Should Disability Items in the Census Be Used for Planning Services for Elders?

Robert J. Calsyn, PhD, Joel P. Winter, PhD, Robert D. Yonker, MA
Should Disability Items in the Census Be Used for Planning Services for Elders?

Robert J. Calsyn, PhD,1,2 Joel P. Winter, PhD,2 and Robert D. Yonker, MA2

Purpose: The main goal of this study was to determine how well the disability questions of both the 1990 and 2000 Census correlated with a standard measure of disability. If the census questions were to correlate moderately well with a standard measure of disability, then Area Agencies on Aging (AAA) and other organizations would be able to use census information in estimating service needs for their catchment (service) area.

Design and Methods: Questionnaires containing both the census disability questions and a standard measure of disability were mailed to 4,508 older adults; 1,514 completed surveys were returned. In order to assess reliability, 472 of the respondents who completed the mail survey were reinterviewed by phone. All three disability measures were collapsed into the following three categories: no needs, instrumental activities of daily living (IADL) needs only, and activities of daily living (ADL) needs.

Results: All three disability measures exhibited moderate to good test–retest reliability. Using a standard measure of disability as the criterion, validity for the 1990 Census measure was quite low (Kappas of approximately 0.35). Validity for the 2000 Census measure was moderate to good (Kappas of approximately 0.60).

Implications: These results suggest that the 2000 Census disability questions may be sufficiently valid for planning purposes. However, additional research with more representative samples of older adults is needed.

Key Words: ADL, IADL, Planning, Needs assessment

Area Agencies on Aging (AAAs) and other human service agencies have routinely used census information to help them plan services for their catchment area. Census data have provided planners with estimates of the percentage of people in their catchment area who are elderly, live alone, and are poor. However, prior to 1990 there was nothing in the census data that provided planners with any information on respondents’ disability level, the single best indicator of service needs of elders (Harlow, 1992). This forced AAAs and other planning agencies to conduct their own needs assessment studies to assess the disability level of respondents living in their catchment area. Because these agencies typically could not afford to draw large samples, their population estimates of disability usually had a large margin of error. In addition, these studies typically have not been able to estimate service needs for catchment areas any smaller than a county.

The 1990 Census contained two disability items. One item asked about the ability of the respondent to perform various activities of daily living (ADLs) such as eating, bathing, and dressing; the other item assessed the ability to perform various instrumental activities of daily living (IADLs) such as the ability to go outside the home to shop or visit the doctor. Somewhat different disability items were included in the 2000 Census.

This study investigated the relationship of the 1990 and 2000 Census disability items with a longer criterion measure of disability that is widely used to estimate the functioning level of older adults. If the census disability measures correlate moderately well with the criterion measure of disability, administrators of AAAs and other agencies could rely more heavily on census data in their planning process. Disability information from the census data could be used to estimate the need for home-based services (such as home-delivered meals, housekeeping assistance, and home health care) for catchment areas as small as a census tract.

We are aware of only one study that compared the 1990 Census disability items with a standard measure of disability. In that study, Rodgers and Miller (1997) found that the census measure underestimated ADL functioning when compared to a criterion measure. That study did not report on the relationship between the 1990 Census IADL item and the criterion measure. To our knowledge, Andresen, Fitch, McLendon, and Meyers (2000) have conducted the only study that compared the 2000 Census disability measure with a longer criterion measure...
sure of disability. Andresen and colleagues reported Kappas of 0.02 to 0.37 between the 3 census disability items and 10 specific ADL questions and concluded that the 2000 Census measure was too inaccurate to be used for estimating disability. However, we believe that the method used by Andresen and colleagues for calculating agreement was too strict. They calculated agreement statistics between the more global census items and each of the specific items from the criterion measure. It is not surprising that the census items that refer to multiple ADLs in a single item did not correlate highly with any specific ADL item on the criterion measure.

We believe a more appropriate examination of the validity of the census items would be to create a categorical measure of disability from the criterion that is comparable to the following categories that can be created from the census questions: no need; IADL need only; and ADL need with or without an IADL need. Although more categories could be formed from the census questions, these three categories are based on a recognized hierarchy of needs that considers ADL problems more limiting than IADL problems (Ferrucci et al., 1998). For planning purposes, this tripartite classification of needs is sufficient because, in general, individuals with ADL problems will have more service needs than individuals who only have IADL problems. Thus, the current study investigated the validity of the 1990 and 2000 Census disability items compared to a criterion measure of disability with all three measures using a tripartite category system for classifying disability.

Methods

Sample and Procedure

We drew our sample from the 1996 voter registration list of St. Louis County, Missouri, because we could identify the age of each member of the population. We were willing to sacrifice some generalizability in order to ensure that each survey was mailed to an eligible participant (i.e., someone over age 65). A random sample of 4,508 older adults, aged 65 or older, was mailed the survey of disability measures. This represented 6.6% of the population of eligible voters in St. Louis County who were over age 65. Twenty-seven surveys (0.6%) were returned as undeliverable or respondent deceased; 1,514 respondents (33.6%) returned a completed survey (846 men and 668 women). Ninety-four percent of the sample were Caucasian.

In order to simulate the Content Reinterview Survey used by the Census Bureau for checking the reliability of their data (McNeil, 1993), 472 of the 1,514 respondents who completed our mail survey were readministered the same instrument via a phone interview within one month of the initial mail survey. In the cover letter sent with the mail survey, all respondents had been informed that they might be recontacted later for a phone interview. We believed it was essential to conduct the phone interviews because McNeil (1993) reported a Kappa of only 0.28 on the 1990 Census ADL item for agreement between mail and phone responses; Kappa for the IADL item was 0.53.

A professional survey research firm conducted the phone interviews, and respondents were paid $10 for the interview. The survey research firm had been given a quota of at least 200 interviews with men and 200 interviews with women, with a minimum sample of 450. No other restrictions were placed on them, including the number of required call-backs and such. In 269 cases no interview was completed because of a wrong number, a disconnected phone, or an answering machine. In 315 cases the phone was answered, but the interviewer was unable to talk to the desired respondent. In 204 instances the interview with the desired respondent was begun, but never completed. No calls were made to the other respondents from the mail sample, because the survey research firm had already reached its quota. The final phone sample consisted of 242 men and 230 women.

To further simulate the procedures used by the Census Bureau, both the mail and phone survey asked the respondent to answer the questions about the health needs of the oldest person living in his or her household. Because of this request, we had some respondents who answered questions about themselves (self) and some respondents who answered the questions about another person (proxy), usually a spouse. In the mail sample there were 1,140 self-respondents and 374 proxy respondents. In the phone sample there were 345 self-respondents and 127 proxy respondents.

Measures

1990 Census Disability Items.—The 1990 Census disability question had two parts and read as follows: “Because of a health condition that has lasted for 6 or more months, does this person have any difficulty . . .

A. Going outside the home alone, for example, to shop or visit a doctor’s office?
B. Taking care of his or her own personal needs, such as bathing, dressing, or getting around the home?”

A “yes” response to part A indicated an IADL problem and a “yes” response to part B indicated an ADL problem. For analysis purposes we created a categorical variable with three levels: no needs, IADL need only, ADL need with or without IADL need.

2000 Census Disability Items.—The 2000 Census contained the following disability items:

1. “Because of a physical, mental, or emotional condition lasting six months or more, does this person have any difficulty dressing, bathing, or getting around inside the home?”
2. “Does this person have a condition that substantially limits one or more basic physical activities such as walking, climbing stairs, reaching, lifting, or carrying?”

3. “Because of a physical, mental, or emotional condition lasting six months or more, does this person have difficulty going outside the home alone to shop or visit a doctor’s office?”

Questions 1 and 3 were similar, but not identical, to the questions used in the 1990 Census. A “yes” response to question 1 indicated an ADL problem. A “yes” response to questions 2 or 3 indicated an IADL problem. For analysis purposes we created a categorical variable with three levels: no needs, IADL need only, ADL need with or without ADL need.

**Criterion Disability Measure.**—The criterion measure of disability used in this study was a slightly modified version of the measure used in the National Health Interview Survey (Fitti & Kovar, 1987). The specific ADL and IADL content areas covered by the items are listed in Table 1. The stem for the items reads as follows:

“Does the oldest person living in the household have any difficulty doing the following tasks because of a health or physical problem?”

**Statistical Analyses**

To determine how well the two census measures of disability correlated with the criterion measure of disability, both Kappa and percent agreement statistics were calculated. Most researchers believe that agreement between measures is overestimated by percent agreement, but underestimated by Kappa, when the sample is unevenly distributed between the categories (Jones, Johnson, Butler, & Main, 1983). We calculated these agreement statistics for the following subgroups as well as for the total sample: (1) male versus female, (2) under age 75 versus 75 and older, and (3) self versus proxy.

**Results**

Table 1 displays information on the percentage of respondents in the mail sample who reported having difficulty performing any of the tasks on the criterion disability measure. Given that this was a community sample, it is not surprising that most participants reported few difficulties, particularly on the ADL items. Trouble doing heavy housekeeping (28.2%) was the only item for which a significant percentage of respondents reported needing help. This item was eliminated from the calculation of the categorical measure of disability for the criterion measure. We believed that this was justified because AAAs were not likely to consider an older adult who only needed assistance with heavy housekeeping as eligible for publicly funded services.

