

Reported activities of daily living: agreement between elderly subjects with and without dementia and their caregivers

TRULS ØSTBYE^{1,2}, SUZANNE TYAS¹, IAN McDOWELL⁴, JOHN KOVAL^{1,3}

¹Department of Epidemiology and Biostatistics, The University of Western Ontario, London, ON N6A 5C1, Canada

²Department of Family Medicine and ³Department of Occupational Therapy, The University of Western Ontario, Canada

⁴Department of Epidemiology and Community Medicine, University of Ottawa, Canada

Address correspondence to: Truls Østbye. Fax: (+519) 661 3766; email: ostbye@uwo.ca

Abstract

Objectives: to determine how accurately information on disability provided by a caregiver (proxy respondent) reflected the opinion of subjects themselves, and if this agreement varied by severity of dementia or relationship of the caregiver to the subject.

Setting and participants: the study was based on data from the Canadian Study of Health and Aging, a multicentre study of dementia and health of Canadians age 65 and over. Eight hundred study subjects and their caregivers were independently interviewed regarding the subjects' activities of daily living (ADL).

Measurements: the percentage of subjects who were independent for individual ADL items and the agreement in these reports between subjects and caregivers were investigated using three-level κ statistics.

Results: index subjects with caregivers other than spouses or offspring required more assistance with ADL. The reported percentage of independence decreased with increasing severity of dementia. There was more agreement between self- and proxy-reported level of independence for physical ADL than for instrumental ADL items. Agreement decreased with increasing severity of dementia. Few statistically significant differences were noted between level of agreement and caregiver relationship.

Conclusion: satisfactory levels of agreement on ADL between cognitively normal subjects and their caregivers indicate that proxy respondents are a reasonable source of information on ADL when data collection from the subjects themselves is not feasible. Since agreement decreases as the severity of dementia increases, caregiver reports may be preferred for elderly patients even with mild dementia in order to facilitate longitudinal assessment of ADL ratings as the dementia progresses.

Keywords: activities of daily living, proxy respondent, data collection methods, dementia, caregivers

Introduction

For diseases such as dementia there is no 'gold standard' source of information: both proxy- and self-reported information are problematic. The general objective of this study was to determine the level of agreement between caregivers (proxy respondents) and elderly people (index subjects) on reported disability, and to assess whether this agreement varied by severity of dementia or relationship of the caregiver to the subject.

Functional status and activities of daily living

Functional status can be defined as the degree to which an individual is able to perform socially allocated roles

free of physical or mental limitations [1]. The most fundamental indicators of functional status assess ability to perform self-care activities, usually known as activities of daily living (ADL). Although originally developed for individuals with substantial disability in long-term care institutions, ADL measures have also been widely used for elderly people living in the community, where they indicate the ability to perform some of the tasks needed for continued independent living [2]. Most surveys have found that the prevalence of ADL disability increases with age [3].

A number of different ADL scales have been developed [4, 5]. The Older Americans Resources and Services (OARS) ADL scale was used in the Canadian

Study of Health and Aging (CSHA), as this scale has been well validated and has been used in population surveys elsewhere [2, 6]. The scale consists of 14 questions divided into two domains: instrumental ADL and physical (self-maintenance) ADL.

Surrogate sources of information

Surrogate sources of information are required when an individual (sometimes termed the index subject) cannot easily provide information personally. Such alternative or proxy respondents may include relatives, friends, neighbours or health care providers.

The level of agreement between the index subject and the proxy respondent may be influenced by characteristics of either the index subject or the proxy. Although the relationship between the subject and the proxy has been thought intuitively to be important, several studies have found little effect of type of relationship on level of agreement [7-10]. Some studies have shown that non-response rates, however, are influenced by the relationship of the proxy respondent to the index subject [11], with spouses showing the lowest non-response [12]. The case or control status of the index subject (i.e., whether the subject has the disease in question or belongs to a comparison or 'control' group) appears to have little or no effect on the level of agreement [8, 10, 13]. However, the disease studied seems to be associated with the level of agreement achieved, with high agreement reported for heart disease and low agreement reported for back pain [14]. The index subject's cognitive status has been reported to be unrelated to the level of agreement between self and proxy reports [12].

Materials and methods

Objectives

The general objective of this study was to determine how accurately information on disability provided by caregivers (proxy respondents) reflects the opinion of the elderly people (index subjects) themselves. Specifically, are there some ADL items for which information from the subjects and the caregivers is very different; does the level of agreement vary by severity of dementia; and is information collected from spouses more accurate than that collected from other family members or friends?

Data source: the CSHA

The CSHA was a national, multicentre study conducted between February 1991 and May 1992. The methods of the main study have been discussed elsewhere in detail [15]. A general survey was administered to a

representative sample ($n = 8949$) of Canadians 65 years old and over living in the community.

Also interviewed were 316 caregivers of patients with dementia and a comparison group of 484 individuals responsible for providing care to someone without dementia (including *potential* caregivers for those in the comparison group who did not need a caregiver [16]). For the 800 patient/caregiver (index subject/proxy respondent) pairs, each member of the pair was interviewed separately. This group is the focus of the present analyses.

The OARS ADL scale [2] was included in both the patient and caregiver questionnaires. The interview questions for the two types of respondents were identical, except that the questions in the patient interview were self-reported, while the questions in the caregiver interview were rephrased to refer to the patient (index subject). Questions were examined individually and a five-category total 'ADL rating', indicating the amount of assistance required, was calculated [2]. This scale rates assistance required for basic activities such as self-care more heavily than assistance for reading or writing. Scores range from 2 (excellent or good level of function) to 6 (total impairment).

The level of agreement between caregiver and index subject was examined with respect to the severity of the subject's dementia and the relationship of the caregiver to the subject. The criteria for the diagnosis and severity of dementia followed DSM-III-R criteria [17]. Patients with severe dementia were excluded from the analysis because of the small number of these cases ($n = 17$), high variability and suspicion that these patients received assistance in completing their ADL 'self' reports. Caregivers were classified as spouses, offspring or 'other'. Caregivers in the 'other' category included friends, neighbours, volunteers, and relatives such as siblings or grandchildren, but excluded all paid help such as nurses or homemakers.

Analyses

Average ADL ratings and the percentage of elderly people and their caregivers who reported independence for individual ADL items provide a general description of the level of agreement found. Significant differences in percentages as reported by subjects and caregivers for individual ADL items were assessed using McNemar's test for matched pairs. Analysis of variance was used to compare self-reported and caregiver-reported ADL ratings by severity of dementia and by caregiver relationship to the subject.

κ statistics were used in this study to summarize level of agreement. κ is defined as the level of agreement corrected for chance, with $\kappa = 1$ indicating perfect agreement, $\kappa = 0$ agreement expected by chance and $\kappa < 0$ agreement less than would be expected by

Reported activities of daily living

Table 1a. Percentage of elderly who can manage each activities of daily living (ADL) item independently by presence and severity of dementia as reported by self (S) and by caregiver (C)

| ADL item | Subjects without dementia (<i>n</i> = 479-484) | | Subjects with dementia | | | |
|------------------------------------|--|--------|----------------------------|---------|--------------------------------|---------|
| | S | C | Mild (<i>n</i> = 143-146) | | Moderate (<i>n</i> = 144-150) | |
| | S | C | S | C | S | C |
| Physical | | | | | | |
| Eating | 99.4 | 98.6 | 96.6 | 96.6 | 92.0 | 85.3 |
| Dressing | 97.1 | 97.3 | 91.0 | 89.0 | 85.3 | 63.3*** |
| Personal care | 97.9 | 97.5 | 91.1 | 95.2 | 86.7 | 71.3*** |
| Walking | 91.1 | 94.0* | 87.0 | 85.6 | 80.0 | 79.3 |
| Getting out of bed | 98.6 | 98.6 | 95.9 | 94.5 | 90.7 | 91.3 |
| Taking bath | 84.3 | 88.2** | 74.7 | 65.1** | 47.3 | 37.2 |
| Using toilet | 99.2 | 99.4 | 95.9 | 94.5 | 91.3 | 85.2 |
| Instrumental | | | | | | |
| Getting to distant places | 85.6 | 82.3 | 61.4 | 41.4*** | 38.9 | 33.3*** |
| Using telephone | 91.9 | 93.2 | 79.5 | 76.0 | 66.7 | 41.3*** |
| Going shopping | 76.2 | 75.4 | 50.1 | 38.2** | 27.5 | 10.1*** |
| Preparing own meals | 89.0 | 84.0** | 69.7 | 43.5*** | 38.9 | 16.7*** |
| Doing housework | 60.4 | 58.8 | 38.2 | 27.8* | 22.8 | 11.7** |
| Taking medicine | 96.1 | 93.8* | 81.1 | 58.7*** | 51.7 | 25.9*** |
| Managing own money | 94.2 | 91.5 | 70.0 | 42.2*** | 42.2 | 12.9*** |
| Mean ADL rating^a | 2.71 | 2.75 | 3.43 | 3.93 | 4.21 | 4.99 |

^aADL rating: 2, excellent/good function; 3, mild impairment; 4, moderate impairment; 5, severe impairment; 6, total impairment. McNemar's test for matched pairs (for individual ADL items); **P* < 0.05; ***P* < 0.01; ****P* < 0.001.

chance [18,19]. Fleiss [20] summarized guidelines to the interpretation of κ values, reporting that κ values exceeding 0.75 represent 'excellent' agreement, values between 0.4 and 0.75 'fair to good' agreement, and values less than 0.4 'poor' agreement. It has been recommended that measures with κ values less than 0.4 not be used [21].

Weighted κ statistics were calculated using a Fortran program which evaluated κ and its standard error using a formula from Fleiss [20]. The weights are those proposed by Cohen [22] and Cicchetti and Allison [23]. Two-level κ statistics for binary outcomes (able to perform a specific task independently or with some assistance/unable to perform a task) and three-level κ for ordinal outcomes (able to perform a task independently/requires some assistance/unable to perform task) were calculated. The two-level κ statistics showed the same pattern and level of agreement as the three-level κ and are not presented here.

Responses of 'not applicable' were treated as missing data. Completeness of data was assessed by calculating the percentage of missing values for each question. Non-response was so low for both proxy and self reports (<0.5%) that further analyses to compare the rate of missing values among the groups were not conducted. As no correction for multiple testing was

made, *P*-values of borderline significance should be interpreted with caution.

Results

In Table 1a, the percentage of subjects who could manage each ADL item independently, according to self or caregiver report, is presented by presence and severity of dementia. The percentages vary depending on the individual ADL item, but independence is generally higher for physical ADL items than for instrumental ADL items in both caregiver and self reports.

The difference in percentages between self and caregiver reports of independence for individual ADL items, assessed by McNemar's test for matched pairs, showed that caregivers were more likely to rate the subjects as requiring assistance than were the subjects themselves. There was also less agreement for instrumental ADL than for physical ADL items. The percentages were most different for pairs where the subject was suffering from dementia, and agreement between self and caregiver reports decreased with increasing severity of dementia.

Using analysis of variance, the ADL rating reflected decreased independence with increasing severity of

Table 1b. Percentage of elderly who can manage each activities of daily living (ADL) item independently by caregiver relationship

| ADL item | Carer | | | | | |
|------------------------------------|------------------------------|---------|------------------------------------|-----------|-----------------------------|--------------------|
| | Spouse (<i>n</i> = 279-283) | | Son/daughter (<i>n</i> = 342-347) | | Other (<i>n</i> = 146-150) | |
| | Self | Spouse | Self | Offspring | Self | Other ^a |
| Physical | | | | | | |
| Eating | 96.5 | 95.1 | 98.9 | 95.4** | 96.0 | 96.0 |
| Dressing | 94.0 | 89.4* | 93.7 | 90.2* | 90.7 | 87.3 |
| Personal care | 95.8 | 94.0 | 93.7 | 90.5* | 92.0 | 92.0 |
| Walking | 92.6 | 92.6 | 86.7 | 88.2 | 82.7 | 86.7 |
| Getting out of bed | 97.5 | 96.5 | 96.0 | 97.1 | 95.3 | 94.0 |
| Taking bath | 84.1 | 80.9 | 70.9 | 72.6 | 70.1 | 67.4 |
| Using toilet | 98.9 | 97.5 | 96.3 | 94.5 | 94.0 | 94.7 |
| Instrumental | | | | | | |
| Getting to distant places | 85.9 | 78.8** | 84.1 | 79.2* | 81.3 | 83.3 |
| Using telephone | 79.3 | 72.9** | 69.6 | 56.2*** | 64.4 | 53.0 |
| Going shopping | 73.8 | 65.6*** | 56.5 | 51.6* | 53.0 | 48.3 |
| Preparing own meals | 77.7 | 66.1*** | 77.2 | 66.4*** | 69.2 | 54.1*** |
| Doing housework | 61.4 | 53.6** | 42.9 | 39.4 | 41.5 | 36.7 |
| Taking medicine | 85.5 | 74.6*** | 84.6 | 75.4*** | 83.9 | 72.5*** |
| Managing own money | 83.3 | 73.1*** | 77.8 | 64.7*** | 74.3 | 61.5*** |
| Mean ADL rating^b | 2.91 | 3.22 | 3.21 | 3.44 | 3.35 | 3.58 |

