Collecting Income Data:
The National Health Interview Survey

James M. Dahlhamer
Statistician (Health), NCHS

Hearing on Minimum Data Standards for the Measurement of
Socioeconomic Status in Federal Health Surveys
March 8-9, 2012

Overview of Talk

❑ Brief overview of the National Health Interview Survey (NHIS)

❑ Brief overview of SES measures in the NHIS

❑ Collection of income data in the NHIS

❑ Changes to the collection/release of NHIS income data to address quality/usability issues

  ▪ Address committee questions on design challenges

❑ General responses to other committee questions

The National Health Interview Survey

- In-person household interview survey
  - Some telephone follow-up

- Household respondent provides basic socio-demographic information on all members of household

  - For each family within a household

    - knowledgeable adult 18 or older responds for him/herself and all other family members
    - *sample adult* 18 or older and *sample child* under age 18 randomly selected for more extensive set of questions

Collection of SES Data in the NHIS

- Family Socio-demographic (FSD) Section
  - Education
  - Employment status
  - Personal earnings (exact amount)

- Family Income (FIN) Section
  - Sources of income (receipt); self-employment, pensions, investments, entitlement programs, other
  - Total family income (exact amount)
    - Follow-up questions for nonresponders
  - Housing tenure (own/rent)
  - Program participation (food stamps)

- Sample Adult Sociodemographic Section
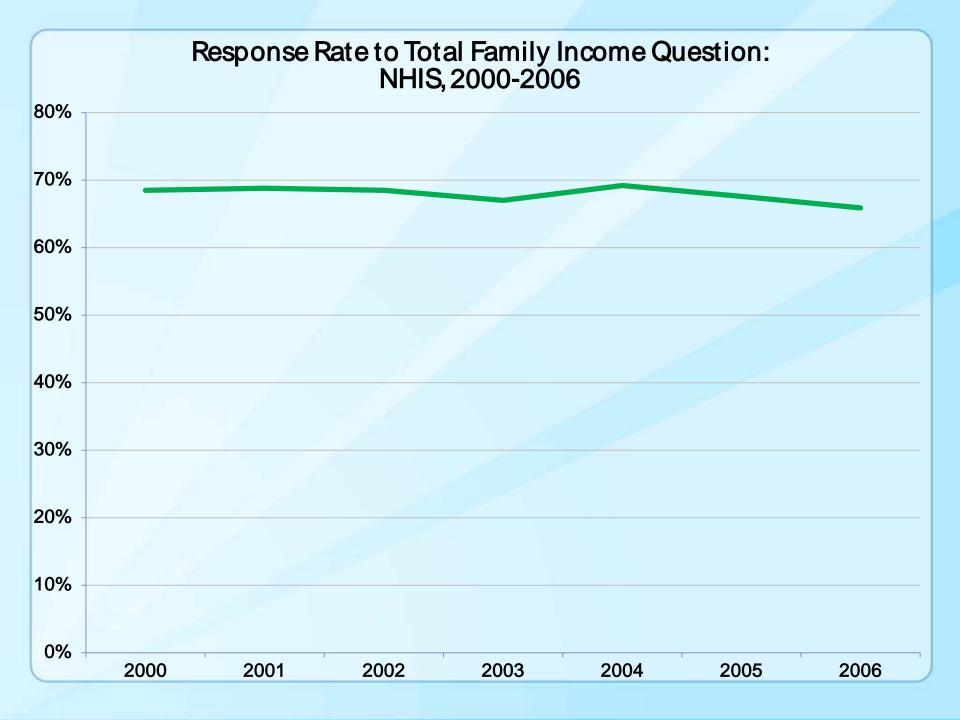 - Occupation and industry; follow-ups on number of employees, tenure, paid sick leave, paid by the hour, etc.

Collection of Income Data in the NHIS

❑ Discuss the collection of income data on the NHIS, with emphasis on the following committee questions:

- What are the methodological issues facing federal surveys in the collection, analysis, linkage and making available data on socioeconomic status variables?

- What is the best way to collect the usual measures of SES, such as income, education, and occupation?

- How can we improve the quality of the data and for whom do we need to improve the quality of data on SES?

Collection of Total Family Income: 1997-2006

Now I am going to ask about the total combined income {for you/of your family} in {last calendar year in 4 digit format}, including income from all sources we have just talked about such as wages, salaries, Social Security or retirement benefits, help from relatives and so forth.

Can you tell me that amount before taxes?

Response Rate to Total Family Income Question: NHIS, 2000-2006

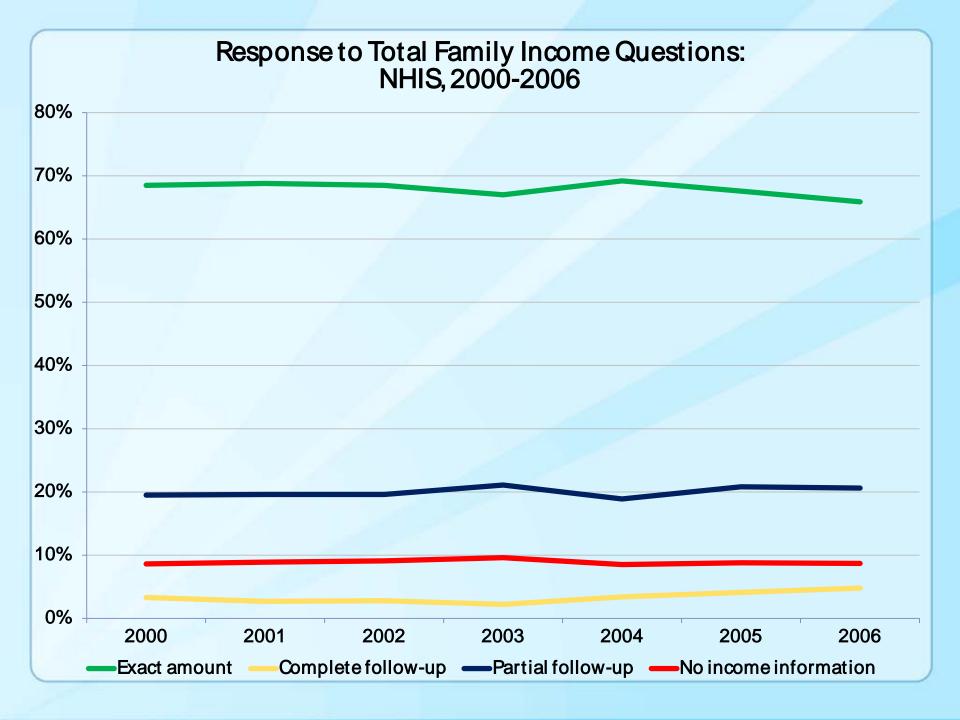# Follow-up Questions for Nonresponders: 1997-2006

You may not be able to give us an exact figure for {your/total combined family} income, but can you tell me if your income in {previous calendar year} was…
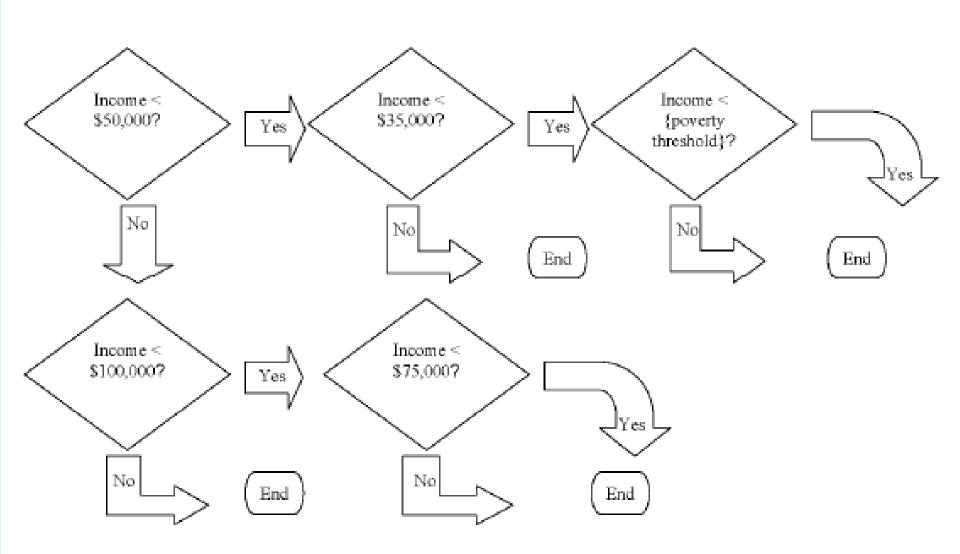
(1) $20,000 or more
(2) Less than $20,000

Follow-up Questions for Nonresponders: 1997-2006, cont.

Of those income groups, can you tell me which letter best represents {your/the total combined FAMILY} income during {previous calendar year}?

- If less than $20,000, R received flashcard with incomes in thousand dollar intervals

- If $20,000 or more, R received flashcard with incomes in thousand dollar intervals up to $35,000; then five thousand dollar intervals up to the final category of "75,000 and over"
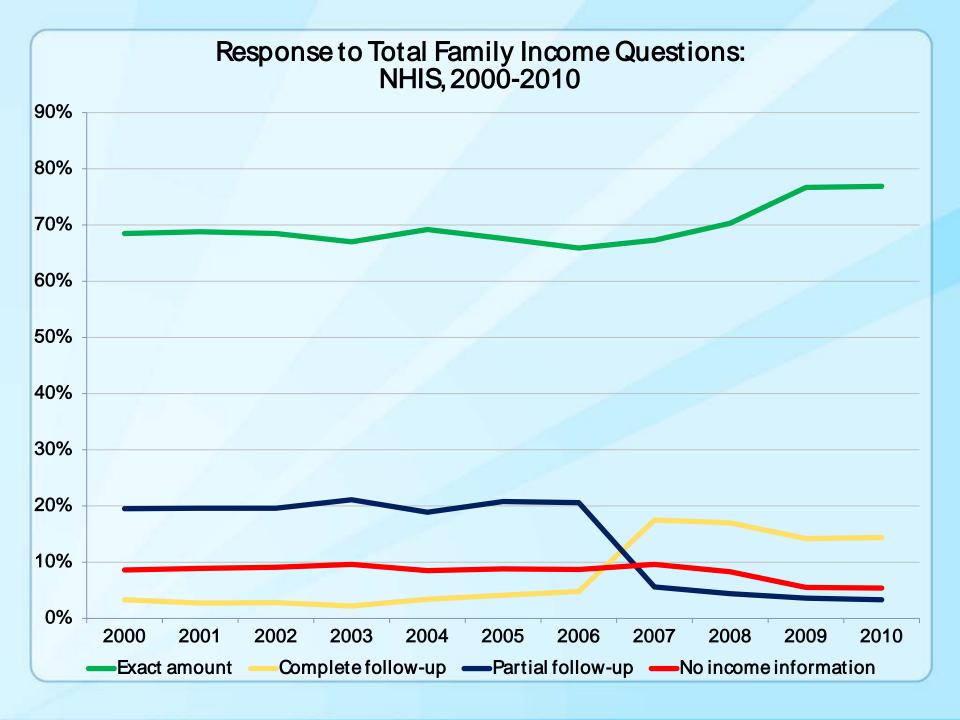
Response to Total Family Income Questions:
NHIS, 2000-2006

