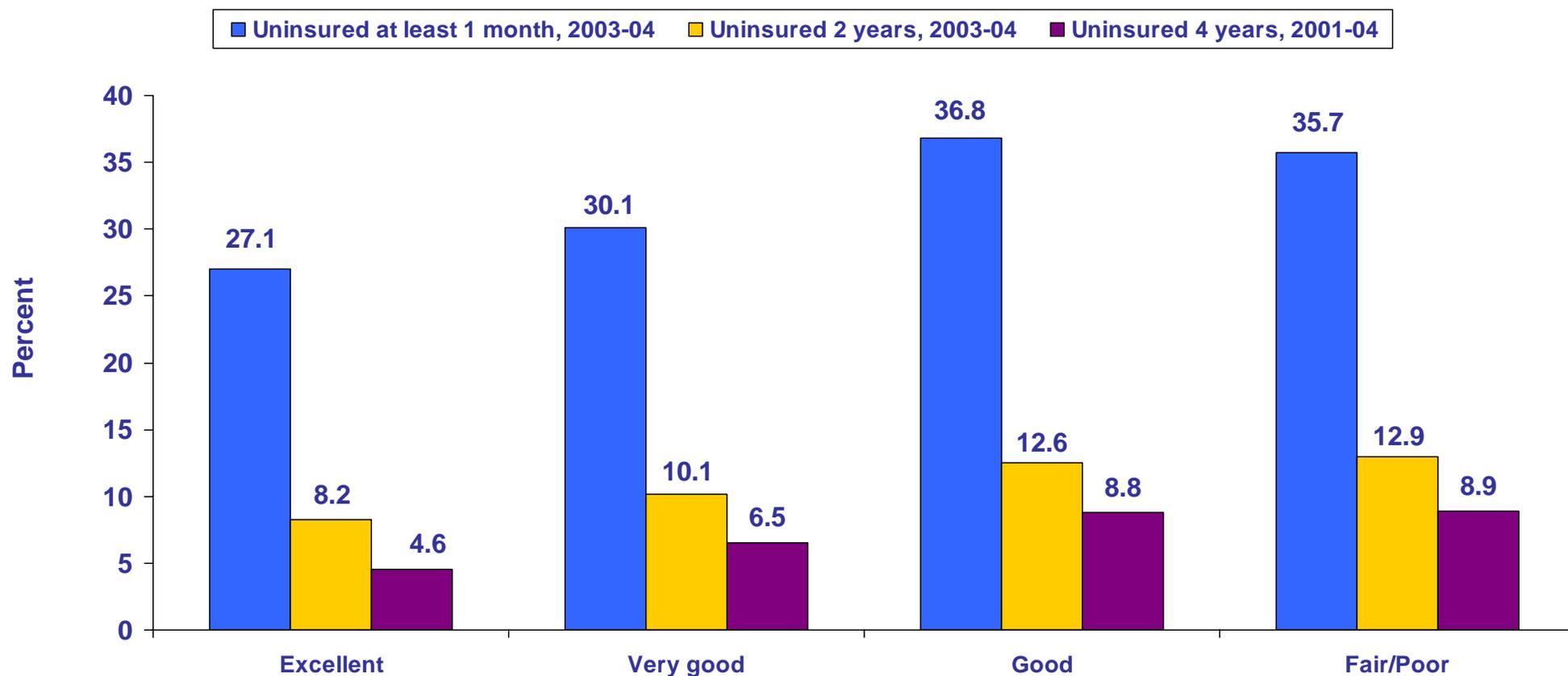
- Within each adjustment cell:

$$A(c) = \frac{\sum_{i \in c} E(i)W_1(i)}{\sum_{i \in c} R(i)W_1(i)}$$

ratio of the sum of weights of all eligible (E) units in the cell to the sum of weights of only the respondents (R) in the cell



## Percentage uninsured by health status, U.S. civilian noninstitutionalized population under age 65, 2001–2004



Source: Center for Financing, Access, and Cost Trends, AHRQ, Household Component of the Medical Expenditure Panel Survey, HC-079 (2003), HC-082 (2004), and HC-086 (Panel 8)