^aExcludes paid caregivers.

^bADL rating: 2, excellent/good function; 3, mild impairment; 4, moderate impairment; 5, severe impairment; 6, total impairment. McNemar's test for matched pairs (for individual ADL items); * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

dementia when reported both by subjects ($F = 123.0$, 2 d.f., $P < 0.0001$) and their caregivers ($F = 296.1$, 2 d.f., $P < 0.0001$). Since the ADL rating summarizes the results of individual ADL items, the total score obtained from caregiver reports again reflected lower levels of independence than that derived from self reports.

Similar results were found when the data were categorized by the relationship of the caregiver to the index subject (Table 1b). Significant disagreement was noted more frequently for instrumental ADL items than for physical ADL items. Where there was statistically significant disagreement between self and caregiver reports, caregivers reported lower rates of independence than index subjects. ADL ratings were higher (i.e., the index subjects were judged less independent) for subjects with caregivers other than spouses or offspring, both when reported by the index subject ($F = 7.84$, 2 d.f., $P = 0.0004$) and by the caregiver ($F = 3.92$, 2 d.f., $P = 0.02$). The possibility that this difference in ADL ratings might reflect a variation in severity of dementia across different caregiver groups was examined. An analysis of caregiver relationship (three levels) by dementia severity (three levels) showed no statistically significant association ($\chi^2 = 3.18$, 4 d.f., $P = 0.53$).

The level of agreement between self and caregiver

reports for the ADL rating was not high in any of the groups, although it was higher in the comparison group without dementia than in patients with dementia (Table 2a). A gradient based on severity of dementia was observed: subjects with moderate dementia showed less agreement with their caregivers than those with mild dementia who, in turn, showed less agreement than those in the comparison group.

Few statistically significant differences were noted between level of agreement and caregiver relationship (Table 2b). The only significant differences were found for 'eating', 'getting out of bed', 'using toilet' and 'managing own money'.

Discussion

Overall, the level of agreement between index subject and caregiver reports was reasonably high for elderly people without dementia, with a κ for ADL ratings of 0.55 (fair to good agreement), decreasing to 0.40 for subjects with mild dementia and 0.34 (poor agreement) for subjects with moderate dementia. Very few ADL items showed excellent agreement. There was also a considerable range in agreement across different ADL items.

Caregiver respondents tended to report lower levels

Table 2a. Agreement^a on activities of daily living (ADL) items and ADL rating between self and caregiver reports by presence and severity of dementia

| ADL item | Subjects and caregivers | | |
|---------------------------|---|----------------------------|--------------------------------|
| | Without dementia (<i>n</i> = 479–484) | Subjects with dementia | |
| | | Mild (<i>n</i> = 140–146) | Moderate (<i>n</i> = 142–150) |
| Physical | | | |
| Eating | 0.50 (0.42–0.57) | 0.34 (0.20–0.49) | 0.27 (0.12–0.42) |
| Dressing | 0.54 (0.46–0.62) | 0.46 (0.27–0.64) | 0.31 (0.19–0.42) |
| Personal care | 0.39 (0.31–0.47) | 0.51 (0.39–0.64) | 0.29 (0.16–0.41) |
| Walking | 0.51 (0.43–0.60) | 0.76 (0.62–0.90) | 0.35 (0.21–0.48) |
| Getting out of bed | 0.26 (0.17–0.34) | 0.57 (0.44–0.70) | 0.42 (0.18–0.66) |
| Taking bath | 0.58 (0.50–0.66) | 0.58 (0.45–0.70) | 0.42 (0.30–0.53) |
| Using toilet | – ^b | 0.81 (0.67–0.95) | 0.41 (0.28–0.54) |
| Instrumental | | | |
| Getting to distant places | 0.43 (0.35–0.51) | 0.41 (0.29–0.53) | 0.25 (0.15–0.34) |
| Using telephone | 0.50 (0.42–0.59) | 0.39 (0.25–0.52) | 0.32 (0.21–0.43) |
| Going shopping | 0.61 (0.54–0.69) | 0.50 (0.38–0.63) | 0.33 (0.22–0.44) |
| Preparing own meals | 0.51 (0.43–0.58) | 0.36 (0.24–0.48) | 0.31 (0.20–0.42) |
| Doing housework | 0.57 (0.49–0.64) | 0.39 (0.27–0.51) | 0.37 (0.26–0.49) |
| Taking medicine | 0.42 (0.33–0.50) | 0.37 (0.23–0.50) | 0.28 (0.18–0.39) |
| Managing own money | 0.36 (0.28–0.43) | 0.27 (0.16–0.39) | 0.18 (0.09–0.28) |
| ADL rating | 0.55 (0.48–0.62) | 0.40 (0.30–0.51) | 0.34 (0.30–0.39) |

