Extending Gerontological Research Through Linking Investigators’ Studies to Public-Use Datasets

Lisa Fredman, PhD,1 William Hawkes, PhD,2 Sheryl Itkin Zimmerman, PhD,3 J. Richard Hebel, PhD,2 and Jay Magaziner, PhD, MsHyg2

Purpose: Public-use datasets can extend data collected by individual investigators in various ways: making external comparisons, providing additional data on individual respondents, and creating internal comparison groups. The authors describe the advantages and limitations of these methods and practical and conceptual issues in combining investigator-initiated and public-use datasets. Design and Methods: These issues are illustrated with a study of functional decline among 674 patients following hospitalization for hip fracture that was augmented with data from a public-use dataset, the Established Populations for Epidemiologic Studies of the Elderly (EPESE). Results: By creating an internal comparison group of EPESE respondents, frequency matched to hip fracture patients on age, sex, and baseline functional limitations, the authors formed a single dataset and performed multivariable analyses of factors associated with functional decline. Implications: Gerontological research may benefit by applying these methods to program evaluations and longitudinal analyses of health outcomes with numerous public-use datasets.

Key Words: Data linkage, EPESE, Hip fracture, Physical functioning

Investigators are using public-use datasets increasingly to extend data collected by gerontological researchers. These datasets are becoming more available (Post, 1996), as are methods to link them to datasets from investigator-initiated studies. In this article we describe three methods of combining investigator and public-use datasets and examine the methodological challenges, benefits, and limitations of each method.

Traditionally, investigators have used public-use datasets to provide external comparison groups to their own samples, such as comparing rates of well-being among caregivers to rates in general populations (George & Gwyther, 1986; Pruchno, Kleban, Michaels, & Dempsey, 1990). By external comparison group, we mean a referent group that is drawn from a different source than the investigator's data, usually because the investigator selected participants on the basis of a specific exposure (such as members of a caregiver support group) and there was no logical group of unexposed persons from the same source (i.e., an internal comparison group; Kelsey, Whittemore, Evans, & Thompson, 1996).

Recently, investigators have used public-use datasets in another way: to supplement data on their samples with endpoints from public-use datasets. Examples include health care utilization and health outcomes from Medicare claims files (Lillard & Farmer, 1997; Lu-Yao et al., 1996; Newschaffer et al., 1998; Potosky, Riley, Lubitz, Mentnech, & Kessler, 1993), cancer incidence and mortality from the Surveillance, Epidemiology and End Results (SEER) dataset and other cancer registries (Lu-Yao et al., 1996; McClish et al., 1997; Newschaffer et al., 1998; Potosky et al., 1993), all-cause mortality from the National Death Index (Fredman, Magaziner, Hebel, Hawkes, & Zimmerman, 1999; Howe, 1998), and outcomes among nursing home residents with Medicaid records (Lipowski & Bigelow, 1996) and with the Systematic Assessment of Geriatric Drug Use via Epidemiology (SAGE) database (Bernabei et al., 1999).

A third method is to merge an investigator’s cohort with public-use data to create an internal comparison group, the result being a single dataset that includes exposed (the investigator’s sample) and unexposed (the public-use sample) individuals. This method allows the investigator to address numerous questions...
that can only be answered with an internal comparison group.

In this article we compare these three methods. We use as an example a study of functional decline following hip fracture in which we merged an investigator’s dataset (Magaziner et al., 2000) with a public-use dataset, the Established Populations for Epidemiologic Studies of the Elderly (EPESE; Cornoni-Huntley, Brock, Ostfeld, Taylor, & Wallace, 1986; Guralnik, Ferrucci, Simonsick, Salive, & Wallace, 1995). We describe the methodologic issues and decisions we made in merging these two datasets and the strengths and weaknesses of the approaches we chose.

Definition of Public-Use Datasets

Public-use datasets may be defined as data from federal or state agencies, and datasets placed in the public domain, that are available to researchers. These datasets may include population health records, such as mortality records, and disease registries, such as cancer registries. They may also include records on health care utilization and costs, such as Medicare and Medicaid records. Other types of public-use datasets include information from studies of nationally representative samples, such as the National Health Interview Survey, as well as studies of specific populations. Examples of these latter datasets that are relevant to gerontologists are the National Long-term Care Surveys, the EPESE study, the Assets and Health Dynamics of the Oldest Old (AHEAD) study, and datasets available through the National Archive of Computerized Data on Aging (NACDA), to name a few (see Appendix A). Datasets such as the Longitudinal Study on Aging (LSOA) are already linked to Medicare records and the National Death Index. These latter datasets vary in terms of the extent of the population they capture and the range and type of sociodemographic and health variables they include.

