THE NEISS SAMPLE (DESIGN AND IMPLEMENTATION) 1997 to Present

Prepared for public release by:

Tom Schroeder
Kimberly Ault
Division of Hazard and Injury Data Systems
U.S. Consumer Product Safety Commission
June 2001

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

The Consumer Product Safety operates a data system known as the National Electronic Injury Surveillance System (NEISS). The NEISS is a probability sample of hospital emergency departments in the United States and its territories. Data collected from the NEISS sample is weighted based on the sample design to produce national estimates of the number of consumer products-related injuries treated in hospital emergency rooms. Additionally, the NEISS data provide a source for follow-up investigations of product-related injuries.

This report documents the changes in NEISS sample design since January 1997. For documentation of the NEISS sample prior to 1997, see the report "The NEISS Sample Design (Design and Implementation) from 1979 to 1996".

NEISS SAMPLE CHARACTERISTICS

January 1997 - December 1998

During 1996, under contract with Westat, Inc., the NEISS sample was updated to reflect changes in the universe of hospitals (Marker & Lo, 1996). A new sampling frame was constructed based on the most currently available (1995) listing of hospitals and emergency room visits purchased from the SMG Marketing Group of Chicago, IL. The sampling frame included hospitals with 6+ beds having an emergency department; excluded were psychiatric and penal institutions. The updated sample contains five strata, four based on size (the total number of emergency room visits reported by the hospital) and one stratum consisting of children's hospitals.

The hospital size strata were constructed as shown in Table 1, with a special stratum for children's hospitals. In selecting the sample, Westat used a resampling method that maximized the probability of retaining hospitals selected for the former 1991 sample. The method was an extension of the Keyfitz procedures for stratified simple random samples. Therefore, 76 of the previously selected hospitals were retained and 26 new hospitals were selected, for a total of 102 hospitals. Of the 76 hospitals retained, 70 were retained from the probability sample and six were retained from the children's sample. Data from these six

hospitals had been collected since January 1995 to increase case findings for children. One hospital was no longer in business when recruitment began in 1997, thus the in scope probability sample is comprised of 101 hospitals, instead of the initially planned 102.

Three hospitals have unique weights because the hospital merged with another hospital after the updated sampling frame was chosen. The specifics of the mergers are listed below:

- One hospital selected in the NEISS sample in stratum 2, merged with another hospital not selected in the NEISS sample in stratum 2. The hospital's administrative records are merged. The merged hospital is considered in the statistical sample as one hospital beginning August 1, 1995. In the 1997 sample, the hospital has a weight of 38.0552.

- Another hospital selected in the NEISS sample in stratum 2, merged with another hospital not selected in the NEISS sample in stratum 2. The merged hospital is considered in the statistical sample as one hospital beginning May 1998. In the 1997 sample, the hospital has a weight of 38.0552.

- One hospital selected in the NEISS sample in stratum 3, merged with another hospital not selected in the NEISS sample in stratum 3. The merged hospital is considered in the statistical sample as one hospital beginning July 1998. In the 1997 sample, the hospital has a weight of 37.6683.

January 1999 to present

Beginning in January 1999, NEISS weights have been adjusted annually to take into account changes in the most recent sampling frame of U.S. hospitals available. This new component of the basic NEISS weight is the ratio of total emergency department visits as listed on the updated sampling frame to the total emergency department visits as estimated from the NEISS sample. Table 2 presents the total emergency department visits (ERVs) from the 1998 sampling frame, the estimated ERVs from the NEISS sample for 1998, and the computed ratio adjustments for January 1, 1999 – December 1999. The section labeled “Adjustments for Changes in the Sampling Frame” discusses the ratio adjustments in more detail. The report “Updated NEISS Weights Using the 1998 SMG Hospital Frame” written by Tom Schroeder (1999) gives a complete description of this ratio adjustment procedure.

One hospital in the small stratum closed its emergency department on December 13, 1999. Thus as of December 14, 1999, the NEISS is considered to be a sample of 100 in-scope hospitals. Table 3 presents the characteristics of the NEISS sample for this period.

Table 4 presents the total emergency department visits (ERVs) from the 1999 sampling frame, the estimated ERVs from the NEISS sample for 1999, and the computed ratio adjustments for January – December 2000. The report "Updated NEISS Weights Using the 1999 SMG Hospital Frame" written by Tom Schroeder (2000) gives a complete description of this ratio adjustment procedure.

Table 5 presents the total emergency department visits (ERVs) from the 2000 sampling frame, the estimated ERVs from the NEISS sample for 2000, and the computed ratio adjustments for January – December 2001. The report "Updated NEISS Weights Using the 2000 SMG Hospital Frame" written by Kim Ault and Tom Schroeder (2001) gives a complete description of this ratio adjustment procedure.

TREND ANALYSIS OF NEISS DATA PRIOR TO 1997

One of the advantages of a long running data series such as the NEISS is the ability to track trends across time. Periodic updates to the sample, such as the one described in this report, can interfere with such analyses. One of the best ways to adjust for these updates is to have an overlap (bridge) during which data are collected from both the old and new samples. Between January and September of 1997, a nine-month overlap study was implemented as part of the sample update. Selecting a new sample from the same frame increases the variances in measures across time. When the new sample is selected from a different frame (as is the case for the NEISS update) there is a potential break in the time series. By comparing the estimates produced by the two samples and the two frames, it is possible to adjust old estimates (backcast) to be consistent with the new sample. Such an overlap can be expected to provide continuing useful information that more than compensates for the one time cost of the bridge sample.
To backcast existing time series for the period from the previous frame update to the current update requires a two-step process. The first step is to estimate the difference between estimates from the two samples for the overlap period. The ratio of total number of injuries estimated from the new sample and new weights,

divided by the estimate for the same period from the old sample and old weights, should be calculated. This is an estimate of change due to fluctuations in the number and type of product-related injuries at existing emergency rooms, the addition of new emergency rooms to the frame, and sampling variation from one sample to another. This measure of change could be computed separately by each stratum. However, given the small sample sizes in some strata, the procedure chosen was to compute an overall measure of change and apply it to all strata.