Table 2 displays the contingency tables for the relationship between the 1990 and 2000 Census disability variables as well as the relationship between both the 1990 and 2000 Census disability variables with the criterion measure of disability, scored as categorical variables, for the entire mail sample. Because of missing data, sample size was reduced to 1,377. Kappa for the relationship between the two census measures was 0.45 and the percent agreement was 80%. Most of the disagreement between the 1990 and 2000 Census measures was in the IADL-only category, with the 1990 Census measure only classifying 6% of the sample in this category and the 2000 Census measure classifying 22% of the sample in this category. Given that the 2000 Census measure had an additional IADL item, this is not surprising.
Examination of the relationship of 1990 Census measure with the criterion measure reveals that most of the errors (i.e., the off-diagonal values) made by the 1990 measure are underestimates of IADL disability. About 24% of the sample was rated as having an IADL disability using the criterion, but only 6% of the sample was rated as having an IADL disability using the 1990 Census measure. On the other hand, 68% of the sample was classified as having no needs using the criterion measure of disability, but 88% of the sample was classified as having no needs using the 1990 Census measure. Kappa was 0.36 and percent agreement was 75%.

The 2000 measure was a significant improvement over the 1990 measure; Kappa was 0.60 and percent agreement was 82% for the relationship between the 2000 Census disability measure and the criterion. Moreover, although the 2000 Census measure still underestimates disability slightly, the errors made using the 2000 measure more evenly distributed between under- and overestimates of disability.

Table 3 contains the Kappa and percent agreement statistics for the relationship between the criterion and the 1990 and 2000 Census disability variables for various subgroups. With the exception of the proxy-mail sample, the Kappas for the 1990 Census were low, approximately 0.35. On the other hand, the Kappas for the various subgroups for the 2000 Census were in the moderate range, approximately 0.60. With two exceptions, subgroup differences in Kappa were minimal. Kappas were lower for the phone survey than the mail survey, and Kappas for self-respondents were lower than for proxies. In general, the percent agreement statistics for both the 1990 and 2000 Census ranged between 75% and 85%, with the values for the 2000 Census being higher.

Table 3. Agreement Between Criterion and Other Disability Measures for Various Subgroups

<table>
<thead>
<tr>
<th>Group</th>
<th>1990</th>
<th>2000</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Kappa % Agree</td>
</tr>
<tr>
<td>Total—mail</td>
<td>1,377</td>
<td>.36 75</td>
</tr>
<tr>
<td>Total—phone</td>
<td>439</td>
<td>.31 74</td>
</tr>
<tr>
<td>Men—mail</td>
<td>775</td>
<td>.35 80</td>
</tr>
<tr>
<td>Women—mail</td>
<td>595</td>
<td>.35 69</td>
</tr>
<tr>
<td>Age &lt;75—mail</td>
<td>800</td>
<td>.25 80</td>
</tr>
<tr>
<td>Age 75+—mail</td>
<td>577</td>
<td>.40 68</td>
</tr>
<tr>
<td>Self—mail</td>
<td>1,031</td>
<td>.28 74</td>
</tr>
<tr>
<td>Proxy—mail</td>
<td>338</td>
<td>.53 78</td>
</tr>
</tbody>
</table>

*Totals do not always sum to 1,377 because of missing data on some variables for a few respondents.

Kappa = 0.59 for the 2000 measure. Given that respondents’ disability levels could have truly changed in the interval between the return of the mail survey and the phone interview, and given that a different measurement modality was employed (mail vs phone), these reliability coefficients are quite respectable. It is also important to note that there was no significant difference between the mean number of difficulties on the criterion measure as a function of whether the data were collected by mail (mean = 0.96) or by phone (mean = 0.99).

Discussion

Past and Future Research

Consistent with the work of Rodgers and Miller (1997), this study also found that the 1990 Census disability questions underestimate disability when compared to a standard ADL/IADL measure. In contrast, this study found that the 2000 Census disability items had adequate reliability and validity. This conclusion is at variance with that of Andresen and colleagues (2000), who used a different approach to measuring agreement. We believe that our approach, which formed comparable categories for both the criterion measure and the census measures, was a more appropriate test of agreement than the approach of Andresen and colleagues, who calculated agreement between the global census questions and specific ADL/IADL items that were not comparable in their scope of inquiry. However, it is important to remind the reader that because the census disability categories are global measures, they do not provide information about specific ADL and IADL needs. If knowledge about functioning on a specific ADL/IADL item is needed, then the broad census disability categories would be inadequate.

