Multiple Frame Sample Design and Estimation for the 2008 National Sample Survey of Registered Nurses – Theory, Implementation, and Assessment

Marshall Fritz[1], Ralph DiGaetano[2], Jim Green[2], and Jay Clark[2]
[1]Health Resources and Services Administration, HHS/HRSA/BHPr/OWPPM,
5600 Fishers Lane, Rm. 9A-55, Rockville MD 20857
[2]Westat, 1600 Research Blvd., Rockville MD, 20850

Abstract
The 2008 NSSRN is the latest in a series of national surveys of RNs carried out roughly every four years since the late 1970s. For the 2008 NSSRN a new sample design and weighting process were used. RNs were sampled independently from State listings of currently licensed RNs. Many RNs were licensed in several States, resulting in multiple chances of sample selection. Probabilistic matching and questionnaire responses were used to identify RNs appearing on multiple listings. Compositing factors based on the number of strata containing an RN were applied to nonresponse adjusted weights assigned to the sampled records of responding RNs. The nonresponse adjustment process was enhanced compared to prior surveys. This paper indicates the theory behind the estimation approach used, the processes of implementing the approach, and the corresponding impact on survey precision and accuracy.

Key words: Multiple frame surveys, probability matching, composite estimation, bias, variance, design effects

1. Overview of the NSSRN

The National Sample Survey of Registered Nurses (NSSRN) is generally considered the leading survey offering estimates representing detailed characteristics of the Registered Nurse (RN) population in the United States. Estimates of interest include demographic items (age, sex, race/ethnicity, etc.), initial and highest level of education, employment status and employment setting, and salary and earnings. After the first survey in 1977, the NSSRN has been administered every four years since 1980. The survey was originally designed to be carried out by mail but has evolved into a multi-mode survey.

This survey is undertaken by the Bureau of Health Professions within the Health Resources and Services Administration which is an Agency within the United States Department of Health and Human Services. The mission of the Bureau is to eliminate health disparities by assuring that there is an adequate supply of nursing personnel to meet the needs for health care among the population of the United States. Statistics forthcoming from this survey are important for assessing the supply and demand for registered nurses now and into the future. The "Findings" report for the 2008 NSSRN will be released in the latter portion of 2010.

2. The Original Sample Design and Sample Weighting Process of the NSSRN

The original sample design for the NSSRN was developed in the mid 1970s by Morris Hansen (1976). This innovative design aimed to meet a number of analytic objectives including a desire to: obtain both State and Nation estimates of reasonable precision; and compare estimates over time with the intent of carrying out the survey annually. Estimates of interest have included demographic items (age, sex, race/ethnicity, etc.), initial and highest level of education, employment status and employment setting, and salary and earnings. The survey was to be carried out by mail (the number of survey modes has expanded over time, now including mail, the Web, and telephone).

The chief target population at the State level has been of a compound nature: if an individual with a currently active RN license is employed as a nurse, the RN is considered a member of the population of the State in which s/he is employed; otherwise, the individual is considered as a member of the population of the State in which s/he resides. Thus, knowing the State or States in which an individual has a currently active RN license does not provide sufficient information to indicate the State target population to which an RN belongs.

The most effective way to sample RNs was determined to be the use of sample frames based on the listings of RNs with currently active licenses maintained by the various State boards of nursing. Since an RN can be licensed in multiple States, the overall probability of selection for each sampled RN had to be determined and direct unduplication of the sample frames was not feasible. In the mid 1970's the possible formats for State listings included computer tapes, paper listings, cards in file cabinets, and computer punch cards. Thus, sample selection often had to be carried out on site by survey or State staff. An additional issue to complicate unduplication is that most RNs are female, and changes in marital status could result in an RN being licensed under different names in different States.

Thus, a prime concern in the development of the sample design was to be able to accurately determine each sampled RN's chance of selection in light of the fact that a non-trivial percentage of the RN population would appear on multiple State sample frames where direct unduplication was not feasible. After much exploratory work, a nested alphabetic cluster design was determined to be the most suitable, satisfying design and estimation requirements and meeting analytic objectives while being operationally feasible to implement. This design had the following features:

1. A hierarchical arrangement of States based on population was established with the States assigned in priority order from smallest to biggest in terms of the number of licensed RNs.

2. 250 clusters of names were established in alphabetical order using State licensure data bases to help make the clusters of roughly equal size.

3. A systematic random sample of 40 of these clusters was selected.

4. These clusters were then partitioned into subclusters of roughly 1/16 of each full cluster to allow varying sampling rates to be established for each State. To achieve the targeted yield for a given State sample, all sampled clusters or a

randomly determined combination of clusters and/or subclusters could be included. Sample selection of all full clusters from a State would represent a 16 percent sample (.16=40/250) while sample selection of a 1/16 subcluster within each cluster for a given State would represent a one percent sample (.01= (40/250) x (1/16)).

5. The random selection of clusters was done in such a way that the sample from a larger State (based on a given set of clusters and/or subclusters) was always nested within the sampled clusters from all States of higher priority (i.e., all States smaller than the State in question).

Thus, for example, the sample for Wyoming (the highest priority State) might be selected using the full cluster name range for all 40 sampled clusters while that for California might use the 1/16 sampled subcluster randomly selected for use for all 40 sampled full clusters whenever a 1/16 cluster was to be sampled from a full cluster. One full cluster for Wyoming might range from "Johnson, B. to Lannom, S." while the corresponding 1/16 cluster for California might range from "Jones, T. to Kelly, A.". This approach served to unduplicate the sample as sampled RNs licensed in multiple States would always be sampled from the smallest, and thus the chance of selection for the NSSRN for each RN could be linked to a unique State, the smallest State in which an RN was licensed.

Another aspect of the design was to obtain from each responding RN a list of the States in which s/he was currently licensed as well as the name under which s/he was licensed so that all States of licensure for each RN could be identified. Rules were developed to determine whether to continue to include or to exclude a sampled RN depending on the highest priority State in which s/he appeared on the sampling frame and whether the corresponding name was sampled from the State.

Another issue to be addressed was the development of sample weights that permitted efficient estimation at both the national and State levels, appropriately reflecting an RN's chance of selection for the study as well as adjusting the weights of participating RNs to reflect those of non-participating RNs. A hierarchical weighting structure based on the priority ordering of States used for sample selection was developed. It involved a series of successive global poststratifications to State licensure totals in priority order but adjusting these totals to account for RNs licensed in a specified State and also in a higher priority State. This is best described through example.

The highest priority State was Wyoming. The weights of all eligible respondents and ineligibles (those individuals who did not meet the eligibility requirements of the survey but who still appeared on a State's list of currently active licenses) from Wyoming were to be adjusted to sum to the total number of licenses on the Wyoming sample frame. The second priority State was Alaska. The weights of all eligible respondents and ineligibles in Alaska except for those RNs who were also licensed in a State in a higher priority State than Alaska (only Wyoming in this case) were to be poststratified to the control total for Alaska. The control total for Alaska was the number of records on the sample frame for Alaska minus the sum of the weights for all eligible responding or ineligible RNs who were licensed in a higher priority State than Alaska in addition to Alaska (only Wyoming in this case). This process continued until all eligible responding and ineligible RNs were assigned weights.

3. Limitations of the Original Sample Design and Weighting Process

While the sample design and weighting approaches were quite innovative, addressing the many hurdles posed by the circumstances (e.g., 51 licensure files appearing in many modes) and analytic objectives for the survey, they did result in some limitations and drawbacks. While design effects were generally of a reasonable level of magnitude, estimates related to race/ethnicity sometimes experienced huge design effects (sometimes ranging from 20-40) due to the use of the alphabetic clustering (see, for example, the report "1984 National Sample Survey of Registered Nurses, Summary of Results", 1986). The initial implementation of the study in 1977 used only 24 sampled clusters, but beginning in 1980 the full set of 40 sampled clusters was employed. From this structure 20 replicate weights were formed, providing relatively few degrees of freedom for variance estimation purposes. Finally, few, if any, auxiliary variables were available on any file for nonresponse adjustment purposes and incorporating them into the weighting process was not feasible. As a result, adjustments for differential nonresponse were not included into the weighting process except by State. Common sources of differential nonresponse such as age and race/ethnicity were thus not reflected in the sample weights, raising concerns that bias could potentially arise in survey estimates.

This same sample design and weighting process was employed from 1980-2004. There were general concerns that the estimates of the numbers of minority RNs were low and of possible undercoverage of the youngest RNs. An attempt to improve the accuracy of minority estimates in the 2000 NSSRN was explored but without great success. An attempt was made to improve the accuracy of the estimates for the youngest RNs in the 2004 NSSRN, again with limited success.

4. A New Sample Design for the 2008 NSSRN

In light of the many changes over time in computing power, software availability, and statistical theory and practice, a new sample design was developed for the 2008 NSSRN. For a start, all 51 licensure data bases were stored electronically. The new sample design had a number of features:

1. Independent systematic random samples were selected from the various sample frames established.

2. Auxiliary variables, such as ZIP code and, often, age, were used from State licensure files for implicit stratification purposes through sorting prior to sample selection.

3. Probability matching was employed to help in the identification of sampled RNs who appeared on more than one State listing of RNs with currently active licenses.

4. Explicit stratification was undertaken to deal with two issues: RNs licensed in two or more neighbouring States and thus with an opportunity to work in either; and RNs who worked in a federal facility in a State where s/he was not licensed (and need not be because it was not a requirement in order to work in the federal facility). This was focused on three States—Alaska, Hawaii, and New Mexico—

where relatively high variation in sample weights had arisen in 2004, and this phenomenon seemed a likely source of some of that variation).

This new sample design was expected to result in increased precision for national level estimates due to both the elimination of clustering as a component of sampling variability and the increased stratification achieved by sorting within strata prior to sample selection. An evaluation of the impact of the sample design and weighting processes on the precision of national level estimates is undertaken later in this paper.