The second step is to adjust for this change. Since the previous NEISS sample design was selected based on 1985 ERV data and the sample frame (prior to 1997) was based on 1995 ERV data, this change represents fluctuations over a 10-year period. While the changes in number of emergency rooms and ERVs has probably been uneven over this period, a reasonable approximation is to assume that one-tenth of the change has occurred each year. Therefore, the adjustment procedure adopted was to add one-tenth of the estimated change to the 1986 estimates, two-tenths to the 1987 estimates, and so on up to nine-tenths of the change to the 1994 estimates. The entire change has been applied to the 1995-1996 estimates. For additional information on analyzing NEISS data over time, see "Trend Analysis of NEISS Data" (February 2000).

Overlap estimates were calculated for individual product codes as well as higher order groupings of product codes. The overall estimate during the overlap period form the old sample was 9,258,592 while the estimate from the new sample was 8,436,476. This represents an 8.9% decrease in the estimate from the new sample as compared to the estimate from the old sample. The difference in estimates fluctuated by product code or product grouping with some estimates increasing and others decreasing when comparing the estimates from the new sample to estimates from the old sample. Table 6 lists the estimates from the new and old sample at the overall level as well as at the product grouping level used in CPSC's annual report.

HOSPITAL SELECTION PROCEDURE AND HOSPTIAL PARTICIPATION

Westat, the contractor for the current NEISS sample redesign, provided CPSC with the sampling frame, a set of primary hospital sample selections and procedures for selecting alternate hospitals to be used as substitutes for hospitals unwilling to participate. In each recruitment process, CPSC recruiters make

repeated and intensive efforts to obtain the participation of every primary sample hospital selected. Only when all approaches to obtain cooperation from a primary hospital fail do the recruiters turn to a replacement hospital.

If the hospital being replaced is the primary selection, CPSC selects the alternate (replacement) hospital according to the procedures established by Westat. If the first alternate hospital refuses to participate, the second alternate is selected, etc., until cooperation is obtained with a replacement hospital. If the hospital to be replaced is not a primary selection, we use the opportunity to try again to recruit that hospital which is the primary sample selection. This has enabled us to obtain the cooperation of some primary hospitals, which initially had refused to participate. Over time, whenever hospitals remain in business but elect to drop out of the system, CPSC attempts to return to the original recruitment order: primary hospital, first alternate hospital, etc. Since the replacement hospitals have the same probabilities of selection as the primary hospitals with which they are associated, they also have the same statistical weights. All hospitals on the system at the time of one of the updates of the NEISS are considered primary hospital selections as of that time.

HOSPITAL WEIGHTS, 1997 to present

Basic Hospital Weights

The "basic" hospital weights used by NEISS are equal to the inverse of the probability of selection for the hospitals in each stratum. The inverse of the probability of selection is simply the total number of hospitals on the 1995-sampling frame divided by the total number of hospitals in the 1997 NEISS sample calculated at the stratum level. Adjustments to these basic weights are made for non-response and hospital mergers. Annual estimates of injuries are derived by summing the monthly estimates for all months of the year.

Adjustments for Non-Response

Shown in Table 7 are the number of in-scope hospitals for the sample, together with the number of hospitals that actually participated in the system, by stratum, for each year and month since January 1997. This table may be used to compute the monthly stratum non-response adjustment factor for the basic hospital weights.

Adjustments for Hospital Mergers

When two hospitals merge and are in different size classes, the probability of selection of the merged hospitals is found by the formula for the union of two events:

$$P = P_1 + P_2 - P_1 P_2$$

Where P_i = probability of selection of hospital i; i = 1, 2

Taking into account any non-response adjustment, the basic merged weight of hospital i is computed as:

$$Wgt_{hi} = \frac{1}{\left(\frac{n_{h1}}{N_{h1}}\frac{r_{h1}}{n'_{h1}}\right) + \left(\frac{n_{h2}}{N_{h2}}\frac{r_{h2}}{n'_{h2}}\right) - \left(\frac{n_{h1}}{N_{h1}}\frac{r_{h1}}{n'_{h1}}\right)\left(\frac{n_{h2}}{N_{h2}}\frac{r_{h2}}{n'_{h2}}\right)}$$

where:

$N_{h1(2)}$ = Number of hospitals in the NEISS sampling frame for stratum h1(2)

$n_{h1(2)}$ = Number of hospitals selected for the NEISS sample for stratum h1(2)

$r_{h1(2)}$ = Number of hospitals participating in the NEISS sample for stratum h1(2)for the time period

$n'_{h1(2)}$ = Number of in-scope hospitals in the NEISS sample for stratum h1(2)

When two hospitals merge and are in the same size class, the situation is more complex because the sampling in a size class is done without replacement. The sample size in a particular size class was a fixed number n, and the total number of hospitals in the size class was N for the original sample and frame. If S denotes the sample, H_1 and H_2 the hospitals, the three possibilities that lead to the retention of the merged hospital in the sample are:

A. $H_1 \in S$, $H_2 \notin S$ Hospital 1 is in the original sample, hospital 2 is not
B. $H_1 \notin S$, $H_2 \in S$ Hospital 2 is in the original sample, hospital 1 is not
C. $H_1 \in S$, $H_2 \in S$ Both hospitals are in the original sample

The probability of event A is:
P(A) = P(H_1 is selected on the first draw and H_2 is not selected on any draw)

$$= n\frac{1}{N}\frac{N-2}{N-1}\frac{N-3}{N-2}\cdots\frac{N-n+1}{N-n+2}\frac{N-n}{N-n+1} = \frac{n}{N}\frac{N-n}{N-1}$$

The leading multiplier n accounts for the possibility that H_1 may be selected on any of the n draws. The probability of event B, P(B), is the same as P(A) from symmetry.

The probability of event C is:

$$P(C) = 2\binom{n}{2}P(\mathrm{H}_1 \text{ is selected n the first draw and } \mathrm{H}_2 \text{ is selected on the 2nd})$$
$$= 2\binom{\mathrm{n}}{2}\frac{1}{N}\frac{1}{N-1}$$

The probability of inclusion of the merged hospital is then:

$$P(A)+P(B)+P(C) = \frac{n(2N-n-1)}{N(N-1)} = \frac{2n}{N} - \frac{n(n-1)}{N(N-1)}$$

Taking into account any non-response adjustment, the basic merged weight of hospital i is computed as:

$$Wgt_{hi} = \frac{1}{\frac{2n_h}{N_h} - \frac{n_h(n_h-1)}{N_h(N_h-1)}} \left(\frac{n'_h}{r_h} \right)$$

Where:

N_h = Number of hospitals in the NEISS sampling frame for stratum h

n_h = Number of hospitals selected from the NEISS sample for stratum h

r_h = Number of hospitals participating in the NEISS sample for stratum h for the time period

n'_h = Number of in-scope hospitals in the NEISS sample for stratum h

Adjustments for Changes in Sampling Frame

The hospital population does not remain static over time. Hospitals close, merge, and open as well as change in the volume of emergency department visits. In order to stabilize the NEISS estimates over time without taking a new NEISS sample and backcasting historical estimates, a ratio adjustment to the basic NEISS weight can occur. NEISS estimates the number of consumer product-related injuries treated in hospital emergency departments. A ratio adjustment takes advantage of knowledge about a highly correlated auxiliary variable, which in this case is the total number of emergency department visits. The total number of emergency department visits can be obtained by purchasing a complete SMG hospital database on a yearly or biyearly basis.

