

## Cross-Sectional and Transitional Weights Statement, 1982 - 1989

The LTCS, and the associated Public Use Analytic File, contain two types of weights, cross-sectional weights generated by the U.S. Census Bureau, and cross-sectional and transition weights calculated from the U.S. Census Bureau weights for specific types of longitudinal analyses conducted by researchers at the Center for Demographic Studies (CDS). The Census Bureau's cross-sectional weights are contained in the National Long Term Care Survey dataset. The transition and cross-sectional weights developed by CDS are provided in the Public Use Analytic file. The CDS transition and cross-sectional weights are designed to be consistent with a longitudinal observational plan attempting to measure temporal changes in individual functioning. A user of the file may either choose to use these weights in longitudinal analyses, or use the Census Bureau weights making his own adjustments for changes in screening procedure, mortality and the timing and classification of nonresponse. The CDS weights were created in several steps starting with the Census Bureau sample weights as a base. The steps are briefly outlined below.

The starting point in creating weights for longitudinal analyses are the Census Bureau's LTCS Screener Final Cross-Sectional weights for 1982, 1984 and 1989 because they provide weights for all persons in each sample. This cross-sectional weight reflects such factors as age, race, and sex, as well as other detailed features (e.g., PSU selection), of the sample design. Each respondent is determined, on the basis of such factors, to represent a number of people (the inverse of the probability of selection) in the U.S. population aged 65 and older. The sum of the cross-sectional weights equal an independent Census estimate of the total number of persons in the United States aged 65 and over at a specific time, e.g., about 30.9 million in 1989 (26.9 million in 1982; 28.0 million in 1984).

In 1989, persons from the two previous survey years (1982 and 1984) who were chronically disabled and living in the community, or who were institutionalized, received a partial screener interview (not containing disability items) to verify their residence prior to receiving a detailed interview for which they automatically qualified based on their past status. All such persons were assigned a screener final cross-sectional weight. A second group, all persons who completed a full screener interview (i.e., aged into the sample in the inter-survey period, or were

previously nondisabled sample persons selected to be screened anew in 1989 for disability status) also received a screener weight, regardless of functional status. Of persons screened, those reporting at least one ADL or IADL limitation lasting, or expected to last, at least 90 days, were also eligible for the detailed community or institutional interview in 1989.

A similar process was carried out for the 1984 survey --except, of course, that automatic eligibility was based only on the 1982 disability or institutional residence status. In 1982, the first survey, there were no automatically eligible persons since all persons were screened in that initial survey. Another difference between 1982 and 1984-1989 is that no detailed interview was administered to institutional residents in 1982.

In weight calculation, it is useful for data processing to define a "group" variable whose categories describe the response status of sampled individuals across survey dates to help keep track of response status and subsampling changes for individuals between survey dates (e.g., screener respondent; screener nonrespondent; person dead at time of attempt to deliver the screen; detailed community interview respondent; detailed community interview nonrespondent; persons who were eligible for a detailed community survey because of previous disability, or who had disability on the screen, but died between the screener and the attempt to deliver the detailed community interview; or who were not sampled in a given year because of the balance of budget restrictions and requirements for sample component precision; etc.). Such group variables represented major components of the sampled population for the weights developed by U.S. Census Bureau sampling statisticians.

For the longitudinal analyses, we developed new group variables with additional detail on, e.g., level of disability. For the 1982-1984 longitudinal version of the file, 25,541 persons were retained although the total number sampled was 35,789 in 1982 with nearly 5,000 persons who passed age 65 between 1982 and 1984 added into the 1984 sample. The major reduction in the total sample of about 40,000 persons was that a subsample of 12,100 was drawn of 26,623 persons who did not report disability on the screen in 1982 because a sample of 10,250 was determined to be sufficient to give a relative standard error of 10% if 1.6% of the population were

institutionalized in the two year interval (1982 to 1984). In fact, of the 12, 100 selected, 11,892 were confirmed alive at the 1982 screen date and 11, 151 both survived to 1984 and were screened. Thus, 25,541 persons were included in the longitudinal file though in fact there are sample persons not in the 1984 sample (i.e., not receiving a sample weight in 1984) who were nondisabled in 1982. Likewise, for variance optimization under budget restrictions, we created a longitudinal file for 1982, 1984 and 1989 containing 30,308 persons drawn from the original sample and the roughly 5,000 additional persons drawn who passed age 65 between 1982 and 1984 and a second such group for 1984 to 1989 (i.e., a total sample pool of about 45,000 in 1989). In both files, the subsampling is adjusted for by modifying (reallocating) the sample weights for individuals in relevant groups to match the independent population estimates.