5. New Estimation Processes for the 2008 NSSRN

The new sample design provided more flexibility in terms of how sample weights and variance estimates could be developed. This provided an opportunity to help reduce the potential for bias arising from differences in characteristics among those subgroups where differential response rates were in evidence.

The files indicating those RNs with currently active licenses could contain different variables, depending on State requirements. Moreover, many variables that were included and might be of interest for nonresponse adjustment purposes had a great deal of missing data. All files contained names and addresses of the RNs. Forty four States also provided date of birth, which had very little missing data. Thus, response rates were examined by age groups. There was evidence of differential nonresponse by age with younger age groups responding at lower rates than older ones. Thus, lower response rates may have been a major contributor to the concern about low estimates of the number and percentage of the youngest RNs.

The new weighting process involved several steps:

1. **Poststratification of the eligible respondents and ineligibles to stratum totals by age group, where age was available for this purpose.** This served to both adjust for nonresponse by age and maintain consistency with State licensure totals (this type of poststratification is sometimes referred to as population-based nonresponse adjustment). For those States where age was not available for weight adjustment purposes, poststratification based on a single global adjustment to stratum level totals, similar to the approach used for weighting under the old sample design, was employed. This adjustment was undertaken at the "sampled record" level.

2. **Probability matching of State licensure data to help determine the number of different strata in which an RN appeared.** Information provided by responding RNs was incorporated into this process as well. The reciprocal of the number of strata served as a compositing factor, assigned to each RN record with a nonresponse adjusted weight. Thus, the nonresponse adjusted weight was multiplied by this compositing factor, resulting in a "final sampled RN record" weight.

3. **Producing a single weight at the "RN level".** For each eligible responding or ineligible RN sampled multiple times across strata, the values of the "final sampled RN record" weights were summed to obtain an "RN" weight. Thus, a

data base was established at the "RN record" level with each eligible responding or ineligible RN having an appropriate "RN" level weight.

4. **Establishing a final RN weight through raking.** The initial RN level weights were raked to licensure totals for the 50 States and D.C. to retain consistency with the total number of licenses appearing on the State files. The weights from the eligible responding RNs served as the final weights for analytic purposes.

The reciprocal of the number of strata containing an RN serving as the compositing factor for that RN as described in Step 2 is an application of the "generalized weight share method" developed by Lavallée (2007).

With the use of an increased number of adjustment factors to the weights to account for both differential nonresponse and multiple chances of selection, the variability of the sample weights was expected to increase somewhat compared to earlier NSSRNs. However, the elimination of clustering and use of stratification was expected to reduce variability, and the more refined approach to nonresponse adjustment of the weights was expected to help reduce bias.

The number of degrees of freedom for variance estimates was substantially increased, to 100 compared to the 20 available in prior NSSRNs. Thus, the reliability of estimates of variance will increase. An evaluation of the impact of the sample design and weighting processes on the potential for reduced bias in national level estimates is also undertaken in this paper.

Before the two evaluations are presented, we will discuss in more detail how we addressed the issue of appropriately reflecting the overall probability of selection of each participating RN since many RNs had multiple chances of selection for the NSSRN. An important component of this was the probability matching effort.

6. Addressing Multiple Chances of Selection in the Weighting Process

Many RNs appeared in more than one sample stratum, resulting in these RNs having multiple chances of selection. As mentioned earlier, this was addressed for the 2008 NSSRN by determining the number of strata in which an RN was found and then multiplying the reciprocal of this number by the weight associated with each RN sampled record characterized as responding or ineligible, an application of Lavallée's general weight sharing method. Two sources of information were used in making this determination: input from the questionnaire for each responding RN and the probability matching effort.

6.1 Input from the Responding RNs

Each RN was asked to supplement the information on the State or States of licensure if the information was incomplete. The available information was provided on the questionnaire for mailed or web-based questionnaires, most of the sample. Telephone interviews were undertaken in roughly 10-15 percent of cases, and a more direct approach was used in these cases.

6.2 Probability Matching Employed for the NSSRN

The information obtained from the questionnaire was not considered reliable enough to serve as the only source of information on where an RN was currently licensed. Moreover, an RN could appear on a licensure file as currently licensed when s/he was not, depending on the frequency of a State's updating procedures. Thus, a probability matching effort was undertaken to serve as an independent source of information on the identification of States where an RN was licensed.

Roughly, 55,000 sampled records were compared to over three million frame records for the 50 States and D.C. At the individual State level, the number of sampled records ranged from approximately 500 to 3,000, and the number of frame records ranged from about 7,000 to over 300,000 records.

