2001 National Household Survey on Drug Abuse

Statistical Inference Report

Contract No. 283-98-9008
RTI Project No. 7190

Authors:

Kortnee C. Barnett-Walker
James R. Chromy
Teresa R. Davis
Steven L. Emrich
Dawn M. Odom
Lisa E. Packer

Project Director: Thomas G. Virag

Prepared for:

Substance Abuse and Mental Health Services Administration
Rockville, Maryland 20857

Prepared by:
RTI

April 2003

2001 National Household Survey on Drug Abuse

Statistical Inference Report

Authors:
Kortnee C. Barnett-Walker
James R. Chromy
Teresa R. Davis
Steven L. Emrich
Dawn M. Odom
Lisa E. Packer

Project Director:
Thomas G. Virag

Prepared for:

Substance Abuse and Mental Health Services Administration
Contract No. 283-98-9008

April 2003

Acknowledgments

This publication was developed for the Substance Abuse and Mental Health Services Administration (SAMHSA), Office of Applied Studies (OAS), by RTI, Research Triangle Park, North Carolina, under Contract No. 283-98-9008. Significant contributors at RTI include Kortnee C. Barnett-Walker, James R. Chromy, Teresa R. Davis, Steven L. Emrich, Dawn M. Odom, and Lisa E. Packer. Brenda K. Porter word processed and formatted the report, Richard S. Straw copy edited the report, and Thomas G. Virag is the RTI Project Director.

Table of Contents

Section		Page
1.	Introduction	1
2.	Sampling Error	3
3.	Degrees of Freedom	5
4.	Statistical Significance of Differences	6
5.	Confidence Intervals	8
6.	Incidence Estimates	10
7.	Suppression of Estimates with Low Precision	14
	References	17

1. Introduction

Starting in 1999 and continuing through 2001, the National Household Survey on Drug Abuse (NHSDA) was implemented as part of a 5-year 50-State sample design to provide national and State estimates of drug use through 2003. A major change to the study protocol was the introduction of computer-assisted interviewing (CAI) methods for both the screening and interviewing of selected respondents.

For the 5-year 50-State design, 8 States were designated as large sample States (California, Florida, Illinois, Michigan, New York, Ohio, Pennsylvania, and Texas) and provided with samples large enough to support direct State estimates. For the remaining 42 States and the District of Columbia, smaller, but adequate, samples were selected to support State estimates using small area estimation (SAE) techniques.

Using the 50-State design, States were first stratified into a total of 900 field interviewer (FI) regions (48 regions in each large sample State and 12 regions in each small sample State). Within FI regions, adjacent Census blocks were combined to form the first-stage sampling units called "segments." Eight sample segments per FI region were fielded during the 2001 survey year. These sampled segments were allocated equally into four separate samples, one for each 3-month period during the year, so that the survey remains in the field year-round. Beginning in 2000, a supplemental sample was added in order to conduct a separate survey titled the Validity of Self-Reported Drug Use in Population Survey, which evaluated and established baseline information on the validity of survey research methods in assessing recent drug use among the general household population. Previous validity research had been conducted on only specific subgroups, resulting in the need to examine the validity of the NHSDA drug use data on the overall population. For more detailed information on the sample design and the validity study, see the 2001 NHSDA sample design report (Bowman, Penne, Chromy, & Odom, 2003).

Also included in the 2001 survey was an experimental study to evaluate the effectiveness of respondent incentives in improving response rates. In the first two quarters of 2001, a randomized, split-sample experimental design was included with the main study data collection of the NHSDA to compare the impact of $20 and $40 incentive treatments with a $0 control group on measures of respondent cooperation, data quality, survey costs, and population substance use estimates. To control for interviewer effects, the same FIs were required to work all of the control and treatment cases in an FI region whenever possible. A total of 9,600 respondents participated in the experiment, and all were included in the computation of 2001

NHSDA estimates. Full documentation can be found in a report on the 2001 NHSDA incentive experiment (Eyerman & Bowman, 2002).

Finally, a supplemental sample was added in quarter 4 to the New York City area following the events of September 11, 2001. In an effort to measure the impact of the attacks on drug use prevalence and mental health service utilization, the targeted sample was increased by 600, 150, and 150 in New York, New Jersey, and Connecticut, respectively. For additional information on this analysis, see the resulting analytic series report (Office of Applied Studies [OAS], 2002).

