Sampling Methodology

The Staff Time Resource Intensity Verification (STRIVE) project's data will be used for two primary purposes: (1) to refine the current Resource Utilization Group (RUG) classification systems, and (2) to develop case mix indices (CMIs) which will form the basis for future versions of the Medicare skilled nursing facility (SNF) payment rates. The data collected are of two major types: resident assessment data and staff time data. These data were collected in a sample of 205 nursing homes.

The two major applications, RUG refinement and CMI development, have rather different data requirements. Both applications require data that can be generalized to the national population of nursing homes. For development of CMIs, it must be possible to accurately generalize the data to the nation as a whole so that the CMIs and payment rates that are computed from the data accurately reflect nursing home costs nationally.

For RUG refinement, in contrast, it is not absolutely essential that the data be collected in a nationally representative sample of nursing homes. The RUG system is based upon relationships between a resident's characteristics (measured by assessment instruments) and staff costs for the resident (measured by staff time devoted to resident care). Numerous staff time studies in different settings and even different countries have demonstrated that these fundamental relationships are relatively invariant across different populations of nursing homes and residents[1][2][3]. Thus, there is a substantial body of evidence that suggests that even if data are not strictly representative of the population as a whole, a valid RUG system can be developed.

However, there are two reasons why a fairly complex sample of nursing homes is necessary to meet the needs of RUG refinement. First, it is important in drawing a sample of nursing homes to exclude facilities that have very poor quality. Poor quality facilities are likely to have low or sub-optimal staffing levels. The relationships between resident care and staff time in such facilities may be attenuated or distorted, and there is general agreement that it is not appropriate to include such facilities in the staff time study. Second, there are a number of relatively rare, but high-cost special populations that should be over-sampled so that there are enough cases to model costs for members of these small but important groups. To meet these needs, a complex, probability-based sample design was developed that involved clustering, stratification, and sampling with probability proportional to size.

It is important to note, however, that due to practical and logistical considerations, it was not possible to adhere strictly to random sampling when selecting facilities or residents for participation in the study. Data collection in this study was dependent to a large degree upon volunteers. State agency and nursing home association staff volunteered to supervise the collection of data in participating facilities. The nursing homes themselves could not be required to participate. Although facilities were recruited from a randomly drawn sample, study staff had no control over which facilities agreed

---

[1] Ljunggren G, Fries BE, Winblad U. "International Validation and Reliability Testing of a Patient Classification System for Long-Term Care" *European J. Gerontology*, 1(6):372-383 (Jul-Aug), 1992

[2] Ikegami N, Morris JN, Fries BE. "Low-Care Cases in Long-Term Care Setting: Variation Among Nations" Age and Ageing, 26 Suppl. 2:67-73, 1997.

[3] Björkgren MA, Häkkinen U, Finne-Soveri UH, Fries BE. "Validity and Reliability of Resource Utilization Groups (RUG-III) in Finnish Long-Term Care Facilities" *Scandanavian J Public Health* 27:228-234, 1999.

to participate. Finally, within larger participating nursing facilities, residents and staff could not be selected randomly for inclusion in the study. Logistical considerations guided the selection of the particular nursing units that were included in the study when the nursing home was too large to be included in its entirety.

For all of these reasons, the STRIVE sample of nursing homes, residents, and staff was not strictly random. The sample design, however, was rigorously derived and was developed to satisfy the competing data requirements of the study, while accommodating practical limitations. Random processes were used wherever possible and, where not possible, standardized procedures were substituted to eliminate judgment-based biases from affecting the sample.

The following details the methodology that was used to draw the STRIVE sample.

Overview of Sample Design

Table 2-1 below presents an overview of the steps involved in drawing the sample. These steps are briefly described in this "Overview" section. For complete detailed documentation of each step, refer to Appendix A-1.

Table 2-1. Overview of Sampling Procedures

Step	Description	Sampling Procedure
1	Identified all certified facilities in the nation.	Definition of population
2	Identified 15 states that agreed to participate in the study.	Self-selection (not random)
3	Applied data-based exclusions using QI/QM data and survey deficiency data. Eliminated poorest quality facilities in each state (5% to 10% of all facilities). Population defined as all remaining facilities (referred to as "eligible facilities").	Redefinition of population
4	Applied geographic restrictions in certain states.	Redefinition of population
5	Stratified eligible facilities within each state into five strata. Some strata were not represented in some states.	Stratification
6	Set targeted number of facilities for each stratum within each state. Targets were based upon number of available facilities, number of facilities data monitors were able to visit, and overall study targets.	Sample size determination (no selection involved)
7	Within each stratum within each state, selected the target number of facilities with probability proportional to size (where size was defined by the number of residents in the facility on a given day). Selected an over-sample allowing for deletions and refusals.	Sample with probability proportional to size
8	Each list of sampled facilities (for each stratum within each state) was put in random order.	Randomization
9	Sample lists within each state were reviewed by stakeholders who eliminated facilities that were closed, unable to participate, or were known to be of very poor quality.	Exclusions based on judgment (not random)