Two different software packages were used for this probability matching effort. The proprietary package WesMatch, developed at Westat specifically to undertake large scale matching efforts, was used for the bulk of the matching between sample and frame records. However, in order to provide an independent evaluation of the rates of both false positives and false negatives among the matches identified by WesMatch, a version of AutoMatch was used to undertake a similar matching effort but limited to only three States. In this way, the matching strategy used with WesMatch could be further refined to help find additional matches accurately and avoid the inaccurate identification of a match. As mentioned earlier, name and address were available for all records and date of birth was provided by 43 States plus D.C. (44 files in all). Other variables that could help determine a match were home phone number, gender, date license received, date license last renewed, and Social Security Number. However, only about half of the States included any of these other variables, and missing data were common on these variables. The availability of date of birth was an important factor in improving the accuracy of the matches.

Another factor helping to improve the accuracy of the matches was the multistate Nurse Licensure Compact coordinated by the National Council of State Boards of Nursing. The Compact is essentially a large-scale reciprocity agreement between member States. In general, if an RN is licensed in one State, s/he may work in any other Compact State without formally applying for a license. As of the reference date for eligibility for the 2008 NSSRN, 22 States were Compact members. One feature of the Compact is that if an RN has obtained a license in a Compact State and is resident in a Compact State (not necessarily the State where s/he initially obtained a license), the RN is to be listed as having a currently active license among the States within the Compact only in the State of residence. This feature served to reduce the number of RNs with multiple chances of selection for the NSSRN. However, this process was not perfectly executed. We did examine the listings of all Compact States, and each had a small percentage of RNs with an address in a different Compact State. The matching effort also identified a very few sampled RNs who appeared on two Compact State listings.

6.3 Comparison of the Two Sources of Information

The information provided by an RN on her or his States of licensure was not necessarily fully consistent with that obtained through matching. For instance, the matching might have identified a second State in addition to the State listing from which an RN was sampled while the returned questionnaire for the RN did not or vice versa. To evaluate the accuracy of the information from the respondent and the matching effort, 21 categories were formed reflecting the number of States of licensure identified by the

respondent, the number identified through matching, and the extent of agreement between the two sources of information. Some of these categories represented records with no inconsistent information (for instance, roughly 75 percent of the sampled RN records were not associated with a second State, neither by the matching nor by the questionnaire responses).

Visual inspection was carried out for roughly 900 records in categories where matching and the respondent provided inconsistent information. From this inspection a rule was developed: if either matching or a questionnaire response identifies a State of licensure, consider it accurate—with one exception. Matches between records from Pennsylvania and those from Delaware were highly inaccurate. No particular reason could be identified for these inaccuracies although a contributing factor could be that Pennsylvania is one of the seven States where date of birth was not available for matching purposes.

7. Evaluating the Impact of the Changes to the Sample Design and Weighting Processes

7.1 Evaluating the Potential for Bias Reduction

To assess if there is evidence of a reduction in bias, a set of 2008 estimates developed from weights constructed in two separate ways were compared. One set of weights was simply the final "official" weights (where poststratification was used to achieve an age-based nonresponse adjustment when possible). The other set of weights was developed for this specific comparison. For these weights, nonresponse adjustment was undertaken using poststratification to State totals, a surrogate for the weighting process used for the NSSRN prior to 2008.

It should be noted that imputation for missing data was done for the first time in 2008. The bias evaluation included imputed data since it involved a set of comparisons using only 2008 data.

These two sets of estimates are highly correlated, so most all differences were expected to attain statistical significance. To assess whether a difference is of practical importance, we have computed a measure of what we have termed a "standardized difference". Specifically, this is the ratio of the difference of the two estimates (one based on the final weight, the other on the surrogate weight) to the standard error of the estimate based on the final weight. If this ratio has an absolute value of 3 or more, this is taken as evidence of an expected reduction in bias with practical significance. Note that estimates based on the final weight, where age was utilized as part of the nonresponse adjustment process where possible, increased the contribution of younger RNs to the sample estimates. The estimates, their standard errors, the difference in the estimates, and the corresponding standardized difference are provided in Table 1 for the estimated percentage of RNs associated with each category for a set of variables.

For illustration purposes, consider the category "diploma program" for the variable "Initial Nursing Education". The estimated percentage of RNs whose initial nursing education was a diploma program was 20.36 percent based on the final weight and 21.62 percent using the surrogate weight (where nonresponse adjustments did not attempt to account for differential response rates by age). The corresponding standard errors were 0.23 and 0.26, respectively. The difference between these estimates was -1.26, and the standardized difference was -5.43 (dividing -1.26 by 0.23, the standard error for the

estimate based on the final weight). Thus, using our proposed criteria, since the absolute value of the standardized difference exceeds 3, this is taken as evidence that a practical difference arises when nonresponse adjustments incorporate age. This is expected to reduce the bias associated with the estimate. Since initial nursing education is correlated with age, finding such a difference for categories within the ""Initial Nursing Education" variable was anticipated. The absolute value of the standardized difference for "baccalaureate degree" also exceeded 3. Thus, accounting for differential response rates between age groups in the nonresponse adjustment process resulted in a wider gap between the estimate of those RNs whose initial nursing education was a diploma program and that for those with a Baccalaureate degree than would have been expected using the previous approach to accounting for survey nonresponse in the weighting process.