The final respondent sample of 68,929 persons for the 2001 NHSDA was representative of the U.S. general population (the civilian, noninstitutionalized population) aged 12 or older in the year 2001. In addition, State samples were representative of their respective State populations.

2. Sampling Error

The national estimates, along with the associated variance components, were computed using a multi-procedure package called SUDAAN: Software for Statistical Analysis of Correlated Data (RTI, 2001). The final, nonresponse-adjusted, and poststratified analysis weights were used in SUDAAN to compute unbiased design-based drug use estimates. All of the variance estimates were calculated using the SUDAAN option called DESIGN=WR, which is unbiased for linear statistics based on multistage clustered sample designs where the first-stage (primary) sampling units are drawn with replacement. In the past, options other than DESIGN=WR had been employed. For more information, see the 2000 NHSDA statistical inference report (Davis, Packer, Heller, & Chromy, 2002).

Because of the nature of stratified-clustering sampling design, key nesting variables were created for use in SUDAAN to capture explicit stratification and to identify clustering. For the 2001 NHSDA, each FI region consisted of its own stratum. Two replicates per year were defined within each variance stratum. The first replicate consisted of those "phasing out" segments (would not be used in the next survey year). The second replicate was made up of those "phasing in" segments (would be fielded again the following year), thus constituting the 50 percent overlap between survey years. Each variance replicate consisted of four segments, one segment for each quarter of data collection.

Estimates of means or proportions, $\hat{p}_d$, such as drug use prevalence, take the form of nonlinear statistics where the variances are not capable of being expressed in closed form. Variance estimation for nonlinear statistics in SUDAAN is based on a first-order Taylor series approximation of the deviations of estimates from their expected values (RTI, 2001).

Estimates of domain totals, $\hat{Y}_d$, corresponding to proportion estimates, $\hat{p}_d$, can be estimated as

$$\hat{Y}_d = \hat{N}_d \cdot \hat{p}_d ,$$

where

$\hat{N}_d$ = estimated population total for domain *d* and

$\hat{p}_d$ = estimated proportion for domain *d*.

The standard error (SE) for the total estimate is obtained by multiplying the SE of the proportion by $\hat{N}_d$, that is,

$$SE\left(\hat{Y}_d\right) = \hat{N}_d \cdot SE\left(\hat{p}_d\right).$$

This approach is theoretically correct when the domain size estimates, $\hat{N}_d$, are among those forced to Census Bureau population projections through the weight calibration process (Chen, Gordek, Murtha, Singh, Westlake, & Yu, 2003). In these cases, $\hat{N}_d$ is clearly not subject to sampling error.

For domain totals, $\hat{Y}_d$, where $\hat{N}_d$ is not fixed, this formula may still provide a good approximation if it can be reasonably assumed that the sampling variation in $\hat{N}_d$ is negligible relative to the sampling variation in $\hat{p}_d$. In most analyses conducted for prior years, this has been a reasonable assumption.

However, for a subset of tables produced from the 2001 data, it was clear that the above approach yielded an underestimate of the variance of a total because $\hat{N}_d$ was subject to considerable variation. In these cases, a different method was used to estimate variances in which a direct variance estimate of the linear statistic that estimates a population total was taken from SUDAAN.

3. Degrees of Freedom

To determine whether the observed difference between estimates is statistically significant, the degrees of freedom are needed to locate the corresponding probability level (*p* value) of the test statistic. A "test statistic" is a random variable that has some predetermined distribution (i.e., normal, chi-square, F) that represents the difference between the estimates. The "degrees of freedom" refer to the amount of variation allowed due to sampling error.

SUDAAN automatically calculates the degrees of freedom as the number of primary sampling units (variance replicates) less the number of strata for the dataset being analyzed. SUDAAN also allows the user to run an analysis on subpopulations of the data through the SUBPOPN statement. However, even though the SUBPOPN statement is used, SUDAAN will complete its analysis using the total degrees of freedom of the entire data unless specified otherwise by the user. This can be done in SUDAAN by entering the appropriate degrees of freedom using the DDF option.

In NHSDA analyses, the degrees of freedom are based on the first-level stratification (i.e., the FI regions). When producing estimates on the national level, there are 900 degrees of freedom. If the analysis only involves certain States, the degrees of freedom change depending on whether the State is a large sample or small sample State. The large sample States (California, Florida, Illinois, Michigan, New York, Ohio, Pennsylvania, and Texas) each have 48 degrees of freedom. All other States (small sample States and the District of Columbia) have 12 degrees of freedom.