In fact, different definitions of the longitudinal population can be made depending upon when the re-weighting for subsampling is applied. In the 1982, 1984 and 1989 longitudinal file, group variables were defined for all three waves of the survey and represent an exhaustive, but exclusive set of survey outcomes. For example, the 1989 group variable both classifies persons in the 1989 sample, and charts the disposition in 1989 of persons in the 1982 and 1984 surveys who were not surveyed in 1989 (e.g., died prior to 1989 or were not sampled in 1989). This longitudinal group variable is available in the CDS Analytic File included with the NLTCS Public Use File (DISABILITY\_GROUP\_SY82, DISABILITY\_GROUP\_SY84, DISABILITY\_GROUP\_SY89). This was the second stage in the development of transition weights.

The third stage in the production of transition weights, used the Census Bureau Cross Sectional Weights, and the detailed group variables, for all three survey years (1982, 1984, 1989) to calculate a set of Cross-Sectional Weights to compare disability prevalence between pairs of survey years ( 1982-1984 and 1984-1989). Calculation of the prevalence weights required allocating nonrespondents across the group variable categories used to describe the population distribution of disability using all available information, e.g., detailed community and institutional nonrespondents were proportionately allocated to survey completers in their own response category.

One analytic issue that will increase in importance with time is the allocation of nonrespondents to the community disabled or institutional populations. The allocation does not affect the total estimate of the elderly chronically disabled and institutional population --or its change. However, as there become intermediate housing options (i.e., between nursing home use and independent living) the boundary between the community disabled and the institutional population has become less distinct. Thus, in 1984 nonrespondents were exhaustively categorized as community or institutional residents. In 1989, three categories were defined, i.e., community nonrespondents, institutional nonrespondents, and community or institutional nonrespondents. Obviously, the 1989 three-part categorization involves an additional analytic distinction to make nonresponse allocations. This distinction affects the proportions in the disabled or institutional populations -- but not the total disabled elderly population. Likewise, just as the Next-of-Kin Survey in 1984 had to be adjusted for in identifying cases, the effects of the six-month institutional follow-up survey in 1989 had to be adjusted for in identifying cases (these adjustments affect very small numbers of cases).

Several other adjustments were made to the U.S. Census Bureau sample weights, and the group variable categories, in producing CDS Cross-Sectional weights. Most important is that the criterion for receiving a detailed community or institutional interview in 1984 and 1989 changed slightly from 1982. Specifically, all persons sampled in 1982 were screened for chronic disability. In 1984 and 1989 persons who had been identified as disabled and being in the community, or institutionalized, in a prior survey were automatically eligible for a detailed community or institutional interview. As a consequence they only received a partial screen and the detailed interview. This allowed tracking of all types of changes in disability over time since persons previously disabled, who subsequently regained function, received detailed interviews. However, to define a disabled population wholly comparable to that in 1982, persons not receiving a screen have to be identified and their disability status evaluated relative to the screen criterion. This is easy for those reporting no disability on the detailed interview. Others reporting only the IADL impairment on "heavy housework " on the community interview were recoded as community residents

receiving a detailed interview who were nondisabled because heavy housework was not a screen IADL, e.g., persons with only that IADL disability would not "screen in" as chronically disabled in 1982. Second, certain subgroups were sampled. In the longitudinal file with 30,308 cases, this required a ratio adjustment to various group weights.

Identifying persons reporting no disability, but who received a detailed instrument, provided analytic opportunities not originally contemplated in the 1982 sample design. First, a sizeable group (of about 2,500 persons in the three surveys combined) of "currently" nondisabled people received a detailed interview --thus providing a basis for examining the nondisabled U.S. elderly Medicare enrolled population in some detail. To do this, the sample weights of that group have to be renormalized to the total size of the U.S. nondisabled population; and their representativeness of the "screen out" nondisabled population examined using available ancillary data like Medicare service use and mortality files. Second, because nondisabled persons are likely to have different nonresponse reasons than disabled persons the accuracy of sample weight adjustments for nonresponse in the nondisabled population could be improved. Third, these weights contained specific adjustments for analyses of longitudinal change. For example, Medicare files with dates of death for each individual, only became complete after the initial Census Bureau weight calculations. An adjustment for differential mortality in group categories was performed prior to the availability of complete individual mortality data on dates of death as part of the cross-sectional weight development process using the available aggregate mortality data normed to the NCHS reported counts of death over the inter-survey interval (Manton et al., 1993). When the complete mortality file, which contained the date of death for each person dying between surveys, and during the interview process, became available, those detailed individual mortality records were used to fine tune mortality weight adjustments -- though the initial and final mortality adjustments were very close.