^aAs determined by three-level weighted κ for the individual ADL items, five-level weighted κ for ADL rating (95% confidence intervals in parentheses).

^bNot reported due to presence of empty cell.

of independence compared with index subjects' self reports, confirming the results of previous studies. The relationship of the caregiver to the subject did not appear to make an important difference: our results thus suggest that the reporting caregiver can be chosen on the basis of other characteristics, such as availability.

For diseases such as dementia there is no 'gold standard' source of information: both proxy- and self-reported information are open to question. The issue of whether even the concept of a gold standard is relevant for a partially subjective indicator can be raised; if ADL scales measure, at least partially, an individual's perception of their own or someone else's disability, can this perception be objectively assessed as incorrect? The quality of ADL information was assessed in the current study without referral to a gold standard by examining the agreement between information provided by the index subject and that reported by family members or other caregivers.

The influence of severity of dementia was summarized in Tables 1a and 2a. Since level of dependence is one of the factors used in assessing dementia severity, the decrease in independence with increasing severity of dementia seen in Table 1a is to be expected. This loss of independence is more obvious from the total ADL ratings than from the individual ADL items.

Conversely, since disability increases with severity of dementia, disability may arguably be a good indicator of dementia severity.

In Table 1b, the fewer statistically significant differences between index subjects and 'other' caregivers than between subjects and spouse or offspring caregivers may be partially a function of the smaller size of the 'other' caregiver group. For example, in the ADL item 'using phone', the absolute difference in the percentage who reported independence between index subjects and their spouses is a statistically significant 6.4%, whereas the 11.4% difference between self and 'other' caregiver reports fails to reach statistical significance.

Average ADL ratings in Table 1b indicate more dependency for index subjects with 'other' caregivers. This poorer functional status of those with caregivers other than spouses or offspring does not appear to be due to differences in dementia severity, but may be due to circumstances related to the lack of immediate family members as caregivers.

Fewer statistically significant differences between caregiver- and self-reported percentages of independence were noted for physical ADL than for instrumental ADL items (Tables 1a and 1b). A consistent finding in the literature is greater agreement for easily

Table 2b. Agreement^a on activities of daily living (ADL) items and ADL rating between self and caregiver reports by caregiver relationship