4. Statistical Significance of Differences

Once the degrees of freedom have been determined, various methods used to compare prevalence estimates could be employed. This section describes some of these methods. Customarily, the observed difference between estimates is evaluated in terms of its statistical significance. "Statistical significance" refers to the probability that a difference as large as that observed would occur due to random error in the estimates if there were no difference in the prevalence rates for the population groups being compared. The significance of observed differences is generally reported at the 0.05 and 0.01 levels.

Due to the change in the NHSDA's design, the precision of its substance use prevalence estimates has improved, but the redesign also has made it difficult to assess long-term trends. It is not considered appropriate to compare 1999 CAI and later CAI estimates with earlier NHSDA estimates to assess trends in substance use because of the major differences between the CAI and the paper-and-pencil interviewing (PAPI) methods. If comparisons such as these are made, it is recommended that they be interpreted with caution.

When comparing 2000 and 2001 prevalence estimates, one can test the null hypothesis (no difference in the 2000 and 2001 prevalence rates) against the alternative hypothesis (there is a difference in prevalence rates) using the standard difference in proportions test expressed as

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\text{var}(\hat{p}_1) + \text{var}(\hat{p}_2) - 2\,\text{cov}(\hat{p}_1, \hat{p}_2)}},$$

where

$\hat{p}_1$ = 2000 estimate,

$\hat{p}_2$ = 2001 estimate,

$\text{var}(\hat{p}_1)$ = variance of 2000 estimate,

$\text{var}(\hat{p}_2)$ = variance of 2001 estimate, and

$\text{cov}(\hat{p}_1, \hat{p}_2)$ = covariance between $\hat{p}_1$ and $\hat{p}_2$.

Under the null hypothesis, *Z* is asymptotically distributed as a normal random variable. Calculated values of *Z* can, therefore, be referred to as the unit normal distribution to determine the corresponding probability level (i.e., *p* value). Because there is a 50 percent overlap in the sampled segments between the 2000 and 2001 NHSDAs, the covariance term in the formula for *Z* will, in general, be greater than 0. SUDAAN is used to compute estimates of Z along with the

associated *p* values such that the covariance term is calculated by taking the sample design into account. A similar procedure and formula for *Z* are used for estimated totals.

Chi-square tests of independence were conducted to examine the effects of subgroup variables with more than two levels on a prevalence measure to determine whether significant differences existed between the subgroup and prevalence variables. Log-linear chi-square testing is done in order to control the error level for multiple comparisons. Tests of the significance of particular subgroups were only performed if the chi-square test indicated that there were overall significant differences. SUDAAN analytic procedures were used in all tests to properly account for the sample design. A detailed description of the test statistic, which is based on the Wald statistic, can be found in the SUDAAN user's manual (RTI, 2001, pp. 317-319).

5. Confidence Intervals

In some NHSDA publications, sampling error was quantified using 95 percent confidence intervals. Because the estimates in the NHSDA are frequently small percentages, the confidence intervals are based on logit transformations. Logit transformations yield asymmetric interval boundaries that are more balanced with respect to the probability that the true value falls below or above the interval boundaries than is the case for standard symmetric confidence intervals for small proportions.

To illustrate the method, let the proportion P_d represent the true prevalence rate for a particular analysis domain "*d*." Then the logit transformation of P_d, commonly referred to as the "log odds," is defined as

$$L = \ln[P_d / (1 - P_d)],$$

where "1n" denotes the natural logarithm.

Letting $\hat{p}_d$ be the estimate of the proportion, the log odds estimate becomes $\hat{L} = \ln[\hat{p}_d / (1 - \hat{p}_d)]$. Then the lower and upper confidence limits of *L* are formed as

$$A = \hat{L} - K\left[\frac{\sqrt{\text{var}(\hat{p}_d)}}{\hat{p}_d(1 - \hat{p}_d)}\right]$$

$$B = \hat{L} + K\left[\frac{\sqrt{\text{var}(\hat{p}_d)}}{\hat{p}_d(1 - \hat{p}_d)}\right],$$

where $\text{var}(\hat{p}_d)$ is the variance estimate of $\hat{p}_d$, the quantity in brackets estimates the SE of $\hat{L}$, and *K* is the constant chosen to yield a level of confidence (e.g., *K* = 1.96 for 95 percent confidence limits).