The "ratio adjustment" applied to the basic NEISS weight is the ratio of the known total number of emergency department visits in the population (from the frame) over the estimate of the total emergency department visits based on the sample of NEISS hospitals. For computing ratio adjustments, Westat has recommended combining the small and medium strata together and the large and very large strata together due to the relatively small number of NEISS hospitals in some of the larger strata. (Marker, et al, 1999)

Within each combined stratum, the ratio-adjusted weights, w^*_{hi}, are computed as:

$$w_{hi}^{*} = w_{hi}\left(\frac{ERV_{up,h^{*}}}{\sum_{h\in h^{*}}\sum_{i} w_{hi}^{'} erv_{up,i}}\right) = w_{hi} * R_{h^{*}} \qquad \textit{(Equation 1)}$$

where

w_{hi} = NEISS basic weight (adjusted for hospital mergers if necessary)

w'_{hi} = NEISS basic weight (adjusted for hospital mergers if necessary: see discussion below)

$ERV_{up,h^{*}}$ = Total ERVs on updated SMG file for combined stratum h*

$Erv_{up,i}$ = Number of ERVs from the updated SMG file for NEISS hospital i

$R_{h^{*}}$ = Ratio adjustment for combined stratum h*

Before applying *equation 1* to compute the ratio adjusted NEISS weights, three issues need to be resolved. Several NEISS hospitals have been replaced since the redesigned sample was selected. It was decided that the ERVs from the set of hospitals that are currently on the NEISS at the time of the update would be used in the denominator of *equation 1.*

With the possible exception of the three hospitals with merger weights in NEISS, the two basic NEISS weights, w_{hi} and w'_{hi}, are equal. A NEISS hospital is given a merged weight if that hospital has merged with another hospital and the NEISS coder cannot distinguish in which hospital emergency department an injury was treated. However, it is possible on the updated frame that the total ERVs from each of the merged hospitals are listed separately, and for estimating total ERVs the hospitals would not be considered merged. If this is the case then w_{hi} equals the merged weight and w'_{hi} equals the NEISS basic weight. The following are how each of the merged hospitals will be treated:

- For one of the merged hospitals in stratum 2: Of the two hospitals that merged, one eventually closed and one built a new facility. For computing ERVs, this will be treated as two separate hospitals on the updated frame, one with zero ERVs and the other with a specified number of ERVs. Thus $w_{hi} \neq w'_{hi}$.

- For the other merged hospital in stratum 2: Both of the merged hospitals have separate ERVs listed on the updated frame and both emergency rooms remain operational. For computing ERVs, this will be treated as two separate hospitals on the updated frame. Thus $w_{hi} \neq w'_{hi}$.

- For the merged hospital in stratum 3: Both emergency rooms remain operational. However, the updated sampling frame lists ERVs for only one of the hospitals. This will be treated as a merged hospital in computing ERV totals. Thus $w_{hi} = w'_{hi}$.

A ratio adjustment in its true form would only sum the total ERVs on the updated frame for the hospitals that where eligible for NEISS on the 1995 frame. However, summing in this manner would not account for the ERVs from any of the new emergency departments that opened in the interim and thus underestimate the total number of injuries. Through NEISS, CPSC wants to estimate the total number of product-related injuries treated at all eligible U.S. hospitals. This would include any new emergency departments that have opened since 1995. Thus, the total ERVs from the new emergency departments are added into their appropriate combined strata.

Final NEISS Weights

The final NEISS weight calculated each month and used for national estimates can consist of the following parts: basic weights, adjustments for non-response, adjustments for merged hospitals, and adjustments for changes in the sampling frame. The final weight (for all non-merged hospitals) can be written as:

$$NEISS_{wt} = \frac{(N_h * n'_h)R_h}{(n_h * r_h)} \qquad \textit{(Equation 2)}$$

where:

N_h = Number of hospitals in the 1995 sampling frame for stratum h
n_h = Number of hospitals selected for the NEISS sample for stratum h
n'_h = Number of in-scope hospitals in the NEISS sample for stratum h
r_h = Number of NEISS hospitals participating in stratum h for the given month
R_h = Ratio adjustment for combined stratum h

Table 8 shows the final NEISS weights for 1997 to 2000.

NATIONAL ESTIMATES OF PRODUCT-RELATED INJURIES FROM NEISS

National estimates for a given month of NEISS are calculated using the following formula

$$Estimate = \sum_i wgt_i x_i \qquad \textit{(Equation 3)}$$

where:

wgt_i = Weight of hospital i for the month

x_i = Number of cases for a specified product or type of injury reported by hospital i for the given month

Except for the unique weights of merged hospitals, the weights of the hospitals are the same within a stratum and *equation 3* can be written as:

$$Estimate = \sum_{h=1}^{m} \sum_{i=1}^{r_h} \left(\frac{N_h}{n_h} \frac{n'_h}{r_h} \right) R_{h*} x_{hi} \qquad \textit{(Equation 4)}$$

where:

m = Number of strata in the NEISS sample during the given time period

N_h = Number of hospitals in the NEISS sampling frame for stratum h

n_h = Number of hospitals selected for the NEISS sample for stratum h

n'_h = Number of in-scope hospitals in the NEISS sample for stratum h

r_h = Number of NEISS hospitals participating for stratum h for the given month

R_{h*} = Ratio adjustment factor for stratum h for the given month

x_{hi} = Number of cases for a specified product or type of injury reported by hospital i in stratum h for the given month

Note that N_h/n_h , the reciprocal of the probability of selection of a hospital in stratum h for the (current) sample, is the basic weight associated with each hospital in stratum h.

