

Care-Receiver and Caregiver Assessments of Functioning: Are There Gender Differences?*

Julie C. Shaver and Diane E. Allan
Centre on Aging, University of Victoria

RÉSUMÉ

Lorsque les prestataires de soins ne sont pas en mesure de répondre aux questions relatives à leurs capacités fonctionnelles, les fournisseurs de soins sont souvent amenés à se substituer à eux. Cependant, les études portant sur les ententes entre les fournisseurs et les prestataires de soins demeurent d'une portée relativement restreinte, mettant surtout l'accent sur les fonctions cognitives ainsi que la relation avec le fournisseur de soins. Cette étude vise à élargir la portée des recherches dans ce domaine par l'étude des comptes rendus effectués par les fournisseurs et les prestataires de soins à l'égard des diverses activités de la vie quotidienne (AVQ) et des activités instrumentales de la vie quotidienne (AIVQ), en fonction du sexe du prestataire de soins. En outre, la recherche étudie l'influence de l'âge du prestataire et de son état de santé mentale et physique sur les ententes. Les données pour cette étude proviennent d'un échantillon de 388 Manitobains âgés de 65 ans et plus. Les résultats (fondés sur le facteur kappa de Cohen) révèlent que les ententes sont moins fréquentes, en ce qui a trait à l'existence d'une incapacité, chez les hommes que chez les femmes, en particulier chez les personnes âgées de 75 à 84 ans. Les différences selon l'état de santé sont également indiquées. L'incidence des résultats sur l'évaluation du fonctionnement est abordée.

ABSTRACT

In situations where care-receivers cannot respond to questions about their functional status, caregiver proxies are often substituted. Yet studies addressing caregiver-care-receiver agreement remain limited in scope, focusing primarily on cognitive functioning and caregiver relationship. This study broadens the range of research in this area by examining caregiver and care-receiver reports of individual IADL and AADL items by gender of care-receiver. As well, the degree to which the care-receiver's age and mental and physical health status influence agreement are investigated. Data for this study come from a sample of 388 Manitoba older adults aged 65 and over. Results (using Cohen's kappa) suggest less agreement on the presence of disability for men than for women, particularly among those aged 75-84. Differences by health status were also revealed. The implications of the findings for assessments of functioning are considered.

* The production of this paper was partially supported by a Michael Smith Foundation for Health Research Unit Infrastructure grant – Centre on Aging, University of Victoria. The data reported here were collected as part of the Manitoba Study of Health and Aging. This project was funded by Manitoba Health. The authors would also like to thank Neena Chappell for her comments.

Manuscript received: / manuscrit reçu : 01/04/04

Manuscript accepted: / manuscrit accepté : 19/10/04

Mots clés: prestataires de soins; fournisseurs de soins; fonctions physiques; différences selon le sexe

Keywords: care-receivers; caregivