



# Integrated Health Interview Series (IHIS)

Building an NHIS Cross-Sectional  
Time Series from 1963-Present

**Pamela Jo Johnson, MPH, PhD**

State Health Access Data Assistance Center

School of Public Health

University of Minnesota

**William C. Block, PhD**

Minnesota Population Center

University of Minnesota

NCHS Data User's Conference  
Washington DC July 10, 2006



ihis

# Purpose

- Overview
- Background to IHIS
- IHIS Collaborators
- Key components to integration project
  - Harmonization
  - Documentation
  - Dissemination
- First public data release
- Future data releases
- Demonstration of data extraction system

# Overview

- The Integrated Health Interview Series (IHIS) is a five-year NICHD-funded collaboration at U of MN
  - School of Public Health
  - Minnesota Population Center
  
- Harmonize, document, and disseminate publicly-available data from NHIS for the period from 1963 to the present

# Background

- NHIS fielded annually since 1957
  - Longest running national health survey in U.S.
  - Survey 100,000 people in 40,000 households, on average, each year
- Data are collected in-person
  - Core questionnaires and
  - Supplemental questions, which may vary by year
- Richest source of data on health and health care for the US population

# Utilization of NHIS Data

- Empirical studies published 1978 – 2003 (n = 999)
  - Search “National Health Interview Survey”
- Random sample of 100 studies using NHIS data
- Reviewed studies and tabulated the data files used
  - 70% used a single year of data
  - < 25% used data from before 1986
  - No studies used all public use data files

# Challenges to NHIS Data Use

- Complexity discourages use of multiple years of data
- Multiple codebooks and files
  - Variable locations, coding schemes, and response categories may change over time
- Critical challenges include:
  - Managing a large number of distinct files
  - Making key documentation easily accessible for each survey year
  - Harmonizing various file structures, sample weighting schemes, and variables

# IHIS Collaborators

- State Health Access Data Assistance Center (SHADAC), School of Public Health
  - **Lynn A. Blewett**, PhD (Principal Investigator)
  - Public health and health policy research
  - Experience with NHIS and other complex survey data
  
- Minnesota Population Center (MPC)
  - **Steve Ruggles**, PhD (Co-principal Investigator)
  - Data harmonization projects
  - Data dissemination infrastructure
  
- With a little help from our friends...
  - NHIS staff (Jane Gentleman, Director)
  - National advisory board
    - Chaired by E. Richard Brown (UCLA)

# IHIS Inventory (unintegrated)

- Current emphasis is 1969 - 2003
  - Data files: ~350
    - Range = 5 to 16 files per year
    - Levels: household, family, person, episodes
  - Variables: ~60,000
    - ~34,000 unique variable names
  - Household records: ~1.5 million
  - Person records: ~4 million
  - Lines of data: ~ 20 million
  - Documentation pages: Too many to count...
    - In 2000, for example, there are over 3,000 pages of documentation describing 2,451 unique variables
- Will also include 2004 – present
- Hope to eventually include 1963 – 1968

# Integrating the NHIS

- IHIS builds on model of the Integrated Public Use Microdata Series (IPUMS)
  - see [www.ipums.org](http://www.ipums.org)
- IPUMS spurred new research using harmonized microdata of U.S. census from 1850-2000
- IHIS will enhance the value of NHIS data and promote new research
  - Access to harmonized data and comprehensive documentation from entire period
  - Consistent comparisons across 40 years of dramatic change in public health

# IHIS Components

- Harmonization
- Documentation
- Dissemination
  - <http://www.ihis.us>

# Harmonization: Translation Table

	A	B	C	F	S	AH
1	<b>code</b>	<b>label</b>	<b>syntax</b>	<b>ih1969</b>	<b>ih1982</b>	<b>ih1997</b>
2	<b>rectype</b>	MARSTAT		P	P	P
3	<b>columns</b>	Legal marital status		60	88	78
4	<b>norecode</b>			0	0	0
5	<b>hide</b>					
6	<b>proj1</b>			P60	P48	P40
7	<b>proj2</b>					
8	<b>proj3</b>					
9	<b>&lt;tt&gt;</b>					
10	0	NIU	NIU		0=Under 14 years	B=Under 14 years
11	0			0=Under 17 years		
12	1	Married	Married	1=Married		
13	2	Married - Spouse present	Married - Spouse present		1=Married - Spouse in household	1=Married, Spouse in household
14	3	Married - Spouse not in household	Married - Spouse not in household		2=Married - Spouse not in household	2=Married, Spouse not in household
15	4	Married - Spouse in household unknown	Married - Spouse in household unknown			3=Married, Spouse in household unknown
16	5	Widowed	Widowed	2=Widowed	3=Widowed	4=Widowed
17	6	Divorced	Divorced	4=Divorced	4=Divorced	5=Divorced
18	7	Separated	Separated	5=Separated	5=Separated	6=Separated
19	8	Never married	Never married	3=Never married	6=Never married	7=Never married
20	9	Unknown marital status	Unknown marital status		7=Unknown	9=Unknown marital status
21	<b>&lt;/tt&gt;</b>					
22	<b>&lt;univ&gt;</b>					
23	ih1969=1981	1969-1981	Persons age 17+.			
24	ih1982=2003	1982-2003	Persons age 14+.			
25	<b>&lt;/univ&gt;</b>					
26	<b>&lt;Notes&gt;</b>					
27	<b>Comments</b>					

