ovation
a division of ICON Clinical Research

Use and Integration of Freely Available U.S. Public Use Files to Answer Pharmacoeconomic Questions: *Deciphering the Alphabet Soup*

Prepared by Ovation Research Group
for the National Library of Medicine
October 20, 2006

Public Use Files: *The Opportunities*

❑ Free or low-cost

❑ Can answer many Pharmacoeconomic questions

- Prevalence

- Incidence

- Utilization

- Unit costs

- Trends over time

❑ Helpful for benchmarking other sources, or for sensitivity analyses

Public Use Files:
Your Questions and the Alphabet Soup

HCUP

Prevalence?

BRFSS

MEPS

Utilization?

Incidence?

NHDS

NAMCS

Time Trends?

Unit Costs?

NHANES

NHIS

Where to Begin?

❑ Identify questions that are appropriate for PUF data

❑ Classify data by source and content

❑ Step through examples matching PUF data to research questions

❑ Review general guidelines and pitfalls to avoid

❑ Provide resources for future use

Questions that PUF Data can Answer

❑ Prevalence of chronic disease

 ▪ How many adults have arthritis in the US?

❑ Incidence of acute diseases or events

 ▪ What is the incidence of acute respiratory failure?

More Questions that PUFs can Answer

❑ Medical resource use associated with a condition or procedure

  ▪ Hospitalizations

  ▪ Emergency Department, Outpatient, or ambulatory visits

❑ Disease Direct Costs

  ▪ Total amounts paid for health care for individuals with a disease

❑ Indirect costs

  ▪ Usually lost earnings attributable to a disease

Where to Begin?

❑ Identify questions that are appropriate for PUF data

❑ Classify data by source and content

❑ Step through examples matching PUF data to research questions

❑ Review general guidelines and pitfalls to avoid

❑ Provide resources for future use

Classification of PUF Data: *Definitions*

❑ Population-based surveys

- ▪ Generalizable to the non-institutionalized population
- ▪ Include socio-demographic information
- ▪ Information usually based solely on respondent self-report

❑ Facility-based visit samples

- ▪ Disease and utilization information from medical records
- ▪ Prevalence inferences must be made cautiously

PUF Data: *More Definitions*

❑ Administrative

 ▪ Based on records of utilization kept by public agencies (e.g., Medicaid, Medicare)

 ▪ Due to privacy concerns (HIPAA) Medicare and Medicaid claims data are now more difficult to obtain

❑ Other

 ▪ Utilities and files available for download

PUF Population-Based Surveys

❑ National Health Interview Survey (NHIS)

- Conducted annually by NCHS

- Certain "priority" conditions asked of all adult and child respondents

- Supplements available in various years

❑ Medical Expenditure Panel Survey (MEPS)

- Conducted annually by AHRQ since 1996

- Household survey includes information on health-care resource use, costs, and insurance coverage

PUF Population-Based Surveys *(cont)*

❑ National Health and Nutrition Examination Survey (NHANES)

- ▪ Six waves of data available

- ▪ Health information from physical and lab examinations

- ▪ Wide range of disease information (e.g., infectious diseases, risk factors for cardiovascular disease)

PUF Population-Based Surveys *(cont)*

❑ Behavioral Risk Factor Surveillance System (BRFSS)

- Conducted annually by state agencies and NCHS

- Gathers information on health behaviors linked to leading causes of death (e.g., heart disease, cancer, stroke)

- Telephone survey

PUF Population-Based Surveys *(cont)*

❑ National Survey on Drug Use & Health (NSDUH)

- Conducted annually by the Substance Abuse and Mental Health Services Administration (SAMHSA)

- Designed to produce drug and alcohol use incidence and prevalence estimates in the general U.S. civilian population aged 12 and older

- Also reports the consequences and patterns of use and abuse

- For 12-17 year olds, hospital or other treatment for substance abuse is recorded

PUF Population-Based Surveys *(cont)*

❑ Longitudinal Studies of Aging (LSOAs)