Similarly, since the standardized difference is greater than 3 and positive, including age, where possible, in the nonresponse adjustment process somewhat increased the expected percentage of those RNs employed in nursing. No race/ethnicity category had an absolute value over 3 for the standardized difference measurements. "Graduation Year", highly correlated with age, had most categories with the absolute value of the standardized differences exceeding 3. These changes almost certainly represent a reduction in bias after taking into account age in the weighting process.

One employment setting, "hospital", had an absolute value of the standardized difference exceeding the "3" threshold, suggesting an important reduction in the bias in the estimate, and that more RNs are employed in hospitals than might have been thought if age had not been taken into account in the nonresponse adjustment process. For the variable "Highest Nursing Education", the results are basically the same as was found for "Initial Nursing Education".

Finally, we consider estimates reflecting the age distribution of RNs. Using the final weight in the development of estimates produces higher estimated percentages for all age groups under the age of 50 compared to those using the surrogate weight, which incorporated only a global, State level nonresponse adjustment. The absolute values of the "standardized difference" exceed 3 for all categories but the age groups 40-44 and 45-49. For instance, for RNs under the age of 25, the standardized difference was 4.21. Again, almost certainly, bias in these estimates has been reduced through the use of age in the nonresponse adjustment process for the weights.

7.2 Evaluation of the Effect on Sample Variability

To assess whether or not the changes in the sample design and estimation processes had an appreciable effect on sample variability, we compared the design effects for a set of 2008 NSSRN estimates to the design effects for the same estimates from the 2004 NSSRN. As mentioned earlier, imputation was carried out for an NSSRN for the first time in 2008. Thus, to help ensure comparability between estimates for the two surveys, 2008 estimates were developed using the 2008 data prior to imputation.

A design effect measures the extent to which the sample variation of a specified estimate resulting from a survey's sample design and weighting processes differs from the expected variation for the same estimate from a simple random sample of the same size. It is computed as a ratio of the estimated variance of a survey estimate to the corresponding variance of a simple random sample of the same size. The number of

respondents to the 2004 and 2008 NSSRNs were approximately 36,000 and 34,000, respectively.

Before discussing these estimates, it should be noted that some category descriptions did not match up precisely between 2004 and 2008 and standard errors were not computed for the same set of categories in one instance (age groups associated with those 65 or older—this grouping was omitted from the table so the sum of the estimated percentages are not intended to sum to 100 across the age groups in the table). It should also be noted that the estimates for Table 1 and Table 2 may differ for the same variables because, noted earlier, imputed data were used for Table 1 but not Table 2. Finally, the estimates for Not Reported can be somewhat different between the two surveys. This can have an effect on the estimated percentages for other categories of that same variable.

In Table 2 we provide a set of variables (all but one also appearing in Table 1, another added in Table 2 instead) for which estimates of the percentage distribution across categories have been computed for both the 2004 and 2008 NSSRNs. Corresponding standard errors and design effects (DEFFs) have been computed as well as the ratio of the 2008 design effect to the 2004 design effect. For instance, for the variable ”Initial Nursing Education” the estimated percentage of those with a diploma program was 25.21 for 2004 and 20.23 for 2008. The corresponding standard errors for 2004 and 2008 were 0.32 and 0.23, respectively. The corresponding DEFFs for 2004 and 2008 were 1.94 and 1.14, respectively, and the ratio of the 2008 DEFF to the 2004 DEFF was 0.59. For all but one category for”Initial Nursing Education”, the ratios were substantially less than 1. For Master’s Degrees the ratio was about 1. With an estimated percentage for those with Master’s Degrees of 0.51 percent, and with only 20 degrees of freedom, the estimated standard error may be relatively unstable. In looking through Table 2 we find that many, but not all, of the ratios exceeding one are associated with 2004 estimates of less than one percent.

For the variable “Employed in Nursing” the ratio of design effects for those employed as a nurse (and those not employed as a nurse, as they are identical by construction) is .76.

For the Race/Ethnicity categories associated with the four largest estimates in 2008 (White, non-Hispanic; Black, non-Hispanic; Asian, non-Hispanic; and Hispanic), the design effects in 2004 ranged from around 20 to about 37, so that the ratios of the design effects were all less than .1. Thus, eliminating the clustering by names as an element of the sample design resulted in an enormous reduction in the variability of the estimated percentages for these demographic groups, as expected.

For the variable “Employment Status in Current Year” the ratios were all under .8 except for the tiny estimate of those “Employed in Nursing, Full/Part-Time Unknown” where it was about 2. For “Employment Setting” the design effects were generally lower than .8 unless the 2004 DEFF was close to 1. When these DEFFs were 1.18 or lower, the ratios were between 1.2 and 1.6. These 2004 DEFFs seem surprisingly low compared to those for other 2004 estimates, suggesting that the relatively low number of degrees of freedom might be a contributing factor to a relatively low estimated standard error. This same phenomenon was observed for estimated percentages associated with the variable “Highest Nursing Education”. Generally, the ratios are under .8, but for those where the ratios are greater than 1, the 2004 DEFFs are quite low. One was as low as .36 (Doctorate in Nursing or related Field) and one at .73 (Baccalaureate in Related Field).