The factor n'_h/r_h is used to adjust each hospital in a given stratum for the lack of participation of one or more hospitals in the stratum, when necessary. If all hospitals in the stratum participate during the given month, the non-response adjustment factor is one. Non-response adjustment factors can be computed for

each stratum of the NEISS since January 1997 by using Table 8. For a given month and stratum, the non-response adjustment factor is obtained by dividing the number in the column labeled "SAMPLE" by the number in the column labeled "PARTICIPANTS." For example, the non-response adjustment factor for the small stratum for October 1999 is 48/47 or 1.021.

When the monthly non-response adjustment factor for the given stratum is multiplied by the basic hospital weight, the result is the adjusted basic NEISS weight for the small stratum for October 1999. The basic weight for the small stratum in October 1999 was 3179/48 or 66.23 (See Table 1). Therefore, the adjusted basic NEISS weight is 1.021 x 66.23, or 67.639.

The last adjustment is for changes in the sampling frame, so the finally NEISS weight for October 1999 is 67.639 x 1.0070 or 68.11.

Bias in a Ratio Adjusted NEISS Estimate

Ratio estimates in general are biased estimates. In practice, the bias is usually negligible and unimportant if the sample is of moderate size. As discussed in Cochran, the bias in each stratum of a stratified sample has an upper bound less than the standard error of the estimate multiplied by the coefficient of variation of the estimated total ERVs.

$$\left| \text{Bias in NEISS Estimate}_h \right| <= \text{SE}(\text{NEISS Estimate}_h) * \text{C.V.}(\text{Estimated ERV}_h)$$

where:

$$\text{C.V.}(\text{Estimated ERV}_h) <= \begin{cases} .10 \text{ in the combined small and medium strata} \\ .09 \text{ in the combined large and very large strata} \\ .11 \text{ in he children's stratum} \end{cases}$$

Bias in a stratified sample is additive, but due to the relatively small upper bounds of the bias in terms of the standard error of the estimate for each stratum, the bias in ratio adjusted NEISS estimates is considered negligible.

SAMPLING ERRORS ASSOCIATED WITH NEISS ESTIMATE

Calculating Variances of NEISS Estimates

Variances of NEISS estimates are calculated using the classical formula for the variance of a total from a stratified sample. With adjustments to NEISS weights made for non-response, hospital mergers, and ratio adjustments, the classical variance formula doesn't fit NEISS exactly. Other methods for approximating the variance are available such as Taylor series approximation or various replication (Jackknife, Balanced Repeated Replication) methods. Taylor series approximations and a 'leave one out' Jackknife approach available using SUDAAN software produce similar if not identical results to the classical variance formula used by Data Systems. Because of the similar results to other methods of calculating variances, Westat recommended that the classical variance formula should be used in calculating variances for annual NEISS estimates.

Because NEISS estimates are based on a sample of hospital emergency rooms rather than on a census of all hospital emergency rooms, they may differ somewhat from the figures that would have been obtained if product-related injuries had been obtained from all hospital emergency rooms in the U.S. Standard (or sampling) errors are measures of the sampling variability, that is, of the variations in the estimates that occur by chance because a sample rather than the entire set of emergency rooms is surveyed. Measures of sampling variation are frequently expressed as coefficients of variation (c.v.'s). The coefficients of variation are the standard errors divided by the estimates. The c.v. is a measure of the proportionate error due to sampling and the standard error is a measure of the absolute error.

The square of the standard error is referred to as the sampling variance. The variance of an estimate based on a sample can be calculated from the sample data, and this has been done for NEISS.

The estimates of variances for NEISS take into account the probabilities of selection, stratification, and weighting. The variance estimating formula currently used is:

$$\sigma_x^2 = \sum_{h=1}^{m} \frac{r_h}{r_h - 1} \sum_{i=1}^{r_h} \left(\frac{N_h}{n_h} \frac{n_h'}{r_h} \right)^2 (x_{hi} - \bar{x}_h)^2 = \sum_{h=1}^{m} \frac{r_h}{r_h - 1} \sum_{i=1}^{r_h} (wgt_{hi} x_{hi} - wgt\bar{x}_{hi})^2$$

m = Number of strata in the sample for the time period
r_h = the number of hospitals participating in stratum h for the time period
N_h = Number of hospitals in the NEISS sampling frame for stratum h
n_h = Number of hospitals selected for the (current) sample for stratum h
n'_h = Number of in-scope hospitals in the (current) sample or stratum h
x_{hi} = the number of injuries reported for the time period in the i-th hospital in stratum h
wgt_{hi} = the weight of hospital i in stratum h for the time period

$$\bar{x}_h = \sum_{i=1}^{r_h} \frac{x_{hi}}{r_h} \quad \text{and} \quad wgt\bar{x}_h = \sum_{i=1}^{r_h} \frac{wgt_{hi}\, x_{hi}}{r_h}$$

The equation above applies to estimates of injuries during any period -- monthly, quarterly, annually, etc. -- with x_{hi} interpreted as the number of injuries during that period. For periods greater than one month, the formula assumes that the sample size is constant over the period. When there have been variations in sample size, r_h is defined as the number of hospitals reporting during all, or the majority of the months in the period. A “majority n” (number of hospitals reporting a majority of the time) or “fractional n” (sum of number of hospital months divided by the number of months) can be used with little difference in the results between the two methods.

This formula actually slightly overstates the true sampling variance, because it does not take into account the effect of the secondary stratification factor, geography. Test calculations indicate that there are only slight differences between calculations using the formula above, and ones that consider the geographic substratification.

Appendix 1 contains sample SAS code for calculating variances associated with a particular NEISS estimate. Appendix 2 contains sample SAS /SUDAAN code for calculating variances associated with a particular NEISS estimate.

Generalized Sampling Errors

"Generalized sampling errors" are also produced for NEISS estimates. These smoothed values are derived from fitting a curve to all calculated sampling errors for a defined set. Generalized sampling errors are commonly used by U.S. Government statistical agencies such as the Bureau of the Census, the Bureau of Labor Statistics, and the National Center for Health Statistics to convey to the

public information about the precision of their estimates. One reason for their use is to reduce the cost of computing and presenting the large number of variances that would be required if each required a separate computation. Another reason is to produce more stable estimates of variances. Sampling errors, when estimated from sample data, have variances of their own; fitting a curve frequently reduces these errors.