# Medical Provider Component

## Purpose

- Compensate for household item nonresponse
- Gold standard for expenditure estimates
- Greater accuracy and detail
- Imputation source
- Supports methodological studies

# Medical Provider Component

## Targeted Sample

- All associated hospitals and associated physicians
- All associated office-based physicians
- All associated home health agencies
- All associated pharmacies

## Data Collected

- Dates of visit
- Diagnosis and procedure codes
- Charges and payments

# MPC: Correction Source for Item Nonresponse

## Source for event level expenditures

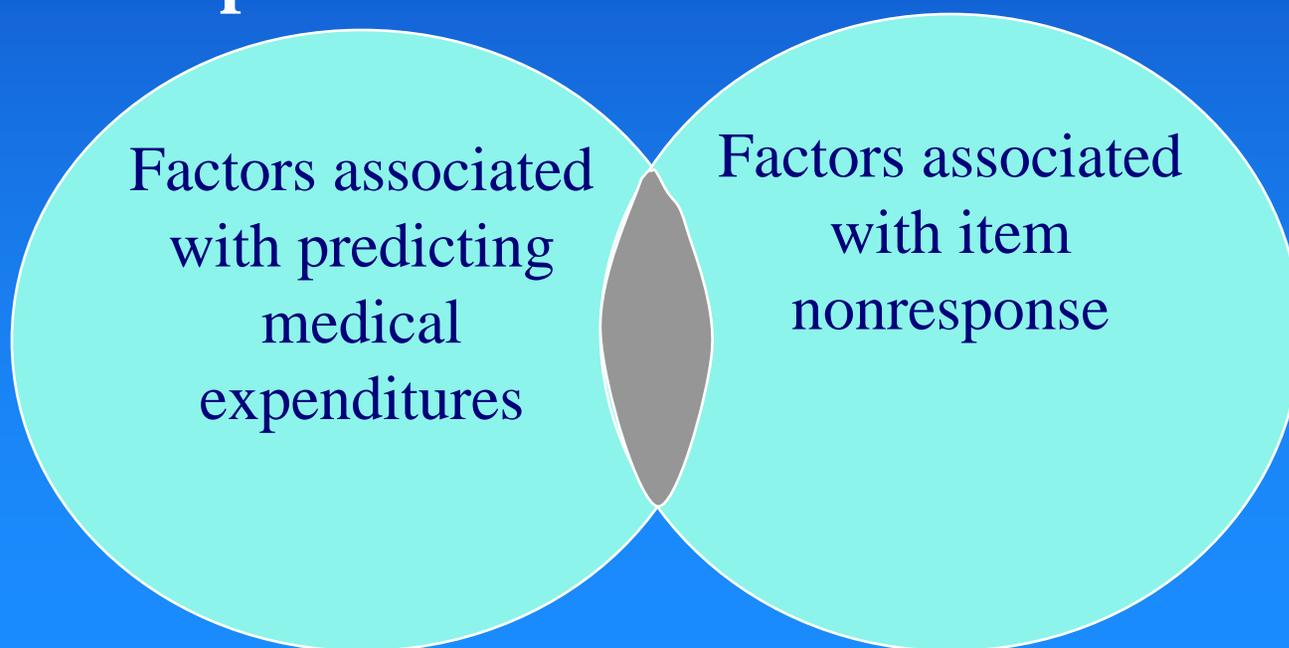
<i>Household</i>	<i>Provider</i>	<i>MEPS value - <math>Y_{ij}</math></i>
<i>Reported</i>	<i>reported</i>	$Y_{ij} = \text{Provider } \$_{ij}$
<i>Nonresponse</i>	<i>reported</i>	$Y_{ij} = \text{Provider } \$_{ij}$
<i>reported</i>	<i>nonresponse</i>	$Y_{ij} = \text{Household } \$_{ij}^1$
<i>nonresponse</i>	<i>nonresponse</i>	$Y_{ij} = \text{Imputed } \$_{ij}$