Issues in Combining Public-Use Data and Data From Investigator-Initiated Studies

Most reviews have focused on concrete issues of combining data from investigator-initiated studies with public-use data (Brenner, Schmidtmann, & Stegmaier, 1997; Howe, 1998; Lillard & Farmer, 1997; Lipowski & Bigelow, 1996; Potosky et al., 1993; Roos, Walld, Wajda, Bond, & Hartford, 1996). These issues include the mechanics of linking an existing dataset to Medicare or Medicaid records (Lillard & Farmer, 1997; Lipowski & Bigelow, 1996) and involve problems that arise when there are multiple persons with the same name or Medicare number (Brenner et al., 1997; Howe, 1998) and decisions when data may be in both Medicare and Medicaid files when an elderly person is eligible for both (Lipowski & Bigelow, 1996). Respondents from the investigator’s dataset might be missing from the public-use dataset that is the source of follow-up data because of respondents moving out of the geographic area covered by the public-use dataset (Adams et al., 1997; Howe, 1998), a change in personal identifiers (Brenner et al., 1997; Lillard & Farmer, 1997), incomplete personal identifiers (Adams et al., 1997; Brenner et al., 1997), reporting or coding errors (Brenner et al., 1997; Howe, 1998; Lillard & Farmer, 1997), or use of health services that are not covered by Medicare, for example (Lillard & Farmer, 1997; Lu-Yao et al., 1996). These concrete problems may differ according to the source of the investigator’s dataset and the type of outcome dataset. For example, linking outcome data to a cohort of persons with a particular disease, such as persons identified in a cancer registry, incurs different difficulties than combining data with a cohort of community-dwelling elderly persons. Likewise, problems encountered in linking an investigator’s dataset to mortality data differ from those encountered in connecting to data on functional status or health care utilization.

Investigators who have data on the respondent’s name, birthdate, and address may link to mortality data from the National Death Index or state-based records. However, linking to other datasets necessitates collecting specific information, such as social security numbers to link to Medicare data. Investigators who anticipate adding respondent-level data from external sources should include this information in a letter of consent for respondents.

Conceptual issues of combining investigator-initiated studies with public-use datasets are inherently connected to these procedural issues, which affect the research questions that may be asked, features of the comparison group, and the ability to control for potential confounders.

The research questions that may be addressed are constrained by the options for linking investigators’ and public-use datasets. Most investigator-initiated studies may be linked to mortality data from the National Death Index or state-based records. These linkages are more feasible than linkages to health care utilization datasets because there is only one outcome, mortality, as opposed to multiple outcomes. By adding mortality data on their respondents, investigators can evaluate factors associated with survival in their sample, or whether survival is better in their sample than an external cohort, such as same-aged residents from their state. Such comparisons have been made in investigations of suicide among psychiatric patients (Tsuang, 1978), mortality following hip fracture (Magaziner et al., 1997), and an evaluation of the Supplemental Food Program for Women, Infants, and Children (Kotelchuck, Schwartz, Andorka, & Finison, 1984).

Datasets that contain longitudinal measures of functional and health status may be combined not only with mortality and health care utilization records, but also with datasets of nationally representative samples or specific populations. These linkages allow investigators to evaluate the relative risk of specific health outcomes and longitudinal changes in functional and health status over time in their cohort.
versus the comparison cohort. An example of this method is our study on functional health decline that combined a cohort of patients hospitalized for hip fracture with data from the EPESE study, described later. Using the public-use dataset to create a comparison group allowed us to answer more research questions than we could have addressed if we had only used a public-use dataset to add outcome data on individuals in our cohort, but it also introduced a greater possibility for confounding and bias of the results. Furthermore, the decision of which public-use dataset to merge with the investigator’s dataset depends on whether the public-use dataset contains the outcome of interest, measures key variables using the same method, and has follow-up intervals of similar lengths.

Another conceptual issue in creating either external or internal comparison groups from a public-use dataset is whether the comparison group is appropriate, that is, assuming that the investigator’s cohort is “exposed” to a health care program or a disease, does the comparison group represent the experience that the “exposed” cohort would have had if they had not been exposed (Rothman & Greenland, 1998). For example, would the EPESE cohort represent the experience of elderly patients hospitalized for hip fracture, had they not fractured a hip? Clearly, certain public-use datasets would make appropriate comparison groups for community-dwelling samples, and others would be better for samples of residents in long-term care facilities.