To fit curves for generalized sampling errors, the relationship between the size of the estimate x_k for the k-th product group and the variance of that estimate, σ^2_{xk} is expressed by the formula:

$$\frac{\sigma_{x_k}}{x_k} = C.V._{x_k} = \frac{1}{a + (b * LN(x_k))}$$

where C.V. is the coefficient of variation, "a" and "b" are estimated by an iterative procedure. The iterative procedure produces estimates of "a" and "b" which minimize the expression:

$$\sum_k x_k \left(C.V._{x_k} - \frac{1}{a + (b * LN(x_k))} \right)^2$$

Recent approximate generalized standard errors and coefficients of variation for annual estimates of NEISS are presented in Table 9.

The standard errors derived from NEISS can be used in the following way: The sample estimate and its standard error enable one to construct confidence intervals, ranges that would include the average results of all possible samples with a known probability. For example, if all possible samples were selected and surveyed, and an estimate and its standard error calculated from each, then:

1. Approximately 95 percent of the time the interval from two standard errors below the estimate to two standard errors above the estimate would include the average result of all possible samples.

2. About 90 percent of similar confidence intervals using 1.6 standard errors would include the average results of all samples.

3. About 68 percent of confidence intervals using one standard error would include the average results of all possible samples.

Table 1

NEISS Sample Characteristics

January 1, 1997 through December 31, 1998

Stratum	Range of Total ERVs	Number of Hospitals in Universe	Total Sample	Out of Scope /1	In Scope Sample
1	1 - 16,830	3,179	48		48
2	16,831 - 28,150	1,059	14	1	13
3	28,151 - 41,130	674	9		9
4	41,131 +	426	23		23
Children's	Various	50	8		8
Total		**5,388**	**102**	**1**	**101**

1/ Out of scope hospitals included hospitals no longer in existence, hospitals without emergency departments, or hospital emergency departments no longer in business.

a) Hospitals no longer in existence:

One hospital selected in the small stratum was no longer in business as a hospital when the recruitment began in January 1997.

Table 2

Ratio Adjustments to NEISS Weights

January 1, 1999 through December 31, 1999

Stratum	1998 ERVs from Emergency Rooms on 1995 Frame	1998 ERVs from Emergency Rooms New to 1998 Frame	Total ERVs	Estimated ERVs from NEISS Sample	Ratio Adjustment
1	24,864,655	128,250	24,992,905	26,587,898	
2	23,919,044	118,767	24,037,811	22,104,280	1.006953
3	23,385,639	0	23,385,639	24,429,219	
4	24,027,969	77,000	24,104,969	25,480,560	0.951529
Children's	1,967,015	94,582	2,061,597	2,355,625	0.875180
Total	98,164,322	418,599	98,582,921	100,957,582	

Table 3

NEISS Sample Characteristics

January 1, 2000 to present

Stratum	Range of Total ERVs	Number of Hospitals in Universe	Total Sample	Out of Scope /1	In Scope Sample
1	1 - 16,830	3,179	48	1	47
2	16,831 - 28,150	1,059	14	1	13
3	28,151 - 41,130	674	9		9
4	41,131 +	426	23		23
Children's	Various	50	8		8
Total		**5,388**	**102**	**2**	**100**

1/ Out of scope hospitals included hospitals no longer in existence, hospitals without emergency departments, or hospital emergency departments no longer in business.

a) Hospitals no longer in existence:

One hospital selected in the small stratum was no longer in business as a hospital when the recruitment began in January 1997.

One hospital in the small stratum closed its emergency department on December 13, 1999.

Table 4

Ratio Adjustments to NEISS Weights

January 1, 2000 through December 31, 2000

Stratum	1999 ERVs from Emergency Rooms on 1995 Frame	1999 ERVs from Emergency Rooms New to 1999 Frame	Total ERVs	Estimated ERVs from NEISS Sample	Ratio Adjustment
1	24,895,553	177,975	25,073,528	26,076,475	
2	23,759,376	269,898	24,029,274	21,927,125	1.022898
3	23,253,312	38,516	23,291,828	24,265,726	
4	23,441,652	77,000	23,518,652	25,787,540	0.935213
Children's	1,984,668	269,222	2,253,890	2,381,013	0.946610
Total	97,334,561	832,611	98,167,172	100,437,879	

Table 5

Ratio Adjustments to NEISS Weights

January 1, 2001 through December 31, 2001

Stratum	2000 ERVs from Emergency Rooms on 1995 Frame	2000 ERVs from Emergency Rooms New to 2000 Frame	Total ERVs	Estimated ERVs from NEISS Sample	Ratio Adjustment
1	24,890,047	292,436	25,182,483	26,484,845	
2	23,810,540	328,161	24,138,701	22,397,879	1.008970
3	23,411,162	137,949	23,549,111	26,982,547	
4	23,573,023	244,811	23,817,834	24,756,286	0.915501
Children's	2,006,918	252,380	2,259,298	2,297,606	0.983327
Total	97,691,690	1,255,737	98,947,427	102,919,163	

Table 6

National Estimates of The Total Number Of Product Related Injuries

January 1 – September 30, 1997

Product Group	91 NEISS Hospital Sample			101 NEISS Hospital Sample			Percent Change
	Cases	Estimate	CV*	Cases	Estimate	CV*	
Overall	225,692	9,258,592	0.06	202,649	8,436,476	0.05	-8.9
1. Nursery Equip.	1,852	68,637	0.09	2,045	64,374	0.08	-6.2
2. Toys	2,824	113,863	0.08	2,908	105,631	0.06	-7.2
3. Sports & Rec.	79,868	3,331,480	0.07	72,914	3,072,952	0.05	-7.8
4. Home Entertain.	1,945	81,088	0.07	1,851	76,785	0.05	-5.3
5. Personal Use	10,933	446,386	0.07	10,443	410,458	0.06	-8
6. Packaging	6,733	269,522	0.07	6,053	247,401	0.06	-8.2
7. Yard & Garden	4,280	204,135	0.08	3,610	177,942	0.07	-12.8
8. Home Workshop	5,842	269,084	0.06	4,897	237,747	0.05	-11.6
9. Home Maintain.	2,375	98,597	0.07	2,107	88,240	0.06	-10.5
10. House Appliance.	2,658	107,349	0.07	2,474	98,003	0.06	-8.7
11. Heat/Cool/Vent	2,419	103,446	0.07	2,249	89,271	0.06	-13.7
12. Housewares	14,774	601,475	0.07	12,093	545,617	0.05	-9.3
13. Home Furnish	37,698	1,521,387	0.07	34,507	1,387,972	0.05	-8.8
14. Home Struct	63,478	2,561,409	0.07	55,094	2,280,250	0.06	-11
15. Miscellaneous	4,988	192,968	0.06	4,760	181,864	0.06	-5.8

* CV= coefficient of variation. The coefficient of variation is the standard deviation of the estimate expressed as a proportion of the estimate.