FLOW DIAGRAM OF FIELD-TESTED INCOME FOLLOW-UP QUESTIONS (UNFOLDING BRACKETS), NHIS 2006, 2nd QUARTER
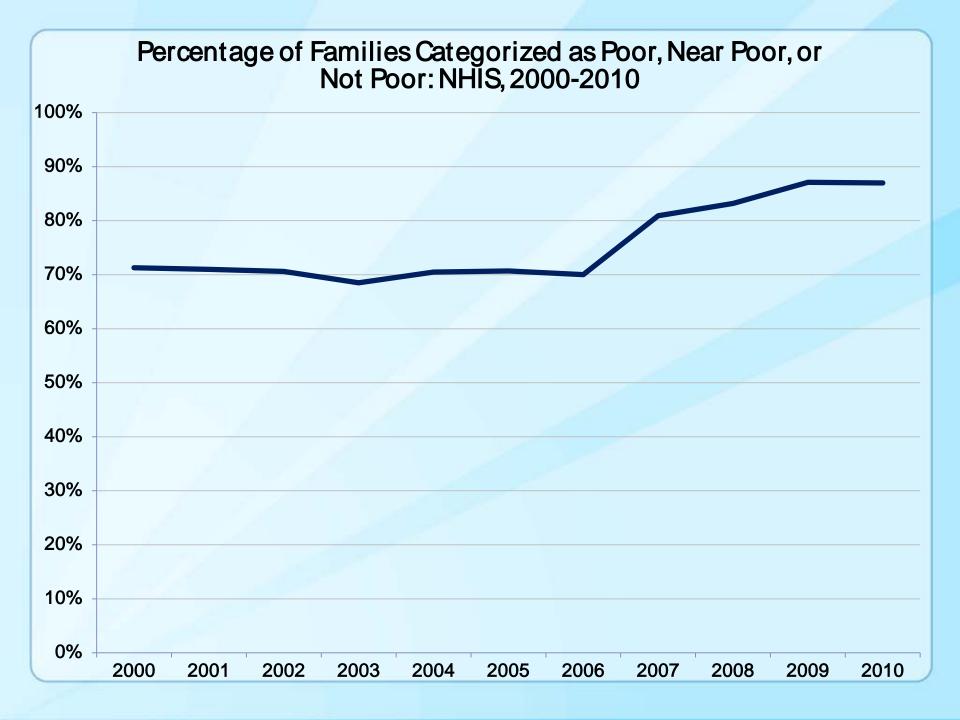
Collection of Total Family Income: 2007-Present

{When answering this next question, please remember to include your income PLUS the income of all family members living in this household.}

What is your best estimate of {your total income/the total income of all family members} from all sources, before taxes, in {previous calendar year}?

Response to Total Family Income Questions:
NHIS, 2000-2010

Percentage of Families Categorized as Poor, Near Poor, or Not Poor: NHIS, 2000-2010

Multiple Imputation (MI) of Personal Earnings and Total Family Income

❑ MI is the process of imputing more than one substitute data value for each missing value

❑ MI is more appealing than single imputation approaches because the extra variation due to imputation is incorporated into the analysis

❑ Five imputation files are released for each year of NHIS data

❑ Increased analyst burden—must work with 5 data files as opposed to 1

Recent Changes to the Collection/Release of NHIS Income Data

- ❑ Release of continuous, top-coded total family income and personal earnings amounts on public-use imputed family income / personal earnings files

  - ▪ Started with 2009 data release

- ❑ Addition of $150,000 and 200% of poverty threshold unfolding bracket questions

  - ▪ Started with the 2011 NHIS

Collection of Personal Earnings

What is your best estimate of {your/ALIAS's} earnings before taxes and deductions from ALL jobs and businesses in {previous calendar year}?

Include hourly wages, salaries, tips and commissions.

- ❑ Asked after collection of employment information, and prior to total family income

- ❑ Asked of all persons 18 years of age or older who worked in the previous calendar year

Recent Changes to the Collection/Release of NHIS Income Data

❑ Experimenting with a verification check involving total earnings and total family income discrepancies

- Discrepancies are verified when:

  - Earnings exceed income by $5,000 or more

  - Income exceeds earnings by $5,000 or more and no other sources of income are reported

- In 2011, random half of families received the verification screens if they met one of the criteria above; currently analyzing data

Collection of Wealth Data?

❑ While income and education may be good indicators of SES in middle age, wealth may be a better measure of SES for older adults

❑ Field tested wealth questions in Q2, 2006
  ▪ Total financial assets (not including home value)
  ▪ Property value; amount still owed on the property
  ▪ Car ownership

❑ Appeared to be value in collecting wealth data
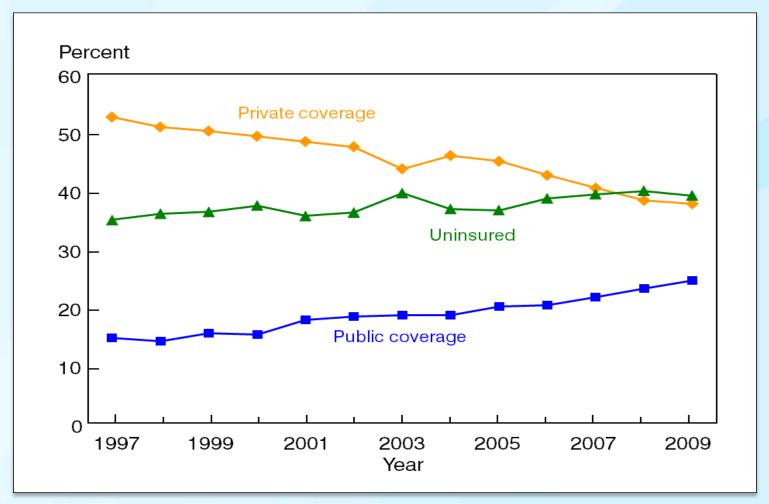 ▪ but high item nonresponse to the exact amount questions: 29.5% - 49.9%