# Documentation: Variable Availability



## Variable Availability

### Technical Household Variables (H)

Variable Name	Variable Label	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00	01	02	03
<a href="#">RECTYPE</a>	Record type	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<a href="#">YEAR</a>	Year	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

### Technical Person Variables (P)

Variable Name	Variable Label	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00	01	02	03
<a href="#">RECTYPEP</a>	Record type, person record	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<a href="#">ASTATFLG</a>	Sample adult flag	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	X	X	X	X	X	X	X
<a href="#">CSTATFLG</a>	Sample child flag	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	X	X	X	X	X	X	X

### Core Demographic Variables (P)

Variable Name	Variable Label	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00	01	02	03
<a href="#">SEX</a>	Sex	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<a href="#">AGE</a>	Age	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<a href="#">MARSTAT</a>	Legal marital status	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

### Geography Variables (P)

# Variable Description & Comparability



## Variable Description

### MARSTAT

Legal marital status

#### **Availability**

1982-2003

#### **Universe**

1969-1981: Persons age 17+.

1982-2003: Persons age 14+.

#### **Description**

MARSTAT reports the person's marital status.

#### **Comparability – General**

Along with universe changes, there were changes in question wording and in the number of recognized categories for this variable.

The question corresponding to MARSTAT remained essentially the same for 1969-1996. At the start of this period, interviewers asked, "Is {person} now married, widowed, divorced, separated, or never married?"; at the end of this period, interviewers asked, "Is {person} now married, widowed, divorced, separated, or has {person} never been married?" The only categories recognized in these questions were those corresponding to legal marital status.

In contrast, beginning in 1997, interviewers initially asked, "{Are/Is} {you/subject's name} now married, widowed, divorced, separated, never married or living with a partner?" The inclusion of "living with a partner" melds legal marital status with something more akin to union status (since a person living with a partner could be legally married, widowed, divorced, or separated). Persons who were living with a partner were, however, asked a follow-up question about their legal marital status, and it is this

# Data Extract Preferences

Sample Selection		
<input type="button" value="Continue to Variable Selection"/>		
<b>Extract File Type</b>		
<input checked="" type="radio"/> Rectangular <input type="radio"/> Hierarchical		
<b>Command Files Format</b>		
<input checked="" type="radio"/> SPSS <input type="radio"/> SAS <input type="radio"/> Stata		
<b>Samples</b>		
<input type="checkbox"/> <a href="#">1982</a>	<input type="checkbox"/> <a href="#">1990</a>	<input type="checkbox"/> <a href="#">2000</a>
<input type="checkbox"/> <a href="#">1983</a>	<input type="checkbox"/> <a href="#">1991</a>	<input type="checkbox"/> <a href="#">2001</a>

# Download Data & Documentation

## Extracts for Download

Click on the extract number in the leftmost column to revise that extract.  
Click [here](#) to refresh your display.

Request	Created	Description	Data	Codebook	Command
<a href="#">158</a>	Jul 17, 2006, 12:26:48	Data for analysis of X, Y, and Z	<a href="#">data</a>	<a href="#">codebook</a>	<a href="#">STATA</a>
<a href="#">157</a>	Jul 11, 2006, 09:58:52	1982	<a href="#">data</a>	<a href="#">codebook</a>	<a href="#">SPSS</a>
<a href="#">156</a>	Jul 10, 2006, 14:54:49		<a href="#">data</a>	<a href="#">codebook</a>	<a href="#">SPSS</a>
<a href="#">155</a>	Jul 10, 2006, 12:54:39		<a href="#">data</a>	<a href="#">codebook</a>	<a href="#">SPSS</a>
<a href="#">154</a>	Jul 10, 2006, 09:58:39	2003	<a href="#">data</a>	<a href="#">codebook</a>	<a href="#">SPSS</a>
<a href="#">153</a>	Jul 10, 2006, 09:57:56		<a href="#">data</a>	<a href="#">codebook</a>	<a href="#">SAS</a>

# First IHIS Release

- First IHIS public data release **November 2006**
- Includes harmonized data from:
  - Household and Person files for 1969-2003
  - Sample Adult/Sample Child files for 1997-2003
  - Health Insurance supplements
  - Access to Care supplements
- For each harmonized variable, on-line documentation provides information about:
  - Years available for download
  - Values, value labels and frequencies for each year
  - Variable universe
  - Question wording
  - Interpretation of the variable
  - Comparability issues over time

# Future Data Releases

- Additional IHIS data will be released to the public on a bi-annual basis
  
- Future data releases:
  - Will accommodate other units of analysis, such as conditions, doctor visits, hospital episodes
  - Will incorporate harmonized variables drawn from other NHIS data files, including:
    - Selected conditions
    - Year 2000 Objectives
    - Child Health
    - Disability
  - We hope to include newly released public-use NHIS data from 1963 - 1968

# For Additional Information



- <http://www.ihis.us>
- [ihis@pop.umn.edu](mailto:ihis@pop.umn.edu)



STATE HEALTH ACCESS DATA ASSISTANCE CENTER

- <http://www.shadac.umn.edu>



- <http://www.pop.umn.edu>

# Thank You

- Pamela Jo Johnson
  - SHADAC, School of Public Health
  - johns245@umn.edu
  
- William C. Block
  - Minnesota Population Center
  - block@umn.edu