- Collaborative project of National Center for Health Statistics (NCHS) and the National Institute on Aging (NIA)

- Study of individuals 70 years of age and over that records health, functional status, living arrangements, and health services utilization over time

- Data obtained from a variety of sources (e.g., surveys, Medicare claims)

- Chronic and acute conditions, as well as cause of death recorded

PUF Population-Based Surveys *(cont)*

❑ National Immunization Survey (NIS)

- Conducted annually by NCHS

- Provides state and local area estimates of vaccination coverage in children between 19-35 months of age

- Specific vaccinations administered to each child as reported by family and (optionally) by provider are recorded

PUF Facility-Based Samples

❑ National Hospital Discharge Survey (NHDS)

  ▪ Conducted annually by NCHS

  ▪ Discharges from non-institutional, non-Federal hospitals

  ▪ Primary and secondary diagnosis and procedure codes

  ▪ Length of stay, discharge status, demographics

PUF Facility-Based Samples *(cont)*

❑ National Ambulatory Medical Care Survey (NAMCS) and National Hospital Ambulatory Medical Care Survey (NHAMCS)

- Conducted annually by NCHS

- Sample of patient visits to office-based physicians (NAMCS) or Emergency or Outpatient Hospital Departments (NHAMCS)

- Primary and secondary diagnosis and procedure codes

- Drug provided/prescribed

- Demographics

PUF Facility-Based Samples *(cont)*

❑ Health Care Utilization Project National Inpatient Sample (HCUP-NIS)

  ▪ Released annually by AHRQ

  ▪ Currently includes 1,004 hospitals in 37 states

  ▪ Similar information as NHDS, plus total charges and median income of patient's residence

  ▪ Small charge (e.g., $200 for 2004)

  ▪ Some statistics can be run online at no charge

PUF Facility-Based Samples *(cont)*

❑ Health Care Utilization Project Kids' Inpatient Database (HCUP-KID)

- ▪ Available for 1997, 2000, and 2003 from AHRQ

- ▪ 2003 release includes hospitals in 36 states

- ▪ Same information as HCUP-NIS, but large enough to study rare conditions in children

- ▪ Small charge (e.g., $200 for 2003)

PUF Facility-Based Samples *(cont)*

- ❑ National Nursing Home Survey (NNHS)
  - ▪ Available for 1995, 1997, 1999, 2004
  - ▪ National sample of nursing home residents and staff
  - ▪ Diagnoses at admission
  - ▪ Use of various medical devices at admission
  - ▪ Demographics
  - ▪ Length of stay, total and daily charges

PUF Facility-Based Samples *(cont)*

❑ National Home and Hospice Care Survey (NHHCS)

- ▪ Available for 1992, 1994, 1996, 1998, 2000
- ▪ Information is collected from home and hospice agencies and their patients
- ▪ Diagnoses and procedures
- ▪ Demographics

PUF Facility-Based Samples *(cont)*

- ❑ Medicare Current Beneficiary Survey
  - ▪ Longitudinal sample of Medicare Beneficiaries by NCHS
  - ▪ Up to 4 years of data for each respondent
  - ▪ Health status, functioning, demographics
  - ▪ Utilization /cost information merged from administrative data
  - ▪ Must sign a data use agreement and purchase the files for $480 per year.
  - ▪ Certain high-level results for each year available online.

PUF Facility-Based Samples *(cont)*

- ❑ Medicare Health Outcomes Survey (HOS)
    - Conducted periodically by CMS (Centers for Medicare and Medicaid Services)
    - Samples Medicare, managed care enrollees
    - Physical functioning and well-being at baseline & follow-up
    - No sampling weights yet for inference to entire Medicare managed care population

PUF Facility-Based Samples *(cont)*

❑ National Compensation Survey (NCS)

- ▪ Conducted periodically by BLS; most recently in 2005
- ▪ Sample of workplaces by size
- ▪ Wage information by occupation, industry, gender
- ▪ Helpful for assigning unit costs for lost work time
- ▪ Learning curve to find the data you need

PUF Facility-Based Samples *(cont)*

❑ National Survey of Ambulatory Surgery

- ▪ Study of ambulatory surgical care in hospital-based and freestanding ambulatory surgery centers (ASCs).