The investigator’s exposed sample and the public-use unexposed sample may have different distributions of important factors related to the outcome, such as age, existing illnesses, or social support. As a result, the comparison of the samples may introduce confounding of the association between the exposure and the outcome. Potential confounding may be minimized by restricting the analyses to subsets of the two datasets that have similar distributions (such as persons aged 75 and older) or by matching on one or more variables, as described in our example of merging the hip fracture and EPESE cohorts. The latter ensures equal distributions of the variables in the two datasets that were matched. Whether or not matching is used, the investigator also may adjust statistically to reduce or eliminate confounding.

Bias resulting from misclassification of outcomes may also occur. This could happen because claims datasets, such as Medicare and Medicaid, might classify a person for reimbursement purposes in a way that differs from how that person would be classified for research purposes (Lipowski & Bigelow, 1996). Additionally, persons might be absent from Medicare files because they received a treatment that was not reimbursed by Medicare or was covered by a different health plan, such as some treatments for prostate cancer (Lu-Yao et al., 1996). Moreover, when investigators link their datasets with registries, matching errors may occur because of coding errors, resulting in mistakenly linking two different persons or missing a correct link between the two datasets (Brenner et al., 1997).

In summary, when combining investigator-initiated data with public-use data, one must consider not only the procedural issues of merging these datasets, but also conceptual issues such as effects on the validity of the study results.

Comparison of Methods to Combine Investigator-Initiated and Public-Use Datasets

Next, we compare the three methods of linking investigator-initiated studies to public-use datasets with regard to the type of public-use data, outcome measures, statistical analyses, and advantages and limitations (see Table 1).

External Comparisons.—Epidemiologists have used public-use datasets to provide external comparisons to investigators’ studies, such as studies of mortality rates in industrial workers versus those in the general population (Grayson & Lyons, 1996; Roscoe, 1997). Similarly, disease rates in particular cohorts have been compared with those in general populations, such as the health status of caregivers compared with that of community-dwelling adults (George & Gwyther, 1986; Pruchno et al., 1990). Benefits of using public-use datasets for external comparisons are that the methods are relatively quick and inexpensive to implement, the external sample is usually well characterized and known, variables are well defined, and established methods exist for computing standardized mortality ratios and for comparing rates of other outcomes.

However, using public-use datasets for external comparisons has drawbacks. These studies can only evaluate rates. The public-use sample may have different characteristics from the investigator’s sample. Public-use samples of community-dwelling elderly persons may be healthier than the investigator’s sample. They are not necessarily a “clean” comparison group in that they may include persons who have the condition that defined eligibility for the investigator’s sample. For example, in studies comparing caregivers to community-based samples, the community-based samples may include persons who are caregivers but who were not so identified. External comparisons allow adjustment for just a few confounders simultaneously, such as age and sex, which may lead to biased results due to factors that were not controlled, such as health status.

Appending Data.—We define appending data as adding individual-level data from a public-use dataset to the individual participants in the investigator’s dataset. Many studies have added mortality data (Adams et al., 1997; Brenner et al., 1997; Howe, 1998) and health care utilization and mortality from Medicare files (Lillard & Farmer, 1997; Lu-Yao et al., 1996; Newschaffer et al., 1998). The addition of these data extends the range of outcomes and research questions that can be addressed. Other benefits are that these public-use datasets are comprehensive, so few participants are likely to have missing outcomes; the investigator can adjust for confound-
Limitations in appending data come from errors in matching the respondent from the investigator’s sample to a record in the public-use dataset. Previous reviews have noted that personal identifiers, such as social security numbers, are not always unique (Brenner et al., 1997; Howe, 1998). Outcome data might be missing for respondents who moved to another state and were not documented in the state registry (Howe, 1998). The data quality has been questioned in datasets that were collected for administrative rather than research purposes: biases may be inherent in the coding protocols. For example, outcome data may be misclassified if the Medicare record lists hospitalization for one condition when another existing condition was the outcome of interest. Appending public-use data does not solve the problem of the lack of a comparison group; this is not the purpose of this method. Exceptions are studies by Wolinsky and colleagues in which pseudo-dates of disease onset were generated for a comparison group created within the Longitudinal Study on Aging (Wolinsky, Fitzgerald, & Stump, 1997). In summary, appending data from public-use datasets allows investigators to evaluate risk of mortality and other outcomes more efficiently than if they had to recontact respondents or their proxies to collect follow-up data, but possible limitations are missing data for some participants and misclassification of some outcomes.