<sup>1</sup>Recalibrated as necessary based on analyses of concordance between sources

# Determination of Factors for Expenditure Imputation

**Hot Deck Imputation:**

**Classification Variables for Donors and Recipients**



# Collection of Rx Data in MEPS

- MEPS: annual nationally representative household survey of the civilian, non-institutionalized population
  - 15,000 households, 40,000 individuals, 195 Primary Sampling Units (e.g. MSAs, counties)
- For each person in household, for each prescription drug purchase, collected each round of survey:
  - Medication name
  - Number of purchases
  - Condition for which prescribed
- Pharmacy follow-back survey with signed permission form:
  - Verifies household purchases

# Collection of Rx Data in MEPS

- ~8,000 pharmacies sampled annually
  - data on prescribed medicines purchased by households
- Data obtained:
  - Medication Name
  - National Drug Code (NDC)
  - Quantity Dispensed
  - Strength and Form
  - Sources of Payment
  - Amount Paid by Each Source

**SAMPLE PRESCRIPTION INFORMATION LIST**  
**MEDICAL EXPENDITURE PANEL SURVEY - U.S. PUBLIC HEALTH SERVICE**

SANDIE M KING  
 DOB: 07/10/65 • FEMALE  
 42816659-666123B5253

WALGREEN'S PHARMACY  
 63335109

*For the patient listed to the left, please provide information about each prescription filled or refilled between January 1, 1996 and December 31, 1996.*

<p><b>1.</b> Date Filled: <u>03/26/96</u>                  NDC: <u>00003-0109-60</u>                  Generic/Trade Name(s): <u>Amoxicillin (BMS)</u>                  Strength: <u>500</u> unit: <u>mg</u>                  Quantity (Pkg. Size/Amt. Dispensed): <u>21</u>                  Total Charge: \$ <u>7.00</u></p>	<p><b>Payments:</b>                  Patient: \$ <u>6.00</u> Other Federal: \$ _____                  Private Ins.: \$ <u>1.00</u> Other State: \$ _____                  Medicaid: \$ _____ Workers' Comp: \$ _____                  VA: \$ _____ Other: \$ _____                  CHAMPUS/CHAMPVA: \$ _____</p>
<p><b>2.</b> Date Filled: <u>03/24/96</u>                  NDC: <u>00085-0647-03</u>                  Generic/Trade Name(s): <u>Intro - A Inj</u>                  Strength: <u>3</u> unit: <u>mmu/vial</u>                  Quantity (Pkg. Size/Amt. Dispensed): <u>12</u>                  Total Charge: \$ <u>370.09</u></p>	<p><b>Payments:</b>                  Patient: \$ <u>18.00</u> Other Federal: \$ _____                  Private Ins.: \$ <u>352.09</u> Other State: \$ _____                  Medicaid: \$ _____ Workers' Comp: \$ _____                  VA: \$ _____ Other: \$ _____                  CHAMPUS/CHAMPVA: \$ _____</p>
<p><b>3.</b> Date Filled: <u>   </u> / <u>   </u> / <u>   </u>                  NDC: _____                  Generic/Trade Name(s): _____                  Strength: _____ unit: _____                  Quantity (Pkg. Size/Amt. Dispensed): _____                  Total Charge: \$ _____</p>	<p><b>Payments:</b>                  Patient: \$ _____ Other Federal: \$ _____                  Private Ins.: \$ _____ Other State: \$ _____                  Medicaid: \$ _____ Workers' Comp: \$ _____                  VA: \$ _____ Other: \$ _____                  CHAMPUS/CHAMPVA: \$ _____</p>

<sup>1</sup> *Total Charge:* Record the dollar amount of the total charge. This is the price of the prescription: the cash price if it is a cash transaction, or the contract price if it is a third party transaction. The prescription price is calculated based on the specific prescription and/or the specific payor's contracted reimbursement rate.  
<sup>2</sup> *Patient Payment:* Record the total dollar amount of the payment made by the patient. Include any payments made to meet the patient's copayment, coinsurance, or deductible.  
<sup>3</sup> *Private Insurance Payment:* Record the total dollar amount of all payments, if any, made by private health insurance sources. Private insurance sources do include insurance that is paid by an employer or an individual. Do not include payments made by public insurance sources, such as Medicaid, Medicare, VA, etc. Those should be included under "Medicaid," "Veterans' Administration," "CHAMPUS/ CHAMPVA," "Other Federal," "Other State," "Workers' Compensation," or "Other," as appropriate.



