



# Developing a Restricted-Use Data Analysis System (R-DAS) for the National Survey on Drug Use and Health

Jonaki Bose, Joe Gfroerer, Art Hughes, Mike Jones  
Center for Behavioral Health Statistics and Quality

**NCS Research Day – August 24, 2011**



# Outline of Presentation

- NSDUH background
- Current laws and strategies regarding release of NSDUH data, and limitations
- Consideration of new data dissemination approaches
- Current status of the Restricted Data Analysis System (R-DAS)

# National Survey on Drug Use and Health (NSDUH): Background

- Representative nationally, the 50 states, & D.C.
- Civilian, noninstitutional population, age 12+
- Face-to-face interview using CAPI/ACASI
- ~ 68,000 respondents surveyed annually
- Sensitive data collected from one or two persons per selected dwelling unit

# Questionnaire Content

- Alcohol, tobacco, and illicit drug use
- Substance use disorders (DSM-IV)
- Mental disorders (SMI, depression, suicide attempts, etc.)
- Treatment for substance use/mental health problems
- Health conditions (HIV/AIDS, etc), criminal activities
- Employment, income, program participation
- Demographics

# Applicable Laws Protecting NSDUH Data

- **Section 501(n) of the PHS Act (42 USC 290aa)** [as amended in the Children's Health Act of 2000, Pub. L. 106-310 – Oct 2000] - Limits the use and dissemination of identifiable information to only the purpose for which it was supplied
- **Title V of the E-Government Act of 2002 - Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA)** (Pub. L. 107-347 – Dec 2002) - Provides a uniform set of confidentiality protections to all identifiable data collected for statistical purposes

# Current Dissemination Strategies

- Reports and detailed tables
- Public-Use Files - Created from restricted-use files where records are recoded, swapped, and subsampled. Variables such as detailed race-ethnicity, state, and county , are deleted.
- Custom requests - Mostly tabular requests

# Limitations to Current Approach

- Custom runs are expensive – require both staff and contractor time; occasionally there are delays
- Special runs are limited – complex analyses, regressions, etc are limited or not possible under current resource constraints
- Restricted NSDUH data are not being used to their fullest potential

# Approaches Considered

- Make microdata available directly to users
- Create a restricted-use DAS for increased access to data hitherto unavailable for user-directed runs.
- Definition of R-DAS: An online data analysis system where users can run analyses without having direct access to the microdata; data is restricted-use, but analysis output is public-use.

# Considerations: Analyses

- What analyses should be permitted?
  - Cross Tabulations
  - Means
  - Regressions
  - Other
- What functions should be permitted
  - Recodes
  - List cases

# Considerations: Data

- What data sets should be included?
  - Permitted by law:
- What variables should be included?
  - State (based on many user requests)
- How should the datasets be presented?
  - Single year indicators
  - Combined years

# Considerations: Output

- What Statistical Disclosure Limitation (SDL) Methods should be implemented?
  - Cell suppression
  - Tabular suppression
  - Precision-based suppressions
  - Minimum cell sizes
  - Subsampling
  - Introduction of noise
  - Rounding
  - Unweighted numbers
  - Weighted totals
  - Number of decimal points displayed

# Considerations: Cost

- In the long run an R-DAS may be more cost effective as fewer special runs are needed.
- However, costs are incurred:
  - Creating system and statistical specifications
  - Preparing datasets for the R-DAS
    - Incorporation of new variables
    - Documentation
  - Adding new data to the DAS
    - Annually
    - Maintaining/updating of R-DAS

# Current Status

- A set of 3 R-DASes will be released using data from 2002 to 2009.
  - A 2-year R-DAS allowing for comparison of trends and estimates based on 2 years of data.
  - A 4-year R-DAS allowing for comparison of estimates from 2002-05 and 2006-09 and estimates based on 4 years of data
  - An 8-year R-DAS which will only allow for pooled estimates using 8 years of data
- Each R-DAS will have only one weight
- Updates planned every two years

# Information on Current NSDUH Data Products

- <http://www.oas.samhsa.gov/nsduhLatest.htm>
- <http://www.icpsr.umich.edu/icpsrweb/SAMHDA/>
- <http://www.oas.samhsa.gov/nsduh/methods.cfm#2K10>
- <http://oas.samhsa.gov/2k4/NSDUH/nsduh.pdf>
- <http://oas.samhsa.gov/2k3/NSDUH/nsduh.pdf>
- <http://store.samhsa.gov/shin/content//SMA11-4641/SMA11-4641.pdf>

# Questions?

Contact Jonaki Bose

[jonaki.bose@samhsa.hhs.gov](mailto:jonaki.bose@samhsa.hhs.gov)

240-276-1257

Art Hughes

[art.hughes@samhsa.hhs.gov](mailto:art.hughes@samhsa.hhs.gov)

240-276-1261