Merging Datasets.—Merging an investigator’s sample with data from public-use datasets has pragmatic benefits. Investigators might not have included a comparison group in the original study because it was too expensive or because the outcome was rare so that a large comparison group or long follow-up period would have been needed. Other reasons are that the data collection methods might be considered too invasive by potential control participants or that it was difficult to identify an appropriate and accessible comparison group. For example, by merging a dataset of elderly women with incident breast cancer identified through the Virginia Cancer Registry with Medicare data on these women and on elderly, female Virginia residents, Newschaffer and colleagues...
created a comparison group of women who were hospitalized for genital prolapse and obtained data on covariables and outcomes for all respondents (Newschaffer et al., 1998).

Benefits of merging an investigator-initiated sample with a public-use dataset are that a single dataset is created and that there is no need for specific information on respondents, such as social security number. The investigator can quantify differences in sample characteristics between the two datasets, control for confounders, evaluate effect modification, and perform complex statistical analyses.

However, finding appropriate datasets to merge with an investigator’s dataset may be difficult. Reasons may include differences in the distribution of key variables and sample characteristics; different data collection methods, such as different scales to measure cognitive impairment or different response options on similar measures; and different decisions about coding responses and missing data. The follow-up intervals may also differ: If the investigator’s study has follow-up at 6-month intervals and the public-use dataset has follow-up at annual intervals, the rate of health decline over 12 months is not necessarily the same as that over a 6-month period. We encountered these challenges in merging the hip fracture dataset with the EPESE dataset. Investigators may minimize these problems at the design phase by using standardized measures and follow-up intervals that are used in existing public-use datasets.

Example: A Study of Hip Fracture Outcome

We combined data from an investigator-initiated study with public-use data to evaluate whether elderly patients hospitalized for hip fracture had more decline in functional limitations over a 2-year period than community-dwelling elderly adults.

Methods

Samples

The investigator-initiated sample included 674 community-dwelling adults, aged 65 years and older, who were admitted to eight Baltimore hospitals in 1990–91 and followed for 2 years (Magaziner et al., 2000). All elderly patients admitted to these hospitals for hip fracture during 1990–91 were eligible for this study. Data were collected through interviewer-administered questionnaires in face-to-face interviews at the time of hospitalization and at 2 months, 6 months, 12 months, 18 months, and 24 months posthospitalization.

The EPESE dataset included prospective data on adults aged 65 years and older living in three communities: East Boston, MA (n = 3,809), Iowa (n = 3,673), and New Haven, CT (n = 2,812; the dataset did not contain prospective data on respondents from the Durham, NC, site). Participants had annual interviews in 1982, 1983, and 1984 and were asked questions on functional status and health. Interviews were conducted face-to-face at baseline and by telephone at 12- and 24-month follow-up. We did not exclude persons with previous hip fractures because the hip cohort included persons who had had previous fractures.

The EPESE cohort appeared well suited to merge with the hip fracture cohort because of the similar age range, time between interviews, and similar key variables. Although the EPESE study began 8 years before the hip fracture study, we assumed that a comparison group of community-dwelling elderly persons from the 1980s would not differ substantially from older adults in 1990 with regard to their rate of change in functional status over a 2-year period.

Hip Fracture Study Variables

At each interview, respondents were asked whether they received help performing each of 16 physical activities of daily living (ADLs) during the past week and how difficult it was to perform each activity. Baseline was considered the week before hospitalization. These questions were modeled after items on the Functional State Index (Jette, 1987) and the Katz ADL scale (Katz, Ford, Moskowitz, Jackson, & Jaffee, 1963) and additional questions on activities related to lower extremity function. Being able to dress oneself was separated into four activities: put on a shirt/blouse; button a shirt/blouse; put on pants; and put socks and shoes on both feet. Taking a bath/shower was separated into two activities: getting in/out of a bath or shower; and taking a shower, bath, or sponge bath.

Covariables included in these analyses were the respondent’s age at baseline, sex, history of stroke, and presence of four other potentially disabling medical conditions: diabetes, hypertension, cancer, and heart disease. Information on other comorbid conditions was collected, but these were the conditions that were also measured in the EPESE study.

