



# epimonitor

THE EPIDEMIOLOGY MONITOR

A monthly update covering people, events, research and key developments

## Enhanced Data Set From The National Health Interview Survey Covering Multiple Years Now Available Free of Charge

### 5,000 Variables Now Available For Study

The Minnesota Population Center and the School of Public Health at the University of Minnesota have just released an enhanced version of their Integrated Health Interview Series (IHIS) which is a more user-friendly version of the public use data sets from the National Health Interview Survey (NHIS). That survey is the principal source of information on the health of the US population and typically covers 100,000 persons in 45,000 households each year.

The survey results have been available online in downloadable form since 2001. However, data from the NHIS are not easy to use to make comparisons across time, says Miriam King, a research associate at the Minnesota Population Center. The questions on the survey are altered and the name of the variables changes over time. Tracking these changes in survey methodology in different types of documentation is a daunting task, according to King. Users have a difficult time navigating through all of the documentation associated with the

- NIH Survey, continues on page 4

- NIH Survey, con't from page 1

NHIS just to understand what is there. For these reasons, the incredible resource that is the survey is underutilized, according to King.

### Survey Made Easy

Now all this difficulty has been overcome or at least reduced by a group of investigators who are devoted to making it easier for people to use publicly collected data for free, such as census and national survey data of different types. The IHIS data set is the most health oriented data set which the group has enhanced to date, according to King.

### Enhancements

To help bring greater consistency to the data sets, the Minnesota group has recoded variables of interest as they have appeared at different times and in different forms scattered in different parts of the survey and have provided documentation of how the recoding was done. As a result, users can now make use of these more friendly versions of the NHIS public data sets to run their analyses. Prior to August 2010, the more friendly datasets had information on the years between 1969 and 2006 and had 1000 recoded variables to choose from. As of August of this year, the IHIS now has the most recent data covering the years 1969-2009 and offers users 5,000 variables to choose from in topic areas such as health insurance, access and use of health care services, health status and health conditions, and health behaviors.

### Future Releases

This is the 6th year of a 10 year project for the IHIS which data have been available now for approximately three years on the Internet. Since the project is ongoing, readers can expect to obtain

more variables as new releases are made. For example, another 2,000 variables are to be released in November 2010 around the time of the APHA meeting

### Desirable Features

What makes the IHIS so special? Actually, it has many desirable features. The data are available free, they cover multiple years, the variables have been traced over time and recoded to a single variable for consistency, documentation of the harmonization that has been carried out is prepared on each variable, and warnings are given about where important changes in survey questions have taken place.

Furthermore, users can visit the IHIS website, register to use the dataset, fill out a simple form, promise to make responsible use of the data, and receive within 2 hours a data extract that contains only the information they are interested in, that is, information only on the years and variables they want. Also, if users need to ask questions about the data sets they have or if they need personalized user support, such help is available from the producers of the more user friendly data sets.

### Contribution Made

It is important to note that the IHIS does not create new variables from the NHIS but instead makes existing variables more comparable across the years by tracking down the information on how the variable of interest was modified over time and creating a single name for each instance in which the variable was used. Thus, the considerable information that might exist on a topic within the NHIS across the years can be rescued from obscurity and made greater use of because the

- NIH Survey, continues on page 7

*"...the incredible resource that is the survey is underutilized..."*

information contained on each topic has been made more recognizable.

### Benefits of IHIS

What are the anticipated benefits of the IHIS? According to the brochure describing the project, it reduces the cost of research by minimizing initial investments in data preparation, simplifies access to data across multiple years and data files, facilitates consistent comparisons across four decades of dramatic change in public health, and documents changes in concepts and comparability issues.

So far the IHIS has attracted 900 registered users and the products that have emerged have included textbook examples, masters theses, and publications. An online bibliography of IHIS uses has been created to keep track of the outputs. Readers who wish to learn more may go to the IHIS website at [www.ihis.us](http://www.ihis.us). ■

---