Step	Description	Sampling Procedure
10	Remaining facilities were contacted in random order until enough facilities in each stratum agreed to participate.	Self-selection (not random)
11	Enumerated nursing units within each selected facility. If facility was sufficiently small, all nursing units were included in study. In larger facilities, units were selected using a standard protocol.	Selection by standard protocol (not random)

- **Step 1.** The data from the Center for Medicare and Medicaid Services (CMS) national MDS repository were used to identify all facilities in the nation that had submitted at least one MDS assessment over a target six-month period. This list represented all certified facilities in the nation and was used to construct the initial sampling frame for each of the states that participated in the study.
- **Step 2.** States were contacted to enlist support for the study. Because participation in the study required state agencies and/or state nursing home associations to provide volunteers who could travel throughout the state and serve as data monitors, not all states that wished to participate could do so, due to budgetary and other constraints. Fifteen states were able to meet the study's staffing requirements and agreed to participate (refer to Table 1-4 for a list of participating states). Note that these 15 states do not represent a random sample of states since they volunteered to participate (i.e., they were self-selected).
- **Step 3.** Two data sets were used to eliminate poor quality facilities from the sampling frame. The first data set consisted of quality indicator and quality measure (QI/QM) data for all facilities in the nation. Based upon advice from a panel of QI/QM development experts, a set of measures were identified which were combined to produce a composite QI/QM score, with a higher score indicating poorer quality. Facilities were rank ordered by this score within each state and the top (worst) 5% within each state were eliminated from the sampling frame.

  The second data set consisted of special focus facility (SFF) scores for every facility in the nation. The SFF scores are produced using an algorithm developed by CMS that considers survey and complaint deficiencies and scores them according to scope and severity. Facilities with the highest scores are those with the most serious quality-related deficiencies. These scores were rank ordered within state and the top (worst) 5% within each state were eliminated from the sampling frame.

  Since 5% of the facilities in each state were eliminated on the basis of QI/QM scores and 5% were independently eliminated on the basis of SFF scores, a total of 5% to 10% were eliminated within each state. This step essentially redefined the population of facilities that comprised the sampling frame. Instead of including ***all*** certified facilities, the sampling frame now consisted of all certified facilities except those of the lowest quality. In the discussion below, these remaining facilities are referred to as "eligible facilities."
- **Step 4.** Due to travel and budget restrictions, four states (Florida, Illinois, Louisiana, and Texas) agreed to participate but only if the study area was restricted to certain sections of the state. Where these restrictions applied, counties were used to define the geographic areas of interest. All eligible facilities in the agreed-upon set of counties comprised the sampling frames for those states with geographic restrictions. For the remaining states, all eligible facilities in each state comprised the sampling frame.
- **Step 5.** As mentioned above, it was deemed necessary to over-sample members of certain resident special populations in order to yield a sufficient number of cases for the RUG classification analysis. A literature review as well as consultation with CMS and the project's

Technical Expert Panel yielded a list of potentially important special populations. National MDS data were used to estimate the relative prevalence of each of these groups in the population and estimates were made of the number of cases in each group that would be obtained in our sample without employing a targeted sample. Several groups were identified that were sufficiently rare and important that targeting was required and that were sufficiently concentrated in facilities that targeting was likely to successfully enhance their representation in the sample. The strata that resulted were as follows:

- Hospital-based (HB) facilities. Residents in HB facilities typically have stays that are considerably shorter than residents in non-HB facilities. Furthermore, HB facilities typically have different staffing patterns and cost structures than non-HB facilities. For this reason, HB facilities were included as a stratum.
- Facilities with a high concentration of residents on ventilators/respirators (Hi-Vent). Residents who are on ventilators/respirators are thought to be costly to care for and to require more intensive staff resources. Using MDS data, facilities were identified in which 12% or more of their residents were on ventilators/respirators. These facilities fell into the Hi-Vent stratum.
- Facilities with a high concentration of residents with HIV (Hi-HIV). Residents with HIV are thought to be costly to care for and to require more intensive staff resources. Using MDS data, facilities were identified in which 10% or more of their residents had HIV. These facilities fell into the Hi-HIV stratum.
- Facilities with a high concentration of Medicare Part A residents (Hi-PartA). The RUG refinement results and the CMIs that result from the STRIVE study are relevant to both the Medicare and Medicaid programs. However, because only a relatively small minority of residents (12.4%) nationally is served in SNFs under Medicare's Part A program, it was necessary to over-sample facilities that served such residents in order to obtain a sufficient number for analysis. Using MDS and SNF claims data, facilities were identified in which 20% or more of their residents were in stays that were paid for under Medicare Part A. These facilities fell into the Hi-PartA stratum.
- All remaining facilities ("Other"). Facilities that did not qualify for any of the four strata described above fell into the "Other" stratum.