EPESE Study Variables

At baseline and follow-up interviews, respondents were asked three questions on mobility functioning from the Rosow-Breslau scale (Rosow & Breslau, 1966; ability to do heavy work around the house, walk half a mile, and walk up and down stairs without help) and seven questions on ADL limitations modified from the Katz ADL scale (Katz et al., 1963). These questions were on limitations in walking across a small room, getting from bed to chair, eating, dressing, grooming, bathing, and using the toilet. A variety of covariables on sociodemographic and health status were assessed (Cornoni-Huntley et al., 1986). For our analyses, we selected those covariables that were listed previously for the hip fracture cohort.

Decisions Regarding the Choice of ADL Variables.—The questions on ADL limitations were not worded exactly the same in the two studies. For walking indoors, hip fracture patients were asked, “In the past week, on average, did you receive help to walk 10 feet or across a room?” Response options were as follows: received no help, used help from

Vol. 41, No. 1, 2001
equipment or a person, or was unable to walk across a room. The EPESE participants were asked, “Was there any time in the past 12 months in which you needed help from some person or from some equipment or device walking across a small room?” Response options were as follows: does not need help, gets help, or is unable to walk across a small room. We remedied these differences by creating the same dichotomous response for each question: unable to walk across a room versus other. Questions on transfer also differed somewhat. The hip fracture respondents were asked, “In the past week, on average, did you receive help getting in and out of bed?” EPESE respondents were asked if there was any time in the past year that they needed help “getting from a bed to a chair.” We made no adjustment for the different reference periods—past week or past year—of the limitation.

We selected five ADL questions that were similarly worded: walking across a room, transferring, bathing, grooming, and eating. We constructed a single dressing question from four items on the hip fracture questionnaire. No other ADL questions were similar enough to include.

**Decisions regarding the choice of covariables.**—Age was operationalized in 5-year age groups because the EPESE dataset included these groups instead of a continuous variable for age. Cognitive status was not included as a covariable because it was measured by the Mini-Mental State Examination in the hip fracture study and by the Short Portable Mental Status Questionnaire in the EPESE dataset. We considered making a dichotomous variable (cognitively impaired, yes or no) based on each of these measures, but we concluded that there would be interpretation problems because of differences between these measures.

**Results**

The hip fracture cohort was older, more likely to be female, and more likely to include participants who had had strokes than each of the EPESE cohorts (see Table 2). Almost 60% of the hip fracture cohort was 80 years or older, versus 20–24% of the EPESE cohorts. Furthermore, hip fracture respondents were more likely than EPESE respondents to report limitations in walking across a room, transferring, bathing, and toileting at baseline, that is, before they were hospitalized for hip fracture.

Because of these differences, we considered restricting our analyses to respondents who were aged 75 years and older. However, discrepancies remained between the two samples: among persons aged 75 years and older, the EPESE sample was younger and less functionally impaired at baseline than the hip fracture sample. Instead, we used frequency matching of each of the EPESE cohorts to the hip fracture cohort on three baseline variables simultaneously—5-year age group, sex, and ability to walk across a room—to force the samples to be similar on these variables and reduce potential confounding in a way that was less restrictive than pair-matching (Kleinbaum, Kupper, & Morgenstern, 1982). By frequency matching, we created three groups of “unexposed” individuals, each group approximately the same size as the hip fracture group.

Table 3 shows that frequency matching on these variables resulted in four samples of 594 respondents each, with equal distributions of age and sex (because we matched on those factors). The distributions of walking across a room, transferring, and grooming were alike. Compared with the unmatched samples, there were smaller differences in the distribution of stroke between the hip fracture and EPESE samples. However, the distribution of limitations in bathing and dressing still differed across cohorts.
Our preliminary results indicated that with matching and controlling for important baseline covariates, hip fracture patients had an increased risk of impairment in walking across a room, transferring, and grooming at 1 and 2 years posthospitalization, compared with community-dwelling elderly adults. When adjusted for covariates, 26% of each sample was limited in walking across a room at baseline. At 1-year follow-up, 54% of hip fracture patients were limited in ability to walk across a room, versus about 20% of each of the EPESE cohorts. At 2 years, about 54% of hip fracture patients had limitations walking across the room, whereas the proportion in each of the EPESE cohorts had returned to about 26%.

If we had used the EPESE dataset as an external comparison group and adjusted for age and sex only in each unmatched sample, the proportion of respondents with limitations walking across a room at 1 year would have been 53% in the unmatched hip fracture cohort, 10% in the East Boston, 9% in the Iowa, and 11% in the New Haven EPESE cohorts. These results would have been confounded by the younger age and better health of the EPESE versus the hip fracture samples.