TABLE 7

Number of Hospitals in the NEISS Sample and Number of Participating Hospitals By Stratum For Each Month Starting With January 1997

		Small Stratum		Medium Stratum		Large Stratum		Verge Large		Children Stratum		Total	
Year	Month	Sample	Participants	Sample	Participants	Sample	Participants	Sample	Participants	Sample	Participants	Total Sample	Total Participants
1997	1	48	45	13	11	9	9	23	18	8	8	101	91
1997	2	48	45	13	11	9	8	23	19	8	8	101	91
1997	3	48	46	13	11	9	8	23	19	8	8	101	92
1997	4	48	47	13	12	9	9	23	20	8	8	101	96
1997	5	48	48	13	12	9	9	23	20	8	8	101	97
1997	6	48	48	13	12	9	9	23	20	8	8	101	97
1997	7	48	48	13	12	9	9	23	22	8	8	101	99
1997	8	48	48	13	13	9	9	23	22	8	8	101	100
1997	9	48	48	13	13	9	9	23	22	8	8	101	100
1997	10	48	48	13	13	9	8	23	22	8	8	101	99
1997	11	48	47	13	13	9	8	23	23	8	8	101	99
1997	12	48	47	13	13	9	8	23	23	8	8	101	99
1998	1	48	47	13	13	9	8	23	23	8	8	101	99
1998	2	48	47	13	13	9	8	23	23	8	8	101	99
1998	3	48	47	13	13	9	8	23	23	8	8	101	99
1998	4	48	47	13	13	9	8	23	23	8	8	101	99
1998	5	48	47	13	13	9	8	23	23	8	8	101	99
1998	6	48	47	13	13	9	8	23	23	8	8	101	99
1998	7	48	48	13	13	9	9	23	23	8	8	101	101
1998	8	48	48	13	13	9	9	23	22	8	8	101	100
1998	9	48	48	13	12	9	9	23	22	8	8	101	99
1998	10	48	48	13	13	9	9	23	22	8	8	101	100
1998	11	48	48	13	13	9	9	23	22	8	8	101	100
1998	12	48	48	13	13	9	9	23	22	8	8	101	100
1999	1	48	48	13	13	9	9	23	22	8	8	101	100
1999	2	48	48	13	13	9	9	23	22	8	8	101	100
1999	3	48	48	13	13	9	9	23	22	8	8	101	100

TABLE 7

Number of Hospitals in the NEISS Sample and Number of Participating Hospitals By Stratum For Each Month Starting With January 1997

		Small Stratum		Medium Stratum		Large Stratum		Verge Large		Children Stratum		Total	
Year	Month	Sample	Participants	Sample	Participants	Sample	Participants	Sample	Participants	Sample	Participants	Total Sample	Total Participants
1999	4	48	48	13	13	9	9	23	23	8	8	101	101
1999	5	48	48	13	13	9	9	23	23	8	8	101	101
1999	6	48	48	13	13	9	9	23	23	8	8	101	101
1999	7	48	48	13	13	9	9	23	23	8	8	101	101
1999	8	48	48	13	13	9	9	23	23	8	8	101	101
1999	9	48	48	13	12	9	9	23	23	8	8	101	100
1999	10	48	47	13	12	9	9	23	23	8	8	101	99
1999	11	48	47	13	12	9	8	23	23	8	8	101	98
1999	12	48	47	13	12	9	8	23	23	8	8	101	98
2000	1	47	46	13	12	9	8	23	23	8	8	100	97
2000	2	47	46	13	12	9	8	23	23	8	8	100	97
2000	3	47	46	13	12	9	9	23	23	8	8	100	98
2000	4	47	47	13	13	9	9	23	23	8	8	100	100
2000	5	47	47	13	13	9	9	23	23	8	8	100	100
2000	6	47	47	13	13	9	9	23	23	8	8	100	100
2000	7	47	47	13	13	9	9	23	23	8	8	100	100
2000	8	47	47	13	13	9	9	23	23	8	8	100	100
2000	9	47	47	13	12	9	9	23	23	8	8	100	99
2000	10	47	46	13	12	9	9	23	23	8	8	100	98
2000	11	47	46	13	13	9	8	23	23	8	7	100	97
2000	12	47	46	13	13	9	8	23	23	8	7	100	97