The strata were defined hierarchically in the order listed above so that they were mutually exclusive. Thus, if an eligible facility qualified for more than one stratum, it was classified into the first one in the list that it qualified for. Using this approach, every eligible facility in each state was classified into one of the five strata listed. Note that some strata were not represented in some states.

- **Step 6.** Targets were set for each stratum within each state. These targets represented the number of facilities that were desired for inclusion in the sample. The targets were set based on a number of factors: (a) the number of facilities that each state's data monitors could visit, (b) the number of facilities that were available within each state's strata (for some of the rarer strata, the number of available facilities was sometimes quite small), and (c) the number of cases that were needed within each stratum for the sample as a whole. The sample-wide targets for each stratum were set on the basis of an extensive set of simulations and sensitivity analyses that were aimed at determining the sample sizes that were required to achieve reasonable levels of precision and statistical power. As the study progressed, state sample targets were sometimes modified because of sample shortfalls in states that had been completed earlier in the study.
- **Step 7.** The SAS SURVEYSELECT procedure was used to select facilities from each stratum within a state. Facilities were selected with probability proportional to size, where "size" was represented by the number of residents in the facility on a particular day, according to the facility's MDS data. The use of probability proportional to size sampling favored the selection of larger, higher-volume facilities, although every eligible facility (even very small facilities) had a non-zero

probability of selection. This method was used for two reasons: (a) it tended to select relatively few facilities that were very small so that scarce project resources were not spent on facilities that would yield a low number of cases, and (b) sampling with probability proportional to size tends to produce better national estimates with lower margins of error when the units to be selected vary widely with regard to size (as is the case among nursing facilities). Because this sampling methodology was used and because larger facilities had a high probability of inclusion in the sample, it will be necessary to develop case weights that can be applied to make accurate national estimates from the data.

It was anticipated that some facilities on the sample lists would be eliminated by stakeholders (see Step 8, below) and that other facilities would not agree to participate in the study. Therefore, the number of facilities that were selected for the sample within each state and stratum was greater than the target number of facilities. In other words, each stratum was over-sampled with the aim of having enough facilities available on the sample lists to meet the target desired, after stakeholder exclusion and facility refusal to participate.

- **Step 8.** Each state/stratum facility sample list was put in random order in preparation for Step 10, below.
- **Step 9.** Each of the stratum sample lists within a state was sent to stakeholders for review. These stakeholders were state Medicaid agency staff as well as regional CMS staff. Stakeholders were asked to review the lists and eliminate any facilities that (a) were closed, (b) were unable to participate in the study due to emergencies, bankruptcies, investigations, etc., or (c) were known to have very poor quality.
- **Step 10.** All facilities on the sample lists that remained after stakeholder review were deemed eligible to be contacted for enlistment in the study. Recall that in Step 8 the lists were put in random order. Facilities that had not been eliminated by the stakeholders were contacted in the order listed until the target number of facilities agreed to participate.
- **Step 11.** Once a facility agreed to participate in the study, it was scheduled for onsite data collection. The number of supervisory staff and the amount of equipment available allowed for a maximum of about 50 residents and associated staff to be included in the study in each facility. If a facility was sufficiently small, all nursing units, residents, and staff were included in the study. However, many facilities could not be studied in their entirety. For these facilities, nursing units were selected for inclusion in the study. Residents and staff who belonged to those nursing units were included in the data collection, while outside residents and staff were excluded.

  In these larger facilities, it was not possible to randomly select nursing units for inclusion in the study. Nursing units were sometimes located on different floors of a building or, in some cases, in different buildings on a campus. In such cases, the only feasible option was to study units that were located adjacent or very near one another. It was often not logistically possible for the data monitors to supervise widely dispersed nursing units. Moreover, if the facility was in a special population stratum (Hi-Vent, Hi-HIV, or Hi-PartA), then priority was given to units with residents in the corresponding special population in order to maximize the number of residents in our sample who belonged to those special populations.

  Therefore, random selection of nursing units was not performed. Instead, a standard protocol was developed for the selection of units and this protocol was followed by project staff in consultation with nursing home management. This procedure minimized the use of judgment-based selection, which might impose unknown biases. Instead, the selection procedure was uniform across nursing homes and applied by project staff who managed the study. When the volunteer data monitors arrived at the facility, the nursing units to be included in the study had already been selected by project staff – the data monitors were not involved in the selection procedures.

It can be seen that the sample design that was used for this study was a three-stage cluster sample with stratification. The three types of clusters selected at the three stages were participating states, participating facilities, and the nursing units within each participating facility. Stratification of facilities was applied within states.