- ▪ Originally conducted from 1994 to 1996, but it was discontinued due to lack of resources.

- ▪ The NSAS is being conducted again in 2006.

- ▪ Data for the NSAS will be collected for approximately 60,000 ambulatory surgery cases in 2006 from a nationally representative sample of hospital-based and freestanding ambulatory surgery

PUF Administrative Data

❑ Medicaid State Drug Utilization Data

- Available for 1996 to present in annual files

- State and national level

- NDC, FDA product name

- Total reimbursed amount and total units reimbursed

PUF Administrative Data *(cont)*

❑ Medicare Limited Dataset Standard Analytic Files

- Available for 1991 – 2004 in a series of files:
  - Part A claims (inpatient, outpatient, SNF, hospice, or HHA)
  - Part B claims (physician/supplier services, durable medical equipment)
  - Denominator
- All entities must apply in order to purchase data
- LDS mask date of service to the quarter of service, and age to 5-year categories
- Total amounts charged and paid provided
- Steep learning curve

PUF Administrative Data *(cont)*

❑ Medicare Payment Rates and RVU

- CMS provides annual RVU files on its web site

- Not as user friendly as what you can buy (e.g., Physician Fees on disk)

- Physician, clinical laboratory, ambulatory surgical procedures, and durable medical equipment available through an interactive web site
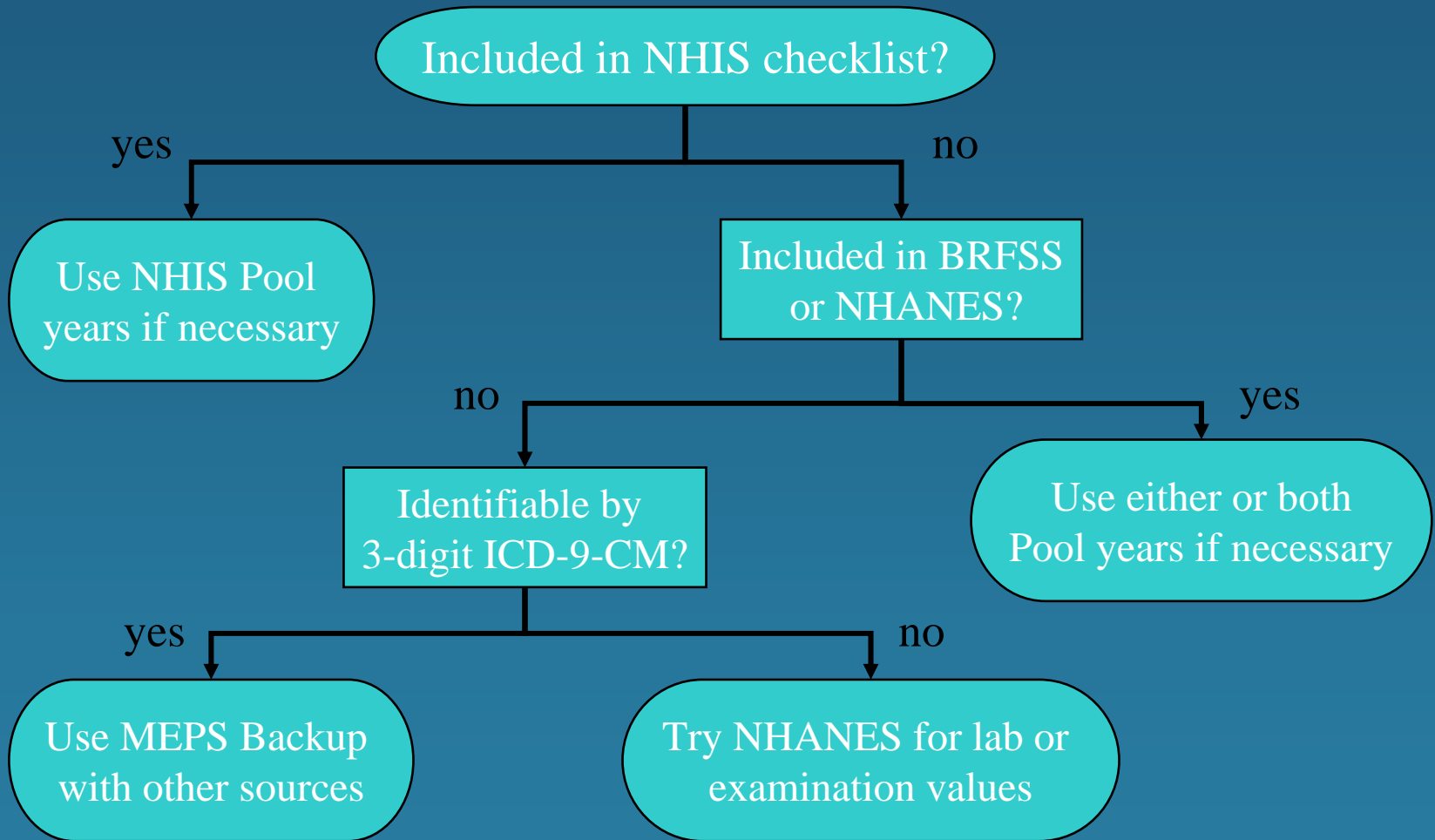
PUF Utilities