Finally, for the categories associated with the variable "Age of Nurse", again the ratios were most all less than .8. However, there was one very large ratio (3.62) with a known explanation. As mentioned earlier, there has been general concern since the inception of the NSSRN that the very youngest RNs may be undercovered by the sample frame. As part of the 2004 NSSRN, a special poststratification effort was undertaken to attempt to improve the estimate of those under the age of 26. Control totals were derived based on examination results for that age group, and the weights of responding RNs under the age of 26 were adjusted to be consistent with these control totals. This poststratification served to reduce the variance of the 2004 estimated percentage of those under the age of 25, and the 2004 DEFF for this estimated percentage was only .43. Thus, the ratio of the 2008 DEFF to that of 2004 was quite large.

The impact of the 2004 poststratification for the youngest RNs on the 2004 DEFF for the estimated percentage of RNs in the under 25 age group was far more dramatic than that of the poststratification to age groups used for nonresponse adjustment purposes in 2008 on corresponding estimated percentages. There are several reasons for this. First, of all, the 2008 poststratification took place earlier in the weighting process, so its overall impact was reduced following the other weight adjustments. The later adjustments included use of the compositing factor and the raking at the end of the weighting process which achieved consistency only with State level totals, not by age group—due to issues with achieving convergence—and resulted in some additional variation in the weights after the poststratification to age group totals for most States. In addition, the poststratification cells for age used for the nonresponse adjustment for the 2008 weights did not correspond exactly with the cells presented in this table, and there were 7 States where date of birth was not provided on the sample frame, so these adjustments did not fully remove the variability associated with these estimated percentages for age group.)

8. Discussion

The original sample design and weighting process for the NSSRN were developed in the mid 1970's to address specific needs and constraints that existed at the time. With the many technological and statistical advances over time, the implementation of a new sample design and weighting process, developed to take advantages of these advances, was expected to result in improved national estimates. Specifically, it was anticipated that reductions in both sampling variability and the potential for bias could be attained. The evaluations undertaken and described in this paper provide evidence that, for national level estimates, these expected gains were realized.

State estimates are also of importance to users of NSSRN data. The variation of the weights at the State level may be a more sizeable source of sample variation than under the previous design and should be examined. Of course, some of this variation was introduced to help reduce the potential for bias through the incorporation of age into the nonresponse adjustment process, and such added variation appears well worth the trade-off by helping to reduce bias. Oversampling was also undertaken to attempt to increase the precision of estimates, such as the number of RNs employed in a State, by increasing sampling rates of those licensed in a given State as well as a neighbouring State.

References

Department of Health and Human Services, *The Registered Nurse Population, March 2008: Findings from the National Sample Survey of Registered Nurses.* U.S. Department of Health & and Human Services, Health Resources and Service Administration, Bureau of Health Professions, National Center for Workforce Analysis

Westat, American Nurses' Association. "Sample Design for the National Survey of Registered Nurses," Volume 1, Technical Report. Submitted to: Department of Health, Education and Welfare, Health Resources Administration, Division of Nursing, Bethesda, Maryland 20014. Contract No. 23175-0815. June 30, 1976.

Westat, "1984 National Sample Survey of Registered Nurses, Summary of Results." Prepared for: Health Resources and Services Administration, Division of Nursing, Bureau of Health Professions, Department of Health and Human Services, Rockville, Maryland 20857. Contract No. 240-83-0090. April, 1986.

Lavallée, P. "Indirect Sampling." Springer, New York (2007).

Table 1: Evaluating Potential for Reduction in Bias: Comparing Final 2008 Estimates with 2008 Estimates Using a Weight that does not Account for Differential Response Rates by Age

	Final Weight				
Description	*Estimated Percentage (Final weight)*	*S.E. of Estimated Percentage*	*Estimated Percentage (Surrogate weight)*	*Difference between Estimates*	*Standardized Difference (Ratio of difference to S.E.)*
Initial Nursing Education					
Diploma program	20.36	0.23	21.62	-1.26	-5.43
Associate degree	45.34	0.33	45.18	0.16	0.48
Baccalaureate degree	33.63	0.33	32.54	1.09	3.32
Master's degree	0.44	0.05	0.42	0.02	0.42
Doctorate	0.03	0.01	0.03	0.00	0.00
Not Reported	0.21	0.03	0.21	0.00	0.00
Employed in Nursing					
Yes	84.77	0.23	84.06	0.71	3.05
No	15.23	0.23	15.94	-0.71	-3.05
Racial/Ethnic background				0.00	
White (non-Hispanic)	83.22	0.25	83.74	-0.52	-2.08
Black/African American (non-Hispanic)	5.40	0.16	5.37	0.03	0.19
Asian (non-Hispanic)	5.53	0.16	5.28	0.25	1.54
American Indian/Alaskan native (non-Hispanic)	0.28	0.03	0.27	0.01	0.33
Native Hawaiian/Pacific Islander (non-Hispanic)	0.31	0.04	0.30	0.01	0.26
Two or more races (non-Hispanic)	1.68	0.09	1.67	0.01	0.12
Hispanic/Latino (white)	3.14	0.13	2.97	0.17	1.27