| ADL item | Caregiver | | |
|---------------------------|------------------------------|------------------------------------|--|
| | Spouse (<i>n</i> = 278–283) | Son/daughter (<i>n</i> = 340–347) | Other (<i>n</i> = 145–150) ^b |
| Physical | | | |
| Eating | 0.65 (0.55–0.75) | – ^c | 0.31 (0.15–0.47) |
| Dressing | 0.43 (0.33–0.53) | 0.49 (0.40–0.58) | 0.55 (0.42–0.68) |
| Personal care | 0.44 (0.34–0.54) | 0.46 (0.37–0.55) | 0.43 (0.30–0.56) |
| Walking | 0.61 (0.50–0.72) | 0.48 (0.39–0.57) | 0.56 (0.42–0.70) |
| Getting out of bed | 0.28 (0.18–0.38) | 0.56 (0.47–0.65) | 0.58 (0.45–0.71) |
| Taking bath | 0.61 (0.51–0.71) | 0.55 (0.46–0.64) | 0.63 (0.49–0.77) |
| Using toilet | 0.40 (0.32–0.48) | 0.55 (0.46–0.64) | 0.72 (0.57–0.87) |
| Instrumental | | | |
| Getting to distant places | 0.60 (0.50–0.70) | 0.44 (0.36–0.52) | 0.45 (0.32–0.58) |
| Using telephone | 0.55 (0.46–0.64) | 0.45 (0.36–0.54) | 0.46 (0.32–0.60) |
| Going shopping | 0.60 (0.50–0.70) | 0.61 (0.53–0.69) | 0.59 (0.46–0.72) |
| Preparing own meals | 0.57 (0.48–0.66) | 0.47 (0.39–0.55) | 0.58 (0.45–0.71) |
| Doing housework | 0.50 (0.41–0.59) | 0.53 (0.45–0.61) | 0.62 (0.50–0.74) |
| Taking medicine | 0.52 (0.43–0.61) | 0.48 (0.40–0.56) | 0.50 (0.35–0.65) |
| Managing own money | 0.61 (0.51–0.71) | 0.40 (0.32–0.48) | 0.52 (0.39–0.65) |
| ADL rating | 0.60 (0.52–0.68) | 0.53 (0.46–0.60) | 0.64 (0.53–0.75) |

^aAs determined by three-level weighted κ for individual ADL items, five-level weighted κ for ADL rating (95% confidence intervals in parentheses).

^bExcludes paid caregivers.

^cNot reported due to presence of empty cell.

observable, objective items [24, 25] and thus physical ADL items might be expected to generally show higher levels of agreement between self and proxy reports. A more skewed distribution also leads to higher agreement.

Agreement declined with increasing severity of dementia, suggesting that caregiver reports are preferable to self reports when the index subject has more than mild dementia. This result differs from that of a previous study [12] which found little relationship between cognitive status of the index subject and level of agreement. Cognitive impairment in that study, however, was conservatively measured on the basis of a few questions in a population of cancer patients, and could have classified subjects with mild or no dementia as impaired.

There was somewhat better agreement in Table 2a for physical ADL items 'walking', 'taking bath' and 'using toilet', and the instrumental ADL items 'going shopping' and 'doing housework'. The item 'managing own money' was the item with the lowest agreement ('poor' agreement for all groups). These results are consistent with a previous study which reported the level of agreement between proxy and self reports to be highest for 'walking' and lowest for 'managing own money' [26].

There were few significant differences in agreement

between index subjects and caregivers by relationship of the caregiver to the subject for the individual ADL items (Table 2b). Although agreement measured by κ statistics was slightly higher for ADL rating in 'other' caregivers, this difference was not statistically significant. It is possible that some characteristic of the caregivers other than their formal relationship to the subject (e.g. amount of time spent with the subject or even the 'quality' of the relationship) could influence agreement more strongly. The finding, however, of no significant association between caregiver relationship and agreement corroborates the results of previous studies examining agreement on outcomes other than ADL ratings [7–10]. Previous findings of an association between relationship of the proxy respondent to the index subject and non-response rates [11,12] could not be examined in this study because of low non-response rates.

Conclusion

For elderly individuals without dementia, it is most reasonable to collect ADL information from the index subjects themselves, although caregiver respondents may sometimes be used for reasons of convenience. Since agreement decreases with increasing severity of dementia, caregiver respondents may be preferred for

elderly patients even with mild dementia in order to facilitate comparison of ADL ratings over time as the dementia progresses.

Key points

- Most older people, including those with dementia, are physically able.
- Self-reported information on physical disability is appropriate in non-demented elderly people.
- There is relatively weak agreement between non-demented old people and their caregivers about whether they need help with everyday activities.
- This disagreement is greater when the patient suffers from dementia.

Acknowledgements

The Canadian Study of Health and Aging was funded by the Seniors' Independence Research Program, with funds administered by the National Health Research and Development Program (project 6606-3954 MC[S]). Personal support for Suzanne Tyas was provided by Health Canada through an NHRDP Research Training Award.