# MEPS is enhanced by links to secondary data sources

- Proprietary drug classification database links to MEPS by NDC (Cerner Multum, Inc)
  - Therapeutic class, subclass
    - For example, cardiovascular drugs (class), beta-blockers (subclass), now easily identified
  - Pregnancy safety category
- Available to public in October 2004
- Other secondary data sets can be linked by NDC
  - Year of approval by FDA; FDA priority code
  - Brand / generic indicator

# Types of Analyses Supported by MEPS Prescribed Medicine Data

- Trends in out of pocket burdens across all major population subgroups
- Prevalence of potentially inappropriate prescribing patterns
- Trends in use and expenditures by therapeutic category: e.g. statins, anti-depressants, analgesics, proton pump inhibitors
- Prediction models of future year's expenditures

# MEPS Validation Studies

- MEPS data on health care utilization based on household reports:
  - Benchmark comparisons of MEPS to National Health Expenditure Accounts (NHEA) and administrative data raise concerns about possible underreporting
- *Goals of Validation Studies:*
  - Better understand differences between survey and administrative data
  - Improve MEPS data collection and editing/imputation procedures
  - Improve alignment between MEPS and the NHEA and administrative data for use in policy development and analysis (e.g. CBO/CRS, OMB...)

# MEPS-Medicare Validation Study Areas

- Validate provider reporting of payment data
  - Compare MPC records to Medicare claims
- Validate household reporting of utilization
  - Compare MEPS household reports to Medicare claims by type of service (inpatient, ED, outpatient department, medical provider visits)
- Validate collection of data on separately billing doctors for hospital-based events.
- Examine completeness of use and expenditure data for MEPS respondents who go into nursing homes or are lost to death and attrition.

# IC - Purpose

- Availability of health insurance
- Access to health insurance
- Cost of health insurance
- Benefit and payment provisions of private health insurance

# IC - Sample

- 30,000 establishments: derived from Census Bureau frame
- Supports national and state estimates
- Employers linked to HC sample
- Data released in tabular form on MEPS website

# Key Administrative Data Available for MEPS Insurance Component

1. Industry
2. Payroll
3. Age of Firm
4. Establishment Size
5. Enterprise Size
6. Location
7. Multi/Single Unit Firm
8. Form of Organization