Thus, four sets of cross-sectional weights were **initially** developed, i.e., 1982, 1984a, 1984b and 1989 weights are defined by dividing the sum of the preliminary weights by the marginal counts from the cross-tabulations of the group variable. The two **interim** weights were defined for 1984 because of uncertainty about certain types of

nonresponse during the 1984 interview period. For example, if a person interviewed in 1982 died before the attempt to deliver the detailed interview in 1984 (such a person would not be screened so that a screen date is not available), he should not be counted as a nonresponse in 1984 (i.e., this person would be a valid decedent associated with his 1982 status). If an attempt is made to deliver a detailed interview in 1984, and during the course of several call backs it was ascertained that the person had recently died, we could have a health-related nonresponse to the survey. Thus, to evaluate the magnitude of these effects, we took the precaution of calculating two sets of 1984 weights.

In addition, one has to consider the different definition of the "left hand" (1984a) and "right hand" estimates (1984b). Specifically, the 1984a groups contain roughly 2.7 million persons dying between 1982 and 1984 because it is tied longitudinally to the 1982 surveys. The 1984b groups contain 10.1 million persons who passed age 65 between 1984 and 1989. Though neither the 2.7 or the 10.1 million persons are used in making 1984a and b prevalence estimates, their presence as part of the calculation of changes 1982 to 1984 or 1984 to 1989 makes slight differences in the 1984 weights. (Such right and left hand sided estimates will exist for any survey wave linked to waves conducted before and after; if the differences between the two sets of weights are small this suggests the sample selection and weighting procedures are consistent over time -- especially with respect to age-in sample selection and mortality adjustments.) The effects on the two 1984 weights were indeed quite small -- they differ in total by 0.1%. However, even though the difference is small (i.e., one part in a thousand), it has to be reconciled for consistent analyses of all three years. One strategy is to average the 1984a,b sample weights so the net expected error is 0.05%. Another is to use the exact dates of death and exact dates of interview attempts in a more complex statistical model for nonresponse (Manton et al., 1991). The group variables used in calculating the three distributions are contained in Table 1. In the tabulations provided, there are slight differences between the 1984 distribution and that reported in the attached papers because the average of the 1984a and b weights were used in the paper but only the 1984a weights are provided in the attached tabulations.

The final step is the development of transition weights. This requires that the status of the population be

tracked as it ages from one study period to the next. Transition weights were generated from the weights used for the analysis of prevalence change for the LTCS respondents whose status could be determined in both 1982 and 1984 or 1984 and 1989 by cross-tabulating survey pairs ( 1982-1984 and 1984-1989). For example, people who received an interview in 1984 and 1989 would be given a weight for the transition population since their status is known in both years. People who were interviewed in 1984, but could not be reached for an interview in 1989, would not be given a weight for the transition population because their status in 1989 is indeterminate. Persons who were interviewed in 1984 and died between 1984 and 1989 were assigned a transition weight (their status in 1989 is known -- they are deceased), as do the samples of people who passed age 65 between the 1982 and 1989 surveys. Thus, the sum of the 1984-1989 transition weights, shown in Table 2, is 38,145,478 -- 20,748,795 people with valid interviews in both 1984 and 1989, 10,122,550 people who became 65-69 years old between the two survey periods and had a valid interview in 1989, and 7,274,133 people who had a valid interview in 1984 and died in the interval. Table 3 presents the same type of information as Table 2 but for the 1982 and 1984 observations. Thus, to successfully analyze changes over time, the nature of the surveys as sampling a population process has to be kept in mind to make appropriate sampling and nonresponse adjustments.

Finally, it is worthwhile noting that additional file improvements will be made of several types. One is that additional Medicare experience and mortality will be linked to the survey records. Two, is that several experimental nonresponse adjustments will be examined to fine-tune estimates. Since nonrespondents who screened-in as disabled or who were automatically eligible for the survey are always counted as disabled, this will tend to reduce disability estimates in later years, i.e., or when ancillary data is used to make better nonresponse adjustment. Thus, the analytic file variables are subject to improvement, e.g., there are a number of decision points in the questionnaire that will be evaluated to see if estimates can be improved. The net effect of these changes will be to produce "best" estimates. They will not alter the broad changes identified in the survey but will serve to reduce certain types of conservative assumptions built into those estimates.



## References

Manton KG, Corder LS, Stallard E: Estimates of change in chronic disability and institutional incidence and prevalence rates in the U.S. elderly population from the 1982, 1984, and 1989 National Long Term Care Survey. *Journal of Gerontology Social Sciences* 47(4):SI53S166, 1993.

Manton KG, Stallard E, Woodbury MA: A multivariate event history model based upon fuzzy states: Estimation from longitudinal surveys with informative nonresponse. *Journal of official Statistics* (published by Statistics Sweden, Stockholm) 7:261-293, 1991.