References

1. Bowling A. *Measuring Health. A Review of Quality of Life Measurement Scales.* Milton Keynes: Open University Press, 1991.
2. Fillenbaum GG. *Multidimensional Functional Assessment of Older Adults. The Duke Older Americans Resources and Services Procedures.* Hillsdale, NJ: Lawrence, Erlbaum, 1988.
3. Fulton JP, Katz S, Jack SS, Hendershot GE. *Physical Functioning of the Aged (United States 1984).* National Center for Health Statistics, Vital Health Statistics Series 10, No. 167. DHHS Publication No. (PHS) 89-1595, Hyattsville, MD, 1989.
4. Katz SC, Ford AB, Moskowitz RW, *et al.* Studies of illness in the aged. The index of ADL: a standardized measure of biological and psychosocial function. *JAMA* 1963; 185: 914-9.
5. Branch LG, Meyers AR. Assessing physical function in the elderly. *Clin Geriatr Med* 1987; 3: 29-51.
6. McDowell I, Newell C. *Measuring Health: a Guide to Rating Scales and Questionnaires.* New York: Oxford University Press, 1987; 303-4.
7. Kolonel LN, Hirohata T, Nomura AMY. Adequacy of survey data collected from substitute respondents. *Am J Epidemiol* 1977; 106: 476-84.
8. McLaughlin JK, Mandel JS, Mehl ES, Blot WJ. Comparison of next-of-kin with self-respondents regarding questions on cigarette, coffee, and alcohol consumption. *Epidemiology* 1990; 1: 408-12.

9. Hatch MC, Misra D, Kabat GC, Kartzmer S. Proxy respondents in reproductive research: a comparison of self and partner-reported data. *Am J Epidemiol* 1991; 133: 826-31.
10. Graham P, Jackson R. Primary versus proxy respondents: comparability of questionnaire data on alcohol consumption. *Am J Epidemiol* 1993; 138: 443-52.
11. Pickle LW, Brown LM, Blot WJ. Information available from respondents in case-control interview studies. *Am J Epidemiol* 1983; 118: 99-108.
12. Farrow D, Samet J. Comparability of information provided by elderly cancer patients and surrogates regarding health and functional status, social network, and life events. *Epidemiology* 1990; 1: 370-6.
13. Marshall J, Priore R, Haughey B, Rzepka T, Graham S. Spouse-subject interviews and the reliability of diet studies. *Am J Epidemiol* 1980; 112: 675-83.
14. Halabi S, Zurayk H, Awaidda R, Darwish M, Saab B. Reliability and validity of self and proxy reporting of morbidity data: a case study from Beirut, Lebanon. *Int J Epidemiol* 1992; 21: 607-12.
15. Canadian Study of Health and Aging Working Group. Canadian Study of Health and Aging: study methods and prevalence of dementia. *Can Med Assoc J* 1994; 150: 899-913.
16. The Canadian Study of Health and Aging. Patterns of caring for people with dementia in Canada. *Can J Aging* 1994; 13: 470-87.
17. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, 3rd ed, revised (DSM-III-R).* Washington, DC: American Psychiatric Association, 1987; 107.
18. Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; i: 307-10.
19. Nelson LM, Longstreth WT, Koepsell TD, Checkoway H, van Belle G. Completeness and accuracy of interview data from proxy respondents: demographic, medical, and life-style factors. *Epidemiology* 1994; 5: 204-17.
20. Fleiss JL. *Statistical Methods for Rates and Proportions.* New York: John Wiley and Sons, 1981; 213-36.
21. Armitage P, Berry G. *Statistical Methods in Medical Research.* Oxford: Blackwell Science, 1994; 447.
22. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull* 1968; 70: 213-20.
23. Cicchetti DV, Allison T. A new procedure for assessing reliability of scoring EEG sleep recordings. *Am J EEG Technol* 1971; 11: 101-9.
24. Magaziner J. The use of proxy respondents in health studies of the aged. In: Wallace RB, Woolson RF, eds. *The Epidemiologic Study of the Elderly.* New York: Oxford University Press, 1992; 120-9.
25. Zimmerman SI, Magaziner J. Methodological issues in measuring the functional status of cognitively impaired nursing home residents: the use of proxies and

performance-based measures. *Alzheimer Dis Assoc Disord* 1995; 8 (Suppl. 1): S281-90.

26. Magaziner J, Simonsick EM, Kashner TM, Hebel JR. Patient-proxy response comparability on measures of patient

health and functional status. *J Clin Epidemiol* 1988; 41: 1065-74.

Received in revised form 18 September 1996