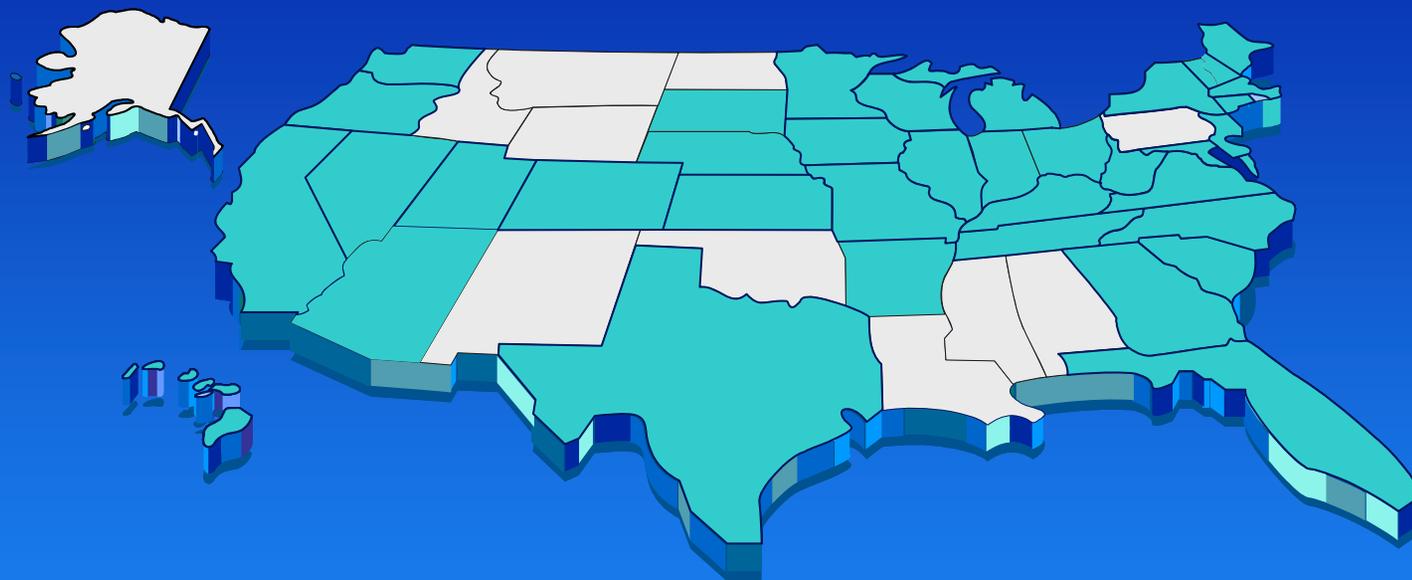
# Uses of Administrative Data in the Insurance Component

1. Sampling
2. Imputation
3. Editing
4. Modeling
5. Table Production
6. Weight Adjustment for Non Response and Control Totals

## Key Improvements Due to Use of Administrative Data in the Insurance Component

1. Reduces Respondent Burden
2. Improves Sampling Precision
3. Helps Find Respondent Errors
4. Improves Weight Adjustment
5. Allows Estimates Be Made for Numerous Key Categories
6. Essential for Modeling and Research

# The Power of Administrative Data: Healthcare Cost & Utilization Project (HCUP)



37 State Partners

90% of All-Payer Hospital Discharges in U.S.

Census of Hospital Care—NOT a Sample

Inpatient, ambulatory surgery, ED databases



# The Making of HCUP Data



*Patient enters hospital*

*Billing record created*



*Hospital sends billing data and any additional data elements to Data Organizations*