Table 8

NEISS Weights by Year, Month, and Stratum From January 1997 to Present

Year	Month	Small Stratum	Medium Stratum	Large Stratum	Very Large Stratum	Children's Stratum
1997	JANUARY	70.6444	89.3961	74.8889	23.6667	6.2500
1997	FEBRUARY	70.6444	89.3961	84.2500	22.4211	6.2500
1997	MARCH	69.1087	89.3961	84.2500	22.4211	6.2500
1997	APRIL	67.6383	81.9464	74.8889	21.3000	6.2500
1997	MAY	66.2292	81.9464	74.8889	21.3000	6.2500
1997	JUNE	66.2292	81.9464	74.8889	21.3000	6.2500
1997	JULY	66.2292	81.9464	74.8889	19.3636	6.2500
1997	AUGUST	66.2292	75.6429	74.8889	19.3636	6.2500
1997	SEPTEMBER	66.2292	75.6429	74.8889	19.3636	6.2500
1997	OCTOBER	66.2292	75.6429	84.2500	19.3636	6.2500
1997	NOVEMBER	67.6383	75.6429	84.2500	18.5217	6.2500
1997	DECEMBER	67.6383	75.6429	84.2500	18.5217	6.2500
1998	JANUARY	67.6383	75.6429	84.2500	18.5217	6.2500
1998	FEBRUARY	67.6383	75.6429	84.2500	18.5217	6.2500
1998	MARCH	67.6383	75.6429	84.2500	18.5217	6.2500
1998	APRIL	67.6383	75.6429	84.2500	18.5217	6.2500
1998	MAY	67.6383	75.6429	84.2500	18.5217	6.2500
1998	JUNE	67.6383	75.6429	84.2500	18.5217	6.2500
1998	JULY	66.2292	75.6429	74.8889	18.5217	6.2500
1998	AUGUST	66.2292	75.6429	74.8889	19.3636	6.2500
1998	SEPTEMBER	66.2292	75.6429	74.8889	19.3636	6.2500
1998	OCTOBER	66.2292	75.6429	74.8889	19.3636	6.2500
1998	NOVEMBER	66.2292	75.6429	74.8889	19.3636	6.2500
1998	DECEMBER	66.2292	75.6429	74.8889	19.3636	6.2500
1999	JANUARY	66.6897	76.1688	71.2589	18.4250	5.4699
1999	FEBRUARY	66.6897	76.1688	71.2589	18.4250	5.4699
1999	MARCH	66.6897	76.1688	71.2589	18.4250	5.4699
1999	APRIL	66.6897	76.1688	71.2589	17.6240	5.4699
1999	MAY	66.6897	76.1688	71.2589	17.6240	5.4699
1999	JUNE	66.6897	76.1688	71.2589	17.6240	5.4699
1999	JULY	66.6897	76.1688	71.2589	17.6240	5.4699
1999	AUGUST	66.6897	82.5162	71.2589	17.6240	5.4699
1999	SEPTEMBER	66.6897	82.5162	71.2589	17.6240	5.4699
1999	OCTOBER	68.1086	82.5162	80.1663	17.6240	5.4699
1999	NOVEMBER	68.1086	82.5162	80.1663	17.6240	5.4699
1999	DECEMBER	68.1086	82.5162	80.1663	17.6240	5.4699
2000	JANUARY	69.2184	83.8228	78.7917	17.3218	5.9163
2000	FEBRUARY	69.2184	83.8228	78.7917	17.3218	5.9163

Table 8

NEISS Weights by Year, Month, and Stratum From January 1997 to Present

Year	Month	Small Stratum	Medium Stratum	Large Stratum	Very Large Stratum	Children's Stratum
2000	MARCH	69.2184	83.8228	70.0371	17.3218	5.9163
2000	APRIL	67.7457	77.3749	70.0371	17.3218	5.9163
2000	MAY	67.7457	77.3749	70.0371	17.3218	5.9163
2000	JUNE	67.7457	77.3749	70.0371	17.3218	5.9163
2000	JULY	67.7457	77.3749	70.0371	17.3218	5.9163
2000	AUGUST	67.7457	77.3749	70.0371	17.3218	5.9163
2000	SEPTEMBER	67.7457	77.3749	70.0371	17.3218	5.9163
2000	OCTOBER	67.7457	77.3749	70.0371	17.3218	5.9163
2000	NOVEMBER	67.7457	77.3749	70.0371	17.3218	5.9163
2000	DECEMBER	67.7457	77.3749	70.0371	17.3218	5.9163

Table 9

Generalized Relative Sampling Errors for NEISS for Estimates of Various Size

Estimated Number of Injuries	Approximate Standard Error	Generalized Sampling Error	Approximate 68% Confidence Interval		Approximate 95% Confidence Interval	
1,200	260	0.22	940	1,460	680	1,720
5,000	750	0.15	4,250	5,750	3,500	6,500
10,000	1,300	0.13	8,700	11,300	7,400	12,600
25,000	2,750	0.11	22,250	27,750	19,500	30,500
50,000	4,500	0.09	45,500	54,500	41,000	59,000
75,000	6,750	0.09	68,250	81,750	61,500	88,500
100,000	8,000	0.08	92,000	108,000	84,000	116,000
125,000	10,000	0.08	115,000	135,000	105,000	145,000
150,000	12,000	0.08	138,000	162,000	126,000	174,000
175,000	14,000	0.08	161,000	189,000	147,000	203,000
200,000	16,000	0.08	184,000	216,000	168,000	232,000
300,000	21,000	0.07	279,000	321,000	258,000	342,000
400,000	28,000	0.07	372,000	428,000	344,000	456,000
500,000	35,000	0.07	465,000	535,000	430,000	570,000
600,000	42,000	0.07	558,000	642,000	516,000	684,000
700,000	49,000	0.07	651,000	749,000	602,000	798,000
1,500,000	90,000	0.06	1,410,000	1,590,000	1,320,000	1,680,000

APPENDIX 1

SAS CODE FOR CALCULATING VARIANCES

```
%include 'g:\users\epds\rdwrite\formats\stratum.fmt';

/*************************************/
/* STRATUM FORMATS                   */
/* 1990 Estimates - $strt90h.        */
/* 1991-1996 Estimates - $strt91h.   */
/* 1997-present Estimates - $strt97h. */
/*************************************/;

DATA IN ERROR(KEEP = HID WT STRATUM DUMMY); SET g.neiss;
   hid = put(hid,$subhosp.);
   stratum = put(hid,$strt97h.);
   dummy = 1;
   WT = WT/10000; *if wt has a format of 7.4 delete this line;
   if wt<1 then delete;
   if stratum in ('C','S','M','L','V') then output in;
   else output error;

PROC PRINT DATA = ERROR;
   TITLE 'Possible Error! Hospital with no stratum';

PROC SORT DATA = IN; BY HID DUMMY;

/*************************************************************/
/* SUM THE TOTAL WEIGHT AND TOTAL NUMBER OF CASES BY HOSPITAL */
/*************************************************************/
;

DATA TOTAL (KEEP = HID TOT_WT TOT_CNT DUMMY STRATUM);
   SET IN; BY HID DUMMY;
   TOT_WT + WT;
   TOT_CNT + 1;
   IF LAST.DUMMY THEN DO;
      OUTPUT;
      TOT_WT = 0;
      TOT_CNT = 0;
   END;
   RETAIN TOT_WT 0 TOT_CNT 0;

PROC SORT DATA = TOTAL; BY DUMMY STRATUM;

      DATA VARPROD (KEEP = ESTIMATE COUNT CV VAR); SET TOTAL;
```

```
         BY DUMMY STRATUM;
         COUNT + TOT_CNT;
         TWCNT + TOT_WT;
         SUMSQ = (TOT_WT)**2;
         TSUMSQ + SUMSQ;

         IF LAST.STRATUM THEN DO;
            IF STRATUM ='S' THEN N = 47;
            ELSE IF STRATUM = 'M' THEN N = 13;
            ELSE IF STRATUM = 'L' THEN N = 8;
            ELSE IF STRATUM = 'V' THEN N = 23;
            ELSE IF STRATUM = 'C' THEN N = 8;

            STR_VAR = (TSUMSQ-(TWCNT**2)/N)*N/(N-1);
            VAR+STR_VAR;
            TSUMSQ = 0;
            ESTIMATE+TWCNT;
            TWCNT=0;
         END;

         IF LAST.DUMMY THEN DO;
            SD = SQRT(VAR);
            CV = SD/ESTIMATE;
            OUTPUT VARPROD;
            ESTIMATE = 0;
            VAR = 0;
            COUNT = 0;
         END;
      RUN;

PROC PRINT DATA = VARPROD;
   TITLE 'VARIANCE ESTIMATES FOR NEISS';
   VAR ESTIMATE COUNT CV VAR;
   FORMAT ESTIMATE 7.0 CV 7.4 VAR 12.;
   SUM ESTIMATE COUNT VAR;
RUN;
```