We also could have used each unmatched EPESE sample as a “standard population” to calculate the proportion of hip fracture respondents with walking limitations as if this cohort had the age–sex distribution of each EPESE sample. This would reduce confounding because of unequal age–sex distributions between the two cohorts. If we had done this, the proportion of hip fracture respondents with walking limitations at 1 year would have ranged from 42% using the East Boston sample to 44% using the Iowa sample.

Discussion

We compared three methods of linking data from investigator-initiated studies to public-use data. Each method has pragmatic and conceptual advantages and disadvantages. Our study of functional decline following hip fracture showed benefits and drawbacks of merging data from an investigator-initiated study with a public-use dataset. Because we merged these datasets, our results were less confounded by differences in the age and health distributions of the hip fracture and EPESE samples than if we had used the EPESE samples as external comparison groups and adjusted only for age and sex within each dataset. Limitations of this method were that our choice of outcome variables, covariates, and follow-up intervals was constrained by the design of the EPESE dataset, as would be the case with the use of any public-use dataset.

In addition, the matched samples were not representative of their target populations. This posed more of a problem for the hip fracture sample than the EPESE samples: The latter samples were intended to create comparison cohorts to imitate the hip fracture cohort, had they not fractured a hip. The matched hip fracture cohort was unrepresentative of its target population because 80 of the original 674 respondents were excluded because of the absence of an age–sex match in the EPESE samples. These respondents were older and frailer than those who were included. The observed results probably underestimated the true differences in functional decline that would have been observed had these 80 respondents been included in our analyses. If these respondents had a different rate of functional decline than the respondents who remained in our analysis, selection bias might have occurred. Thus, our results should be interpreted with caution.

Our results also must be interpreted in light of the data collection methods in the EPESE study. Data were collected by face-to-face interview at baseline and by telephone at 12 and 24 months. This variation in methods might have accounted for some of the change observed between 12 and 24 months in the proportion of EPESE respondents who reported limitations in walking across a room. We could have avoided this problem by excluding 12-month outcomes data, but we included it to better portray functional decline following hip fracture.

In conclusion, linking investigator-initiated datasets to public-use data is a relatively low-cost method to expand researchers’ datasets. The determining factor in combining datasets is how similar the public-use data are to the investigator-initiated data. Linking to public-use datasets increases the range of questions that researchers can answer, and it may be used to create comparison groups for addressing complex associations when a comparison group might be too difficult or expensive to recruit or for program evaluations where randomized trials are unethical or too cumbersome to perform.

References


Appendix A

Web Sites for Selected Public-Use Datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Web Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medicare</td>
<td><a href="http://www.hcfa.gov/stats/pufiles.htm">www.hcfa.gov/stats/pufiles.htm</a></td>
</tr>
<tr>
<td>National Health and Nutrition Examination Survey I-Epidemiologic Follow-up Study Surveillance, Epidemiology and End Results</td>
<td><a href="http://www.cdc.gov/nchs/aboutmajor/nhefs">www.cdc.gov/nchs/aboutmajor/nhefs</a></td>
</tr>
<tr>
<td>National Death Index</td>
<td><a href="http://www.seer.ims.ncri.nih.gov">www.seer.ims.ncri.nih.gov</a></td>
</tr>
<tr>
<td>Established Populations for Epidemiologic Studies of the Elderly (EPESE)</td>
<td><a href="http://www.cdc.gov/nchs/sd/ndi.htm">www.cdc.gov/nchs/sd/ndi.htm</a></td>
</tr>
<tr>
<td>Longitudinal Study on Aging (LSOA)</td>
<td><a href="http://www.icpsr.umd.edu/NACDA/archive.html">www.icpsr.umd.edu/NACDA/archive.html</a></td>
</tr>
<tr>
<td>National Long-Term Care Surveys (NLTCS)</td>
<td><a href="http://www.cdc.gov/nchs/products/catalogs/subject/lsoa">www.cdc.gov/nchs/products/catalogs/subject/lsoa</a></td>
</tr>
<tr>
<td>Assets and Health Dynamics of the Oldest Old (AHEAD)</td>
<td><a href="http://www.cds.duke.edu/NLTCSSINTRO.HTML">www.cds.duke.edu/NLTCSSINTRO.HTML</a></td>
</tr>
<tr>
<td></td>
<td><a href="http://www.Aspe.hhs.gov/daltcp/ongoing.htm">www.Aspe.hhs.gov/daltcp/ongoing.htm</a></td>
</tr>
<tr>
<td></td>
<td><a href="http://www.umich.edu/~hrswww">www.umich.edu/~hrswww</a></td>
</tr>
</tbody>
</table>