Applying the inverse logit transformation to *A* and *B* above yields a confidence interval for $\hat{p}_d$ as follows:

$$\hat{p}_{d,\,lower} = \frac{1}{1 + \exp(-A)}$$

$$\hat{p}_{d,\,upper} = \frac{1}{1 + \exp(-B)} \,,$$

where "exp" denotes the inverse log transformation. The lower and upper confidence interval endpoints for percentage estimates are obtained by multiplying the lower and upper endpoints of $\hat{p}_d$ by 100.

The confidence interval for the estimated domain total, $\hat{Y}_d$ as estimated by $\hat{Y}_d = \hat{N}_d \cdot \hat{p}_d$, is obtained by multiplying the lower and upper limits of the proportion confidence interval by $\hat{N}_d$. For domain totals $\hat{Y}_d$, where $\hat{N}_d$ is not fixed, the confidence interval approximation assumes that the sampling variation in $\hat{N}_d$ is negligible relative to the error in $\hat{p}_d$.

6. Incidence Estimates

To assist in the evaluation of trends in the initiation of drug use, NHSDA data also were used to generate estimates of drug use incidence or initiation (i.e., the number of new users during a given year). The 2001 incidence estimates were generated using combined 1999, 2000, and 2001 data to further increase precision and add stability to the trends. Incidence rates measure the rapidity with which new drug users arise and can suggest emerging patterns of drug use.

The measure of incidence is defined as the number of new cases of drug initiation divided by the person time of exposure. For diseases, the incidence rate, *IR*, for a population is defined as the number of new cases of the disease, *N*, divided by the person time, *PT*, of exposure or

$$IR = \frac{N}{PT}.$$

The person time of exposure is measured as the net time that individuals in the population during an observed period of time are at risk of developing the disease. This period of time can be for the full period of the study or for a shorter period. The person time of exposure ends at the time of diagnosis (e.g., Greenberg, Daniels, Flanders, Eley, & Boring., 1996, pp. 16-19). Similar conventions were followed for defining the incidence of first use of a substance.

Begun in 1999 and continued through 2001, the NHSDA questionnaire allowed for the collection of year and month of first use for recent initiates. The month, day, and year of birth for the initiates were also obtained directly or imputed in the process. In addition, the questionnaire call record provided the date of the interview. By imputing a day of first use within the year and month of first use reported or imputed, the key respondent inputs, in terms of exact dates, can be computed. Exposure time can be determined in terms of days and converted to an annual value.

Having exact dates of birth and first use also allowed the determination of person time of exposure during the targeted period, *t*. Let the target time period for measuring incidence be specified in terms of dates. For the period 1998, for example, the specification would consist of

$$t = [t_1, t_2) = [1\ Jan\ 1998, 1\ Jan\ 1999),$$

a period that includes January 1, 1998, and all days up to but not including January 1, 1999. The target age group can also be defined by a half-open interval as $a = [a_1, a_2)$. For example, the age group 12 to 17 would be defined by $a = [12,18)$ for youths at least age 12, but not yet age 18.

If person *i* was in age group *a* during period *t*, the time and age interval, $L_{t,a,i}$, can then be determined by the intersection

$$L_{t,a,i} = [t_1, t_2) \cap \left[DOB_i\ MOB_i\ YOB_i + a_1, DOB_i\ MOB_i\ YOB_i + a_2\right) ,$$

where the time of birth is defined in terms of day (DOB_i), month (MOB_i), and year (YOB_i). Either this intersection will be empty ($L_{t,a,i} = \varnothing$), or it was designed by the half-open interval, $L_{t,a,i} = [M_{1,i}, M_{2,i})$, where

$$M_{1,i} = Max\left\{t_1, \left(DOB_i\ MOB_i\ YOB_i + a_1\right)\right\}$$

and

$$M_{2,i} = Min\left\{t_2, \left(DOB_i\ MOB_i\ YOB_i + a_2\right)\right\}.$$

The date of first use, $t_{fu,d,i,}$ is also expressed as an exact date. An incident of first drug *d* use by person *i* in age group *a* occurs in time *t* if $t_{fu,d,i} \in [M_{1,i}, M_{2,i})$. The indicator function, $I_{i(}d,a,t)$, used to count incidents of first use is set to 1 when $t_{fu,d,i} \in [M_{1,i}, M_{2,i})$, and to 0 otherwise. The person time exposure, measured in years and denoted by $e_i(d,a,t)$ for a person *i* of age group *a* depends on the date of first use. If the date of first use precedes the target period ($t_{fu,d,i} < M_{1,i}$), then $e_i(d,a,t) = 0$. If the date of first use occurs after the target period or if person *i* has never used drug *d*, then