APPENDIX 2

SUDAAN CODE FOR CALCULATING VARIANCES

Example to calculate estimates and coefficient of variations of NEISS estimates by gender

```
DATA ALL;
   INPUT HOSPITAL $CHAR3. STRAT $CHAR1.;

CARDS;
1 M
2 S
3 S
4 S
5 V
6 S
7 M
8 S
9 S
10 V
11 S
12 S
13 S
14 S
15 S
16 M
17 S
18 V
19 L
20 S
21 S
22 S
23 S
24 S
25 L
26 V
27 M
28 S
29 S
30 M
31 V
32 C
33 M
34 S
35 S
36 S
```

37 S
38 S
39 S
40 S
41 S
42 S
43 S
44 S
45 S
46 M
47 V
48 V
49 S
50 S
51 V
52 L
53 V
54 S
55 L
56 M
57 V
58 V
59 V
60 V
61 M
62 M
63 L
64 L
65 L
66 C
67 L
68 L
69 V
70 V
71 C
72 C
73 C
74 C
75 C
76 C
77 S
78 S
79 S
80 V
81 M
82 S
83 V
84 M

```
85 V
86 S
87 S
88 S
89 S
90 V
91 S
92 V
93 M
94 S
95 S
96 V
97 V
98 S
99 S
100 V
101 S
;

PROC SORT DATA=DATASET; BY STRAT PSU;
PROC SORT DATA=ALL; BY STRAT PSU;

/* CREATE AT LEAST ONE RECORD FOR EVERY HOSPITAL. */
/* DATASET MUST CONTAIN VARIABLES NAMED PSU AND STRAT. */

DATA ALL;
   MERGE DATASET (IN=A) ALL (IN=B); BY PSU STRAT;
   IF ^A AND B THEN IN_STUDY = 0;
   IF A THEN IN_STUDY = 1;
   OUTPUT;

PROC SORT; BY STRAT PSU;

DATA ALL;
   FORMAT STRATUM 1.;
   IF STRAT = 'S' THEN STRATUM = 1;
   IF STRAT = 'M' THEN STRATUM = 2;
   IF STRAT = 'L' THEN STRATUM = 3;
   IF STRAT = 'V' THEN STRATUM = 4;
   IF STRAT = 'C' THEN STRATUM = 5;
   IF STRAT = ' ' THEN STRATUM = .;

PROC SORT; BY STRATUM PSU;

/*************************************************/
/* DATA SET IS NOW SORTED BY STRATUM PSU         */
/* READY FOR USE IN SUDAAN                       */
/*************************************************/;
```

```
/* SUDAAN ESTIMATES OF TOTALS */;

DATA _NULL_;
   TITLE1 'VARIANCE ESTIMATES FOR GENDER';
   TITLE2 'SUDAAN: DESIGN = WR';

PROC DESCRIPT DATA = "ALL" FILETYPE = SAS DESIGN = WR;

   SUBPOPN IN_STUDY = 1;
   NEST STRATUM PSU;
   WEIGHT WT;

   SUBGROUP DISP;
   LEVELS 2;
   VAR NEISS;
   TABLES SEX;
   PRINT NSUM TOTAL SETOTAL / TOTALFMT=F10. SETOTALFMT = F10.
       STYLE = NCHS;
   OUTPUT NSUM TOTAL SETOTAL / TOTALFMT=F10. SETOTALFMT = F10.
       FILENAME = SUDAAN1;

DATA SUDAAN1; SET SUDAAN1;
   FORMAT CV 7.4;
   CV = SETOTAL/TOTAL;

PROC PRINT;
   VAR SEX TOTAL NSUM CV SETOTAL;

RUN;
```

REFERENCES

Ault, Kimberly & Schroeder, Tom, "Updated NEISS Weights Using 2000 SMG Hospital Frame," U.S. Consumer Product Safety Commission, May 12, 2001.

Cochran, W. G. *Sampling Techniques*, 3rd ed., Chapter 6, John Wiley, New York, 1977.

Kessler, Eileen & Schroeder, Tom, " National Electronic Injury Surveillance System (NEISS) Estimated Generalized Relative Sampling Errors", U.S. Consumer Product Safety Commission, September 1998.

Marker, David & Lo, Annie, "Update of the NEISS Sampling Frame and Sample", Westat, October 11, 1996.

Marker, David, Lo, Annie, Brick, Mike, & Davis, Bill, Westat, Inc. "Comparisons of National Estimates from Different Samples and Different Sampling Frames of the National Electronic Injury Surveillance System (NEISS)", Westat, January 25, 1999.

Schroeder, Tom, " National Electronic Injury Surveillance System (NEISS) Improving Estimated Generalized Relative Sampling Errors", U.S. Consumer Product Safety Commission, April 1996.

Schroeder, Tom, "Updated NEISS Weights Using 1998 SMG Hospital Frame," memorandum, U.S. Consumer Product Safety Commission, February 12, 1999.

Schroeder, Tom, "Updated NEISS Weights Using 1999 SMG Hospital Frame," memorandum, U.S. Consumer Product Safety Commission, January 19, 2000.

Schroeder, Tom, "Trend Analysis of NEISS Data", U.S. Consumer Product Safety Commission, February 2000.