Description	*Final Weight*		*Estimated Percentage (Surrogate weight)*	*Difference between Estimates*	*Standardized Difference (Ratio of difference to S.E.)*
	Estimated Percentage (Final weight)	*S.E. of Estimated Percentage*			
Hispanic/Latino (Black/African American)	0.09	0.02	0.08	0.01	0.45
Hispanic/Latino (Two or more races)	0.14	0.03	0.13	0.01	0.37
Hispanic, Other	0.20	0.03	0.19	0.01	0.31
Graduation Year					
Before 1961	2.79	0.11	3.08	-0.29	-2.64
1961-1965	3.49	0.11	3.85	-0.36	-3.33
1966-1970	5.28	0.14	5.84	-0.56	-3.97
1971-1975	8.42	0.18	9.32	-0.90	-4.95
1976-1980	11.41	0.21	12.48	-1.07	-5.12
1981-1985	11.84	0.20	12.24	-0.40	-2.00
1986-1990	10.39	0.20	10.52	-0.13	-0.64
1991-1995	13.72	0.22	13.33	0.39	1.74
1996-2000	12.96	0.23	12.00	0.96	4.27
2000-2004	9.46	0.20	8.41	1.05	5.17
After 2004	10.24	0.20	8.92	1.32	6.60
Employment Setting, if employed in nursing					
Hospital	52.29	0.31	50.90	1.39	4.44
Nursing Home Extended	4.42	0.14	4.52	-0.10	-0.73
Nursing Education	3.21	0.11	3.32	-0.11	-0.99
Public Health/Community Health	8.58	0.19	8.77	-0.19	-1.00
School Health Service	2.76	0.10	2.87	-0.11	-1.12
Occupation Health	0.62	0.05	0.63	-0.01	-0.20
Ambulatory Care	8.83	0.18	8.85	-0.02	-0.11
Insurance Claims/Benefits	1.61	0.08	1.66	-0.05	-0.61
Other	1.70	0.08	1.76	-0.06	-0.74
Not reported	0.75	0.06	0.76	-0.01	-0.16
Not employed in nursing	15.23	0.23	15.94	-0.71	-3.05
Highest Nursing Education					
Diploma in Nursing	13.92	0.22	14.73	-0.81	-3.65
Associate Degree in Nursing or related field	36.02	0.33	35.64	0.38	1.14
Baccalaureate in Nursing (or related field for 2008)	34.87	0.31	33.87	1.00	3.25
Baccalaureate in related field	1.86	0.09	2.00	-0.14	-1.63
Masters in Nursing (or related field 2008)	9.47	0.19	9.66	-0.19	-0.98
Masters in related field	2.80	0.11	2.98	-0.18	-1.59
Doctorate in Nursing (or related field 2008)	0.43	0.04	0.46	-0.03	-0.71
Doctorate in related field	0.50	0.04	0.53	-0.03	-0.70
Not reported	0.13	0.03	0.14	-0.01	-0.40

Description	Final Weight: Estimated Percentage (Final weight)	Final Weight: S.E. of Estimated Percentage	Estimated Percentage (Surrogate weight)	Difference between Estimates	Standardized Difference (Ratio of difference to S.E.)
Age of Nurse					
<25	2.60	0.11	2.15	0.45	4.21
25 to 29	6.81	0.14	5.65	1.16	8.47
30 to 34	9.21	0.13	7.93	1.28	10.00
35 to 39	10.84	0.14	9.50	1.34	9.57
40 to 44	11.43	0.22	11.14	0.29	1.34
45 to 49	14.42	0.20	14.12	0.30	1.51
50 to 54	16.19	0.24	17.97	-1.78	-7.54
55 to 59	12.96	0.21	14.34	-1.38	-6.70
60 to 64	8.15	0.16	9.02	-0.87	-5.47
65-69	4.58	0.12	5.06	-0.48	-4.00
70 to 74	1.75	0.08	1.94	-0.19	-2.29
75+	1.07	0.06	1.18	-0.11	-1.72

Table 2: Evaluating Changes in Precision: Comparing 2004 Estimates with 2008 Estimates Prior to Imputation (using 2008 Final Weight)