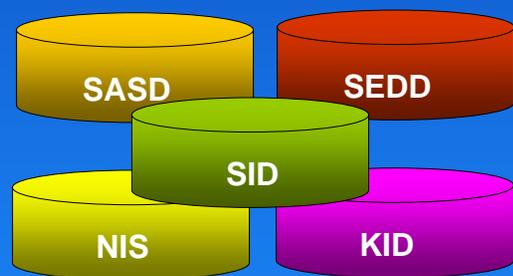


*AHRQ standardizes data to create uniform HCUP databases*

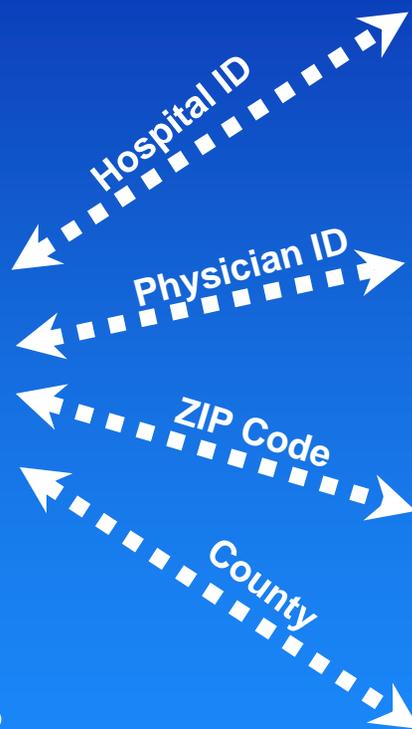


*States store data in varying formats*

# Standard Links to Other Databases



**HCUP Databases**



AHA Annual Survey  
Medicare Cost Reports

State License Files  
CMS Provider File

Census & Vendor Files

HRSA Area Resource File



# Encounter Data Made More Powerful Through Linkages

Linkage Variable	Information Added
Hospital ID	Structural characteristics, e.g. bedsize, ownership Financial status, cost-to-charge ratios
Physician ID	Physician specialty Office location
County (patient or hospital)	Population demographics Health resources
Zipcode (patient or hospital)	Community SES, demographics Data for travel distance calculations

# Encrypted Patient Identifiers in HCUP

HCUP receives encrypted IDs (privacy protection)

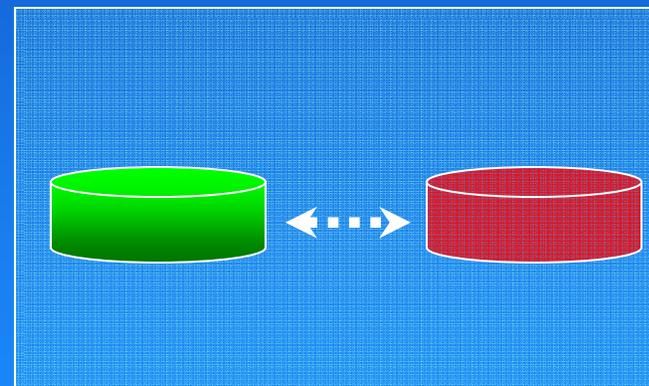
Records for the **same patient** can be linked in some states' HCUP SID, SEDD, SASD for measuring readmissions, revisits, episodes of care



WITHIN  
INSTITUTIONS



ACROSS  
INSTITUTIONS



ACROSS  
DATABASES

# Other Patient Links with Statewide Encounter Data

<b>Linked Database</b>	<b>Information Added</b>
<b>Vital records</b>	Birthweight Date of death
<b>Disease registry</b>	Detailed information about patient's disease, e.g. cancer
<b>Ambulance records, police reports</b>	Information on motor vehicle accidents
<b>State program files</b>	Enrollment status in program, e.g. Medicaid Program service use



# Linkage Challenges

Type of Link	Challenges
Patient ID	<p>State sensitivity about collection &amp; release</p> <p>ID method not uniform across states</p> <p>Encryption methods change over time</p> <p>Data cleaning needed to identify ID errors</p>
Hospital & Physician ID	<p>Some states sensitive about release</p> <p>ID not uniform across databases</p> <p>Hospital mergers and closures over time</p> <p>Physician ID may be shared across practice</p>
Patient zipcode	<p>Some states sensitive about release</p> <p>Zipcode boundaries change over time</p> <p>Codes for foreign or homeless pts not uniform</p>

## ***Studies Using Linked HCUP or State Encounter Data***

- Racial/ethnic disparities in readmissions for diabetes
- Incidence/cost of motorcycle injury to inform decisions on state helmet laws
- Multiple ED visits and ED visit-admission patterns for substance abuse patients
- Financial status of safety net hospitals
- Impact of motor vehicle exhaust on pediatric asthma admissions

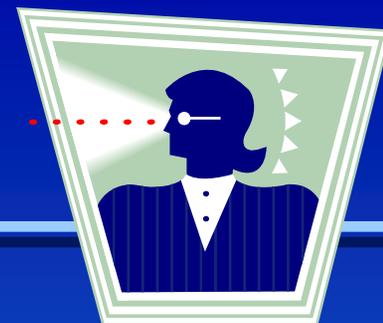
## **Links for the Future**

<b>Type of link</b>	<b>Information Added</b>
<b>Claims and EMR within hospital</b>	<b>Detailed clinical data for better quality measurement</b>
<b>Patient links</b>	<b>User-friendly links, expand states with links, variables for readmissions and revisits</b>
<b>Hospital organization information</b>	<b>Organizational culture, clinical integration, governance, HIT</b>
<b>HospitalCompare</b>	<b>Quality measures on process of care HCAHPS- consumer assessment</b>
<b>Nurse staffing files</b>	<b>Number and type of nurses at hospital</b>
<b>Market area files</b>	<b>Measures of hospital competition</b>