- ICD-9-CM Diagnosis Codes
  - Codes with short definitions updated yearly by CMS

- ICD-9-CM Conversion File
  - Records changes in diagnosis and procedure codes through time

- Statistical Abstract of the United States
  - Published annually by US Census Bureau
  - Population information can be used for prevalence/incidence denominators
  - CPI information to standardize costs

Where to Begin?

❑ Identify questions that are appropriate for PUF data

❑ Classify data by source and content

❑ Step through examples matching PUF data to research questions

❑ Review general guidelines and pitfalls to avoid
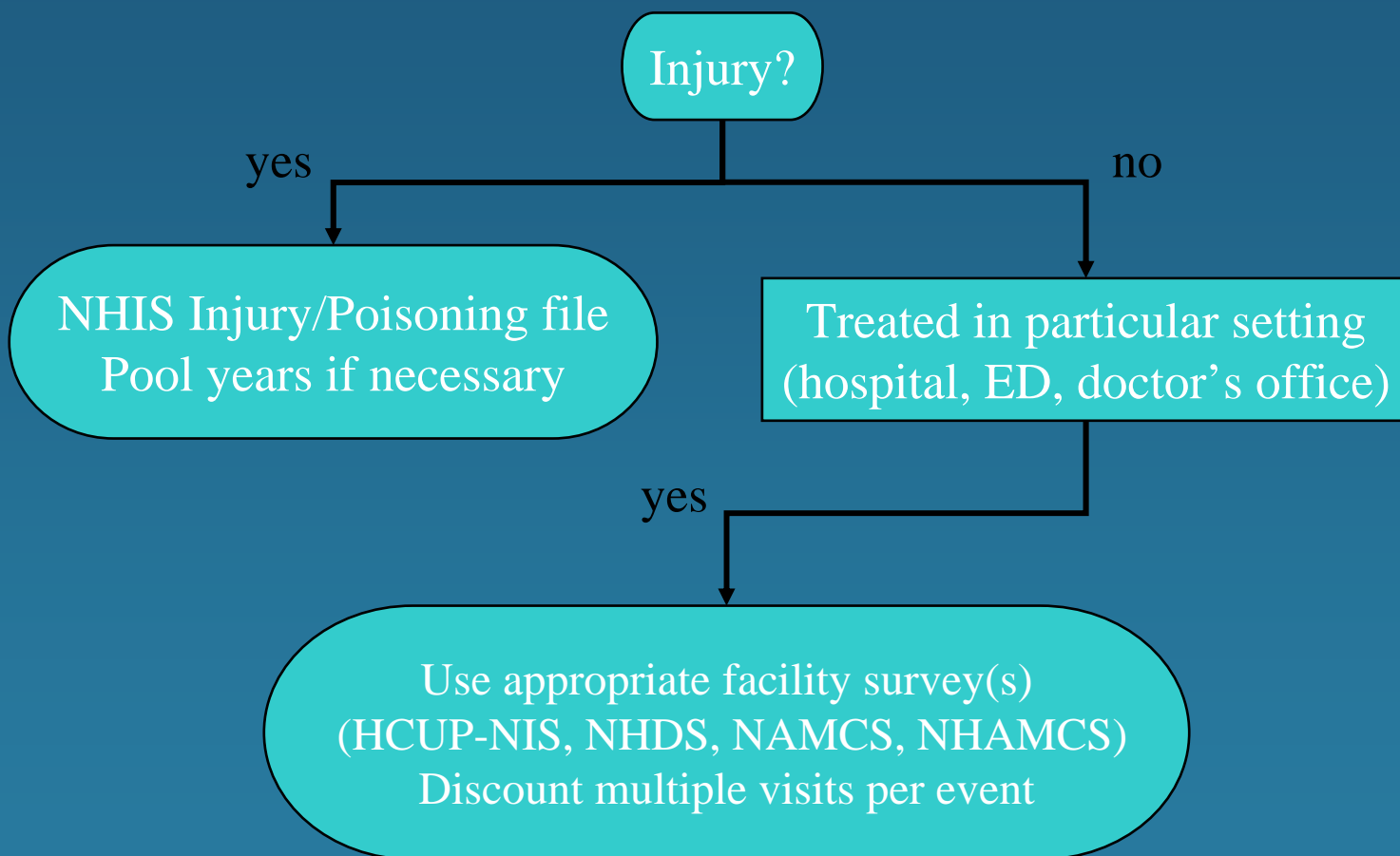
❑ Provide resources for future use

Answering Your Questions: *Prevalence of a Chronic Condition*

Included in NHIS checklist?

yes

no

Use NHIS Pool years if necessary

Included in BRFSS or NHANES?

no

yes

Identifiable by 3-digit ICD-9-CM?

Use either or both Pool years if necessary

yes

no

Use MEPS Backup with other sources

Try NHANES for lab or examination values

Prevalence of A Chronic Condition: *Example Using NHIS*

❑ Prevalence of adults with asthma in the non-Institutionalized Population in 2001.

❑ Use NHIS

  ▪ "Have you ever been told by a doctor or other health professional that you have asthma?" included in adults condition sample

❑ The Answer

  ▪ Weighted estimate: 22.2 million out of 203.8 million adults (10.9%) in 2001 had been diagnosed with asthma.