Description	2004: Estimated Percent	2004: S.E. of Estimated Percent	2004: DEFF	2008: Estimated Percent	2008: S.E. of Estimated Percent	2008: DEFF	Ratio of 2008 to 2004 DEFFs
Initial Nursing Education							
Diploma program	25.21	0.32	1.94	20.23	0.23	1.14	0.59
Associate degree	42.18	0.54	4.27	45.18	0.33	1.50	0.35
Baccalaureate degree	30.49	0.47	3.72	33.51	0.33	1.65	0.44
Master's degree	0.51	0.05	1.74	0.43	0.05	1.81	1.04
Doctorate	0.02	0.01	1.95	0.03	0.01	1.35	0.69
Not reported	1.58	0.09	1.86	0.61	0.05	1.17	0.63
Employed in Nursing							
Yes	83.23	0.27	1.87	84.77	0.23	1.41	0.76
No	16.77	0.27	1.87	15.23	0.23	1.41	0.76
Racial/Ethnic background							
White (non-Hispanic)	81.82	0.89	19.03	81.38	0.26	1.44	0.08
Black/African American (non-Hispanic)	4.21	0.57	28.78	5.20	0.15	1.61	0.06
Asian (non-Hispanic)	2.90	0.54	36.99	5.31	0.16	1.77	0.05
American Indian/Alaskan Native (non-Hispanic)	0.32	0.03	0.99	0.27	0.03	1.05	1.06
Native Hawaiian/Pacific Islander (non-Hispanic)	0.19	0.04	2.98	0.29	0.04	1.68	0.56
Two or more races (non- Hispanic)	1.42	0.09	2.07	1.63	0.09	1.55	0.75
Hispanic/Latino (White)	1.32	0.27	19.93	2.10	0.10	1.50	0.08
Hispanic/Latino (Black/African American)	0.10	0.02	1.42	0.05	0.02	1.94	1.36
Hispanic/Latino (Two or more races)	0.11	0.03	3.02	0.09	0.02	1.97	0.65
Hispanic, Other	0.12	0.03	2.71	0.20	0.03	1.72	0.64
Not reported	7.48	0.19	1.86	3.48	0.13	1.64	0.88

Description	2004 Estimated Percent	2004 S.E. of Estimated Percent	2004 DEFF	2008 Estimated Percent	2008 S.E. of Estimated Percent	2008 DEFF	Ratio of 2008 to 2004 DEFFs
Employment Status in 2004							
Employed in Nursing Full Time	58.32	0.44	2.85	59.89	0.33	1.48	0.52
Employed in Nursing Part-Time	24.76	0.35	2.35	19.88	0.27	1.48	0.63
Employed in Nursing, Full/Part-Time Unknown	0.15	0.02	0.98	5.00	0.17	1.95	1.99
Not employed in nursing	16.77	0.27	1.87	15.23	0.23	1.41	0.76
Employment Setting							
Hospital	46.77	0.43	2.65	52.29	0.31	1.32	0.50
Nursing Home Extended	5.26	0.12	1.03	4.42	0.14	1.49	1.45
Nursing Education	2.18	0.10	1.67	3.21	0.11	1.33	0.79
Public Health/Community Health	8.93	0.15	0.99	8.58	0.19	1.54	1.56
School Health Service	2.68	0.10	1.37	2.76	0.10	1.20	0.88
Occupation Health	0.77	0.06	1.68	0.62	0.05	1.31	0.78
Ambulatory Care	9.55	0.22	2.00	8.83	0.18	1.40	0.70
Insurance Claims/Benefits	1.50	0.07	1.18	1.61	0.08	1.42	1.20
Other	3.85	0.16	2.47	1.70	0.08	1.32	0.53
Not reported	18.49	0.29	1.99	15.98	0.24	1.49	0.75
Highest Nursing Education							
Diploma in Nursing	17.54	0.27	1.80	13.83	0.22	1.39	0.77
Associate Degree in Nursing or related field	33.73	0.49	3.84	35.85	0.33	1.62	0.42
Baccalaureate in Nursing (or related field 2008)	31.71	0.45	3.34	34.69	0.31	1.46	0.44
Baccalaureate in related field	2.46	0.07	0.73	1.71	0.09	1.44	1.98
Masters in Nursing (or related field 2008)	8.81	0.18	1.44	9.35	0.19	1.38	0.96
Masters in related field	3.24	0.10	1.14	2.60	0.11	1.69	1.49
Doctorate in Nursing (or related field 2008)	0.40	0.02	0.36	0.42	0.04	1.35	3.73
Doctorate in related field	0.50	0.04	1.15	0.44	0.04	1.35	1.18
Not reported	1.61	0.08	1.45	1.12	0.07	1.44	1.00
Age of Nurse							
<25	2.12	0.05	0.43	2.53	0.11	1.56	3.62
25 to 29	5.90	0.13	1.09	6.66	0.13	0.95	0.88
30 to 34	8.36	0.19	1.68	8.99	0.12	0.63	0.37
35 to 39	9.95	0.23	2.11	10.55	0.13	0.62	0.29
40 to 44	14.03	0.23	1.57	11.08	0.21	1.56	1.00
45 to 49	17.49	0.26	1.67	14.17	0.20	1.05	0.63
50 to 54	15.93	0.32	2.73	15.76	0.23	1.37	0.50
55 to 59	11.62	0.22	1.68	12.43	0.20	1.22	0.72
60 to 64	7.22	0.20	2.13	7.81	0.15	1.05	0.49
Not reported	1.00	0.05	0.90	2.94	0.11	1.45	1.61