Clearly, more research is needed with different samples and different criterion measures of disability before a definitive conclusion can be made on the value of census data in planning services for elderly people. Not only should future researchers try to replicate our results, but they should try to expand them by seeing how well the 2000 Census disability variables predict other important variables such as service utilization and service awareness (Calsyn & Roades, 1993; Calsyn, Roades, & Klinkenberg, 1998; Wolinsky & Johnson, 1991). It is also important to find out if the single item on cognitive functioning that was added to the 2000 Census correlates with longer, more widely accepted measures of cognitive functioning. Our study did not include a standard measure of cognitive functioning, so we are unable to make a recommendation with regard to the utility of this new item in planning services for the elderly population.

Study Limitations

The generalizability of the results of this study are limited by a number of factors. First, our sample only included registered voters, primarily Caucasian,
from one urban county of the Midwest. Moreover, only about a third of the sample returned our mail questionnaire. Thus, our results may not generalize to other samples, particularly minority elders, frail elders, and those older adults who may be less conscientious in returning surveys. Similarly, because our protocol for the phone sample did not require interviewers to contact specific respondents, the phone sample may have underrepresented older adults with hearing problems and more frail elders who are less likely to answer the phone. In addition, the most active older adults may also have been underrepresented in our phone sample because they were less likely to be at home when called.

Second, although we tried to simulate the procedures used by the Census Bureau in conducting both the mail survey and the phone interview, we cannot know with any certainty whether our data collection procedures were more or less reliable than the census procedures. One could speculate that our procedures might produce somewhat more reliable data because our instruments only focused on disability information, whereas the census instruments asked about a wider range of issues. Third, we used only one criterion measure of disability. Rodgers and Miller (1997) have shown that different measures of disability used in various national samples have produced different estimates of the number of respondents who are disabled. Thus, the results of our study might have been somewhat different if we had chosen a different criterion measure of disability.

Implications for Service Planning

The data from the phone sample indicated that the 1990 and 2000 Census measures had moderate to good reliability, equal to that of the longer criterion measure of disability. With regard to validity, the 2000 Census measure of disability was clearly superior to the 1990 Census measure. We believe that both the percent agreement and Kappa values for the relationship between the 2000 disability measure and the criterion measure of disability were high enough in this study that AAAs and other planning groups should consider using the 2000 Census disability data for estimating service needs in their catchment area. Even though the 2000 disability measure slightly underestimates disability, and it is not as accurate as more conventional ADL/IADL measures, the ability of the 2000 Census data to provide estimates of disability for small geographic areas far outweighs the loss of precision. Because the disability items only appear on the “long form,” estimates of disability for areas smaller than census tracts will probably have large standard errors. However, this is a vast improvement over the typical needs assessment, which can only provide estimates of need for large geographic areas such as counties. In addition, the vast array of other information in the census files allows planners to make separate estimates of disability for various subgroups of elders, categorized according to dimensions including the following: age, race, income, and living situation.

Planners should also consider using one of the sophisticated computer programs that are now available for analyzing census data (e.g., ARCVIEW, 2000; GeoStats, 2000). Using this type of computer software would allow planners to create maps based on census information that could pinpoint those areas within an agency’s catchment area that had: (1) the highest need for assistance with ADL/IADL needs; (2) the highest poverty levels; (3) the highest percentage of elders living alone, and such. Technical assistance in using this software is usually available from most major universities, because university researchers in many fields are now doing research using Geographic Information Systems (GIS). Most of the Census 2000 data that planners would want should be available in Summary File 3, which is scheduled for release in the summer of 2002.

Despite our enthusiasm for the utility of the 2000 Census for planning services for elders, it is important to note the limitations of census data. First, the utility of the 2000 Census for planning purposes will dissipate over time, particularly for those geographic areas that experience a greater migration of inhabitants. Also, because the census disability categories only classify respondents as having ADL or IADL needs, agency planners cannot glean from the census data the specific ADLs and IADLs for which their older adult population needs assistance. In addition, the census disability questions only measure intrinsic disability (i.e., the ability of the individual to perform the activity without assistance). Consequently, service planners cannot obtain any information from census data on residual disability (i.e., the ability of an individual to perform an activity with assistance from equipment or another individual), which is also important in planning services (Freedman, 2000). Finally, the census data provide no information regarding the following issues, which are important in planning services for the elderly population: service awareness, prior service utilization, and perceived barriers to receiving services. Thus, census data cannot totally replace conventional needs assessment studies. Nevertheless, the 2000 Census does provide AAAs and others with meaningful information on disability levels and other social indicators that describe the older adults living in their catchment area.

References


Received January 31, 2001
Accepted June 4, 2001
Decision Editor: Laurence G. Branch, PhD