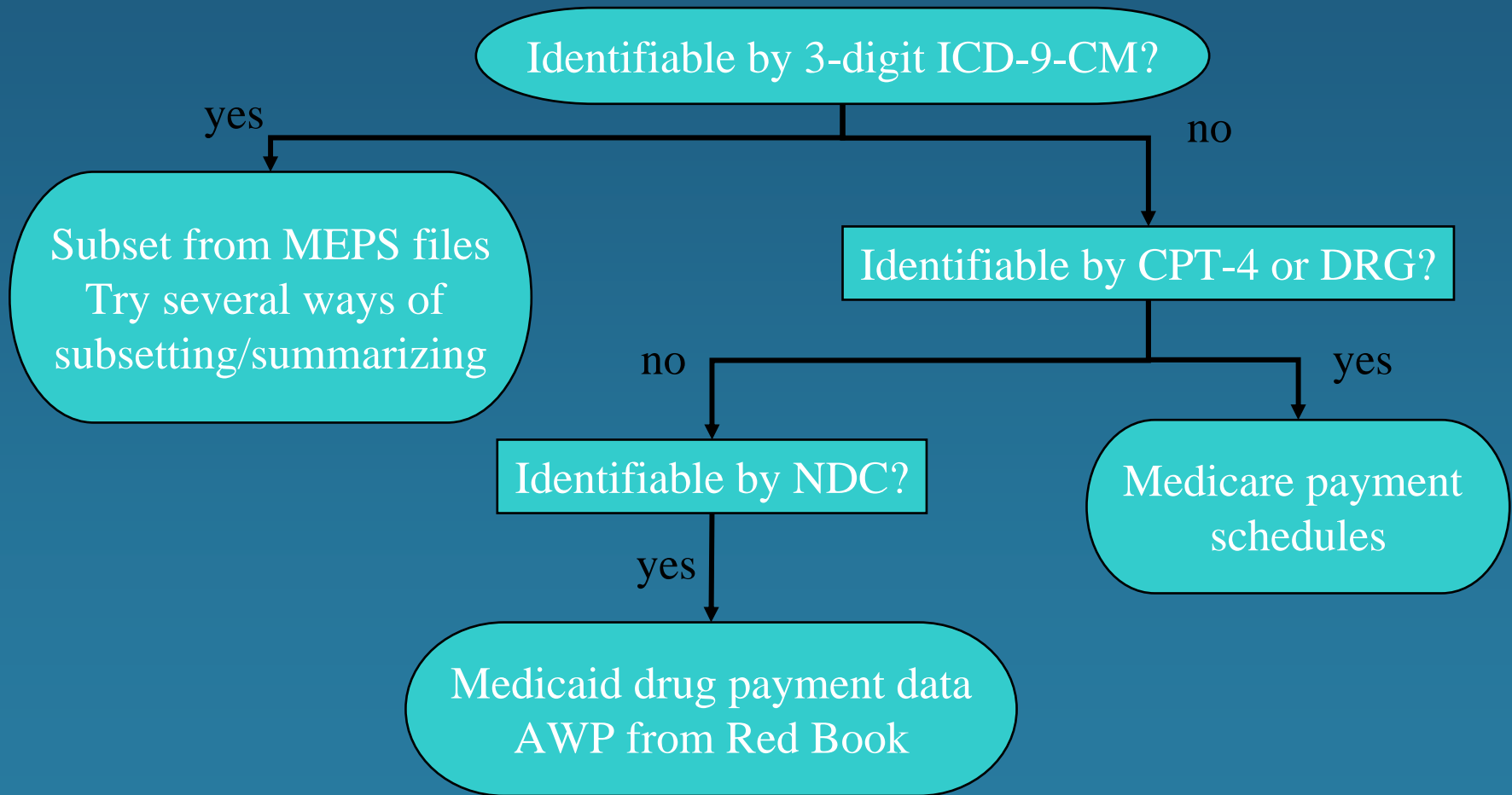 ▪ May include individuals who no longer experience symptoms

Answering Your Questions: *Incidence of Acute Disease or Injury*

Injury?

yes → NHIS Injury/Poisoning file
Pool years if necessary

no → Treated in particular setting (hospital, ED, doctor's office)

yes →

Use appropriate facility survey(s)
(HCUP-NIS, NHDS, NAMCS, NHAMCS)
Discount multiple visits per event

Incidence of an Acute Condition:
Example Using NHIS & Facility Surveys

❑ Report to National Academies of Sciences Institute of Medicine on the Epidemiology of Poisoning (Cisternas and Blanc)

❑ PUF Sources

  ▪ NHIS injury/poisoning supplement

  ▪ NAMCS, NHAMCS, and NHDS

❑ Details

  ▪ Poisoning identifiable through 3-digit ICD-9-CM and E-Codes

  ▪ Known underreporting in NHIS

  ▪ Supplemented by NAMCS, NHAMCS and NHDS

    • Discounted for multiple visits per individuals

    • Discounts based on episode-of-care information and expert opinion

Answering Your Questions: *Unit Costs for Direct Medical Utilization*

Identifiable by 3-digit ICD-9-CM?

yes

no

Subset from MEPS files
Try several ways of
subsetting/summarizing

Identifiable by CPT-4 or DRG?

no

yes

Identifiable by NDC?