# ***Vision for Future: Joining Forces***

- **Where we need to be:**
  - For consumer choice, P4P, quality improvement—  
need timely, cheap, actionable, credible all-payer data at the level where decisions are made
- **Where we are now:**
  - Widespread use of administrative data, with its advantages and disadvantages
  - The promise of an EMR, but much work before it can be used for these purposes
- **The vision**
  - Join forces, building on administrative data & potential of EMR, to create robust data for future
  - Pave way for merged admin and EMR data

## Joining Forces: Examples



- Use HIT to improve **timeliness**
- Add clinical detail for **accuracy, credibility**
  - condition present on admission
  - 20 lab values
- Expand **outpatient** reach (e.g. ED, physician data)
- Pilot **cross-site** data, new data **links**
- New **tools** for expanded data
- Continue privacy & **data security**

## HCUP Contacts

- Irene Fraser, Ph.D.
- Jenny Schnaier, M.A.
- Roxanne Andrews, Ph.D.
  
- HCUP website  
<http://www.hcup-us.ahrq.gov>
- HCUPnet  
<http://hcup.ahrq.gov>
- AHRQ Quality Indicators website  
<http://www.qualityindicators.ahrq.gov>

# Evidence Generation

## *new* DEcIDE Research Network

Developing Evidence to Inform Decisions about Effectiveness

- The main purpose of the DEcIDE network is to expeditiously develop valid scientific evidence about the outcomes, comparative clinical effectiveness, safety, and appropriateness of health care items and services
- The network is comprised of academic and clinic-based centers with access to electronic health information databases and the capacity to conduct accelerated research.

## What Will DEcIDE Primarily Do?

- A. Analyze existing health care databases to compare the effectiveness & outcomes of treatment.
- B. Analyze existing disease, device, and other registries.
- C. Conduct methodological studies to improve research on clinical effectiveness of treatments.

# National Center for Health Workforce Analysis: Area Resource File

**Area Resource File (ARF) is a health resource information system that enables policymakers, researchers, planners and others to analyze the current state of health care access at the county level.**

**Content includes geographic codes and classifications; health professions supply and detailed demographics; health facility numbers and types; hospital utilization; population characteristics and economic data; environment; and health professions training resources.**

**Sponsored by HRSA**

# AHRQ Data Center

- Provides researchers access to non-public use MEPS data (except directly identifiable information) and other restricted data sets;
- Mode of data analysis
  - on a secure LAN at AHRQ, Rockville
  - task order agreement with data contractor
  - combinations of both.

# User Supplied Secondary Data

- **Data Type and/or Source**
  - Area Resource File
  - Health Care Market Variables @ zip code level
  - Proprietary county level HMO variables
  - State and MSA level data from Interstudy Publications
  - State level Medicaid and poverty variables
  - County level unemployment rates
  - State level data from BLS
  - NHIS
  - Urban Institute
  - Academy for Health Services Research and Policy
  - Census Bureau
  - HCFA
  - Proprietary state level data
  - State income tax rates
  - Centers for Medicare and Medicaid Services
- **Research Focus**
  - Changes in Medicaid and SCHIP
  - Access to Care Issues
  - Changes in Health Insurance Coverage
  - Disparities in Health Care Expenditures for Families
  - State Level Health Care Expenditures

# Limitations

- Greater restrictions in data access for public use
- Competing demands on host sample frames
- More frequent survey contacts reduce overall response rate
- Requires greater coordination across data sources and organizations

# Summary

- Capacity of integrated survey designs to serve as cost efficient sampling frames
- Capacity of integrated survey designs to reduce bias attributable to nonresponse
- Related enhancements to data quality and analytical capacity
- MEPS applications
- Limitations
- Discussion questions