$$e_i(d,a,t) = \frac{M_{2,i} - M_{1,i}}{365}.$$

If the date for first use occurs during the target period, $L_{t,a,i}$, then

$$e_i(d,a,t) = \frac{t_{fu,d,i} - M_{1,i}}{365}.$$

During leap years, the denominator used to compute person time exposure is set to 366. Note that both $I_i(d,a,t)$ and $e_i(d,a,t)$ are set to 0 if the target period, $L_{t,a,i}$, is empty (i.e., person *i* is not in age group *a* during time *t).* The incidence rate is then estimated as a weighted ratio estimate:

$$IR(d,a,t) = \frac{\sum_i w_i I_i(d,a,t)}{\sum_i w_i e_i(d,a,t)} \ ,$$

where the w_i are the analytic weights.

Before exact date data were available for computing incidence rates (i.e., before 1999), a person was considered to be of age *a* during the entire time interval *t* if his or her *a*th birthday occurred during time interval *t* (generally, a single year). If the person initiated use during the year, the person time exposure was approximated as one-half year for all such persons rather than computing it exactly for each person.

Because of the new methodology, all incidence rates from 1999 and forward are not strictly comparable with prior year estimates. However, because they are based on retrospective reports by survey respondents (as was the case for earlier estimates), they may be subject to some of the same kinds of biases.

Bias resulting from differential mortality occurs because some persons who were alive and exposed to the risk of first drug use in the historical periods shown in the tables died before the 2001 NHSDA was conducted. This type of bias is probably very small. Incidence estimates are also affected by memory errors, including recall decay (tendency to forget events occurring long ago) and forward telescoping (tendency to report that an event occurred more recently than it actually did). These memory errors would both tend to result in estimates for earlier years (i.e., 1960s and 1970s) that are downwardly biased (because of recall decay) and estimates for later years that are upwardly biased (because of telescoping). There is also likely to be some underreporting bias because of social acceptability of drug use behaviors and respondents' fears of disclosure. This is likely to have the greatest impact on recent estimates, which reflect more recent use and reporting by younger respondents.

Because the incidence estimates are based on retrospective reports of age at first use, the most recent year available for these estimates is 2000, based on the 2001 NHSDA. Estimates for the year 2000 are based only on data from the 2001 survey, estimates for the year 1999 are based only on data from the 2000 and 2001 surveys, and estimates for earlier years are based on the combined 1999 to 2001 data. For two of the measures, first alcohol use and first cigarette use,

initiation before age 12 is common. A 2-year lag in reporting for "all ages" estimates is applied for these measures because the NHSDA sample does not cover youths under age 12. The 2-year lag ensures that initiation at ages 10 and 11 is captured in the estimation.

7. Suppression of Estimates with Low Precision

Direct survey estimates, noted by an asterisk (*), are not reported as they are considered to be unreliable due to unacceptably large sampling errors. The criterion used for suppressing all direct survey estimates was based on the relative standard error (RSE), which is defined as the ratio of the standard error (SE) over the estimate.

For proportion estimates ($\hat{p}$) within the range $[0 < \hat{p} < 1]$, rates and corresponding estimated numbers of users were suppressed if

$$\text{RSE}\left[-\ln(\hat{p})\right] > 0.175 \text{ when } \hat{p} \le 0.5$$

or

$$\text{RSE}\left[-\ln(1-\hat{p})\right] > 0.175 \text{ when } \hat{p} > 0.5.$$

Based on a first-order Taylor series approximation to estimate RSE $[-\ln(\hat{p})]$, and RSE $[-\ln(1-\hat{p})]$ the following was used for computational purposes:

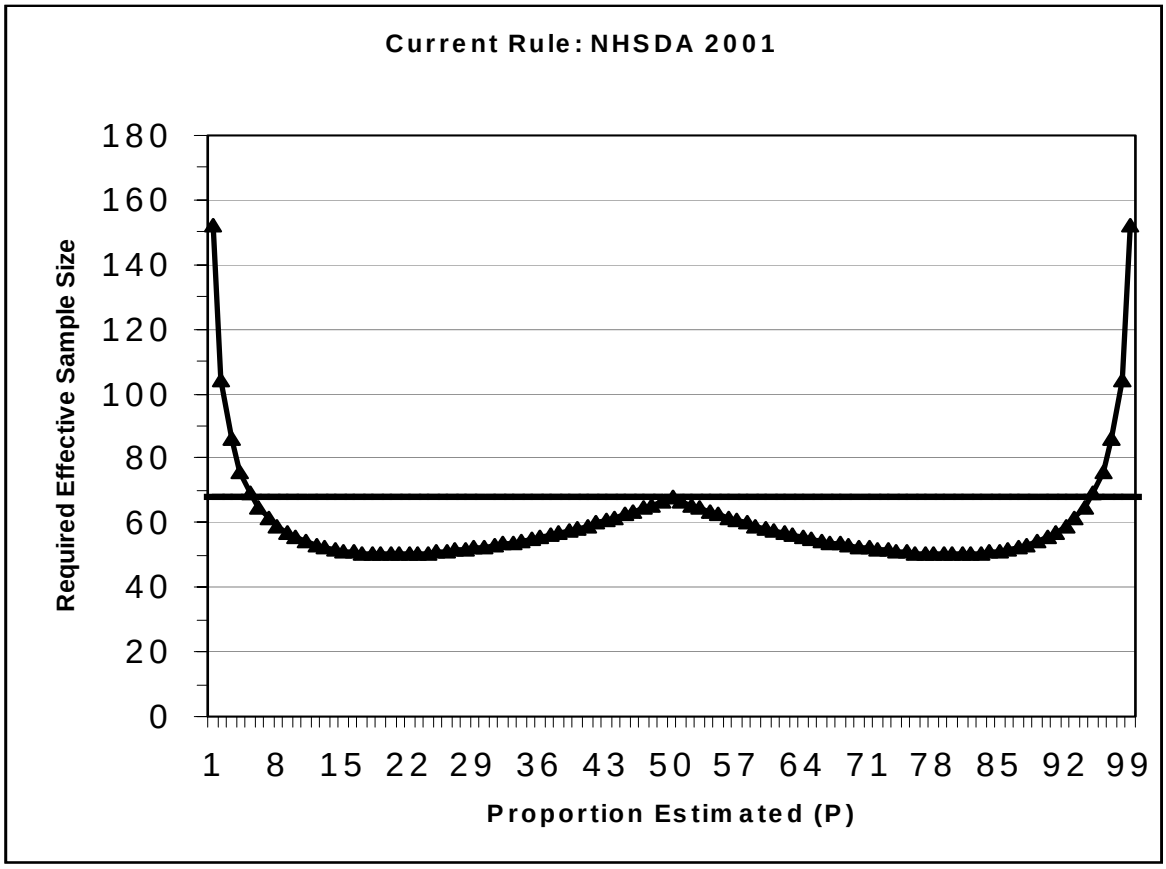
$$\frac{\text{SE}(\hat{p})/\hat{p}}{-\ln(\hat{p})} > 0.175 \text{ when } \hat{p} \le 0.5$$

or

$$\frac{SE(\hat{p})/(1-\hat{p})}{-\ln(1-\hat{p})} > 0.175 \text{ when } \hat{p} > 0.5.$$

The separate formulas for $\hat{p} \le 0.5$ and $\hat{p} > 0.5$ produce a symmetric suppression rule; that is, if $\hat{p}$ is suppressed, so will 1- $\hat{p}$. This is an ad hoc rule that requires an effective sample size in excess of 50. When $0.05 < \hat{p} < 0.95$, the symmetric properties of the rule produce a local maximum effective sample size of 68 at $\hat{p} = 0.5$. Thus, estimates with these values of $\hat{p}$ along with effective sample sizes falling below 68 are suppressed. A local minimum effective sample size of 50 occurs at $\hat{p} = 0.2$ and again at $\hat{p} = 0.8$ within this same interval; so, estimates are suppressed for values of $\hat{p}$ with effective sample sizes below 50. See Figure 1 for a graphical representation of the required minimum effective sample sizes as a function of the proportion estimated.

Figure 1. Required Effective Sample as a Function of the Proportion Estimated



A minimum nominal sample size suppression criteria (*n = 100*) that protects against unreliable estimates caused by small design effects and small nominal sample sizes was employed in 2001. Prevalence estimates are also suppressed if they are close to 0 or 100 percent (i.e., if $\hat{p} < .00005$ or if $\hat{p} \geq .99995$).

Estimates of other totals (e.g., number of initiates) along with means and rates (both not bounded between 0 and 1) are suppressed if

$$\text{RSE}\left(\hat{p}\right) > 0.5.$$

Additionally, estimates of mean age of first use were suppressed if the sample size was smaller than 10 respondents; also, the estimated incidence rate and number of initiates were suppressed if they rounded to 0.

The suppression criteria for various NHSDA estimates are summarized in Table 1.

Table 1. Summary of 2001 NHSDA Suppression Rules

Estimate	Suppress if:
Prevalence rate, $\hat{p}$, with nominal sample size, *n*, and design effect, *deff*	The estimated prevalence rate, $\hat{p}$, is less than 0.00005 or greater than or equal to 0.99995, or $\frac{\mathrm{SE}(\hat{p})/\hat{p}}{-\ln(\hat{p})} > 0.175$ when $p \le 0.5$, or $\frac{se(p)/(1-p)}{-\ln(1-p)} > 0.175$ when $\hat{p} > 0.5$, or Effective *n* < 68, or *n* < 100 where Effective *n* = $\frac{n}{deff}$. Note: The rounding portion of this suppression rule for prevalence rates will produce some estimates that round at one decimal place to 0.0 percent or 100.0 percent but are not suppressed from the tables.
Estimated number (numerator of $\hat{p}$)	The estimated prevalence rate, $\hat{p}$, is suppressed. Note: In some instances when $\hat{p}$ is not suppressed, the estimated number may appear as a 0 in the tables; this means that the estimate is greater than 0 but less than 500 (estimated numbers are shown in thousands).
Mean age at first use, $\bar{x}$, with nominal sample size, *n*	RSE $(\bar{x}) > 0.5$, or *n* < 10
Incidence rate, $\hat{r}$	Rounds to less than 0.1 per thousand person years of exposure, or RSE $(\hat{r}) > 0.5$
Number of initiates, $\hat{t}$	Rounds to fewer than 1,000 initiates, or RSE $(\hat{t}) > 0.5$

References

Bowman, K. R., Chromy, J. R., Odom, D. M., & Penne, M. A. (2003). Sample design report. In *2001 National Household Survey on Drug Abuse: Methodological resource book* (prepared for the Substance Abuse and Mental Health Services Administration, Contract No. 283-98-9008). Research Triangle Park, NC: RTI.

Chen, P., Gordek, H., Murtha, C., Singh, A., Westlake, M., & Yu, D. (2003). Sampling weight calibration for the 2001 NHSDA. In *2001 National Household Survey on Drug Abuse: Methodological resource book* (prepared for the Substance Abuse and Mental Health Services Administration, Contract No. 283-98-9008). Research Triangle Park, NC: RTI.

Davis, T. R., Packer, L. E., Heller, D. C., & Chromy, J. R. (2002). Statistical inference report. In *2000 National Household Survey on Drug Abuse: Methodological resource book* (prepared for the Substance Abuse and Mental Health Services Administration, Contract No. 283-98-9008). Research Triangle Park, NC: RTI.

Eyerman, J., & Bowman, K. (2002). *2001 National Household Survey on Drug Abuse: incentive experiment: Combined quarter 1 and quarter 2 analysis* (Prepared for the Substance Abuse and Mental Health Services Administration, Contract No. 283-98-9008). Research Triangle Park, NC: RTI.

Greenberg, R. S., Daniels, S. R., Flanders, W. D., Eley, J. W., & Boring, J. R. (1996). *Medical epidemiology*. Norwalk, CT: Appleton & Lange.

Office of Applied Studies. (2002). *Impact of September 11, 2001 events on substance use and mental health* (DHHS Publication No. SMA 02-3729, Analytic Series A-18) Rockville, MD: Substance Abuse and Mental Health Services Administration.

RTI. (2001). *SUDAAN user's manual: Release 8.0*. Research Triangle Park, NC: Author.