Conclusion: Design Considerations

- ❑ How much detail to collect in non-income surveys?

- ❑ Reference periods

- ❑ Defining family/household composition

- ❑ Mode of data collection

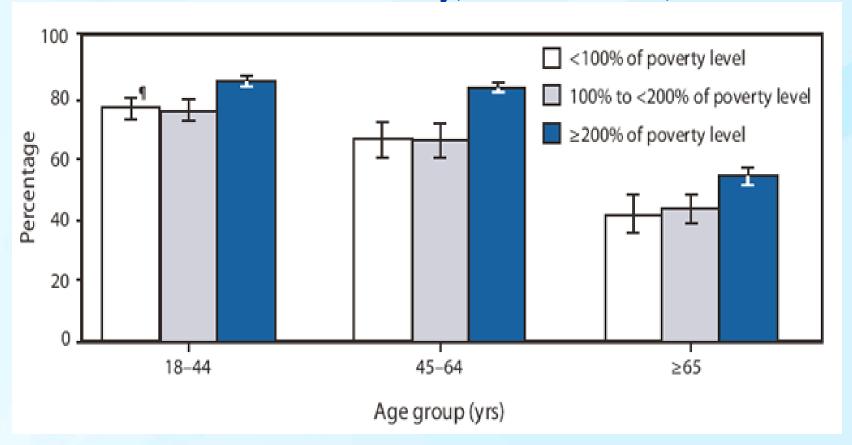- ❑ Consistency checks

- ❑ Imputation

Uses of Income

❑ Income and poverty ratio are important classification variables for health policy-related analyses

❑ Important IVs/covariates in research on heath status, health outcomes, healthcare access and utilization

❑ Income and poverty are frequently used within NCHS/DHHS reports and publications

Percentage with Health Insurance, by Coverage Type, and Percentage Uninsured at the Time of Interview, for Near Poor Adults 18-64 Years of Age: United States, 1997-2009
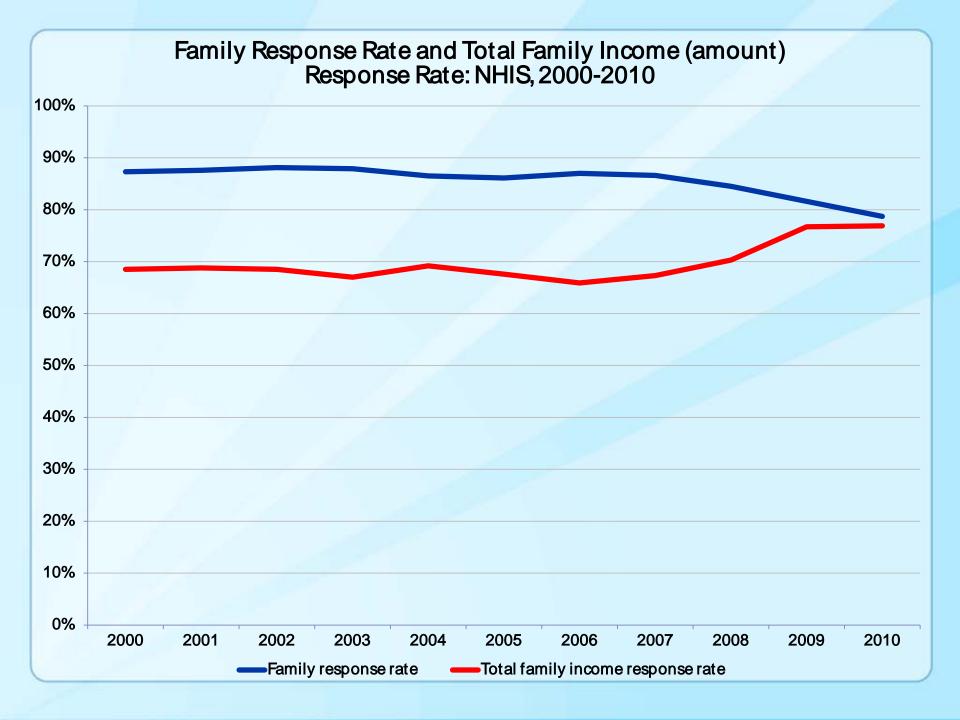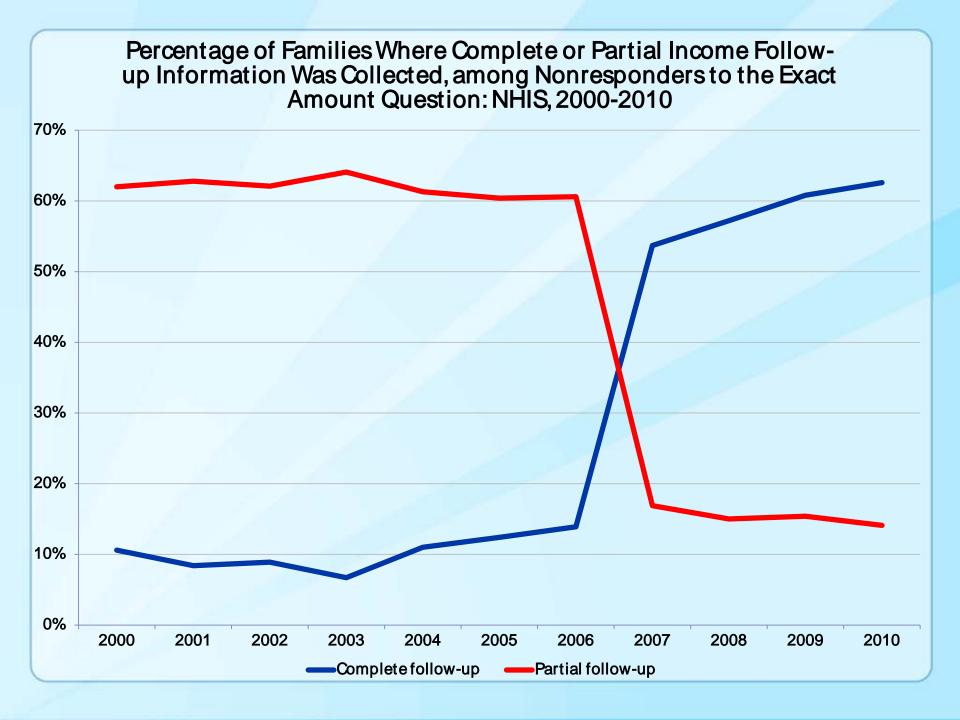
Percentage of Women Aged ≥18 Years Who had a Papanicolaou (Pap) Smear Test During the Preceding 3 Years, by Age Group and Poverty Status --- National Health Interview Survey, United States, 2008

Family Response Rate and Total Family Income (amount)
Response Rate: NHIS, 2000-2010

Percentage of Families Where Complete or Partial Income Follow-up Information Was Collected, among Nonresponders to the Exact Amount Question: NHIS, 2000-2010

2006 (Quarter 2) Field Test of Bracketing Questions

- ❑ Screened-out families (on the basis of race/ethnicity) received the family interview with one of two income sections:
  - Same as regular instrument
  - Income sources on a flashcard, bracketing follow-up questions to total family income, wealth questions

- ❑ Summary of findings
  - Bracketing follow-up questions achieved higher response rates than control follow-ups
  - More information available for calculation of poverty level
    - Control: poverty level for 68.5% of families
    - Bracketing questions: poverty level for 85.8% of families

Standard in the Measurement of SES Variables

❑ Likely more agreement on what to collect than on how to collect it

- Most federal surveys collecting income, education, employment status/occupation

❑ 2008 Mathematica Report: "Income Data for Policy Analysis: A Comparative Assessment of Eight Surveys"

- "…many differences in the income and poverty concepts used, and different surveys provide markedly differing estimates of income and poverty" (xvii).