Medicare payment
schedules

yes

Medicaid drug payment data
AWP from Red Book

Unit Costs for Direct Medical Utilization: *Example Using MEPS & AWP*

❑ Cisternas et al. "A Comprehensive Study of the Direct and Indirect Costs of Adult Asthma." *J Allergy Clin Immunol*. 111 (6):1212-1218.

❑ Public Cost Sources

  ▪ MEPS for office and emergency visits and hospitalizations

  ▪ Red Book (AWP) for drug prices

❑ Details

  ▪ Used positive paid amounts (not charges) from MEPS

  ▪ Pooled several years of MEPS and standardized cost to a base year using medical component of the CPI

  ▪ Calculated weighted average of AWP for all NDCs in drug classes based on market share

Where to Begin?

❑ Identify questions that are appropriate for PUF data

❑ Classify data by source and content

❑ Step through examples matching PUF data to research questions

❑ Review general guidelines and pitfalls to avoid

❑ Provide resources for future use

Suggestion #1:
Answer the Following Questions First

- ❑ What segments of the population are affected by this condition/procedure?
  - ▪ Adults? children? elderly? nursing home residents?

- ❑ In which patient settings does a treatment of interest occur?
  - ▪ Inpatient? ED? Doctors office?
  - ▪ Usually resolved in one patient encounter?

- ❑ Which ICD-9-CM, CPT, or NDCs are used to identify your disease or treatment?
  - ▪ Is identification straightforward, or is there inherent uncertainty?

Suggestion #2: *Know When to Pool*

❑ As a general rule, estimates from NCHS surveys require N>50

❑ Can usually pool data from several waves/years

❑ Check record layout and coding for each variable of interest in every year

   ▪ Variables such as race, insurance status have changed through time

Suggestions #3: *Apply Common Sense*

- ❑ Conduct a thorough literature review and compare your results to past studies

- ❑ If time/budget permit, use several data sources and compare results

- ❑ If complicated file merging is necessary, ensure you have technical expertise (SAS, SPSS) on hand

- ❑ Conduct sensitivity analyses
  - ▪ Even using the same source, several definitions of your population can be applied

PUF Pitfall #1: *Coding Inaccuracies*

❑ The Problem

  ▪ Medical coding is an art, not a science

  ▪ Variation across coders

  ▪ Upcoding to maximize reimbursement

❑ Mitigation

  ▪ For diagnoses, compare results from primary field selection to any field inclusion

  ▪ Compare distributions between payors/type of insurance for consistency

PUF Pitfall #2: *Self-Report &Recall Bias*

❑ Problem for population-based surveys (MEPS, NHIS, BRFSS)

- Respondents may not know their diagnosis
- Respondents may forget diagnosis history

❑ Mitigation

- Augment with estimates from administrative sources
- Augment with estimates from facility-based surveys

PUF Pitfall #3:
Different Cost Perspectives

❑ Perspectives can include charges, paid amounts, or allowed amounts

❑ Pick a perspective and try to estimate an adjustment factor to apply to other perspectives

❑ Example

  ▪ Your study has chosen paid amount perspective

  ▪ HCUP-NIS only provides charges

  ▪ Develop a charge-to-paid ratio using similar hospitalizations from MEPS

PUF Pitfall #4:
Combining Sources from Different Years

❑ Check for changes in ICD-9-CM or CPT coding if applicable, e.g., hepatitis or AIDS

❑ Pick a base year

❑ If combining costs, adjust to base year using medical component of the CPI (from Statistical Abstract of the United States)

❑ Don't forget to divide the annual weight by the number of pooled years and adjust denominator appropriately.

Conclusions

❑ PUF data are a cost effective resource for Pharmacoeconomic questions

❑ Many are freely downloadable via the web

❑ However, these sources should be used carefully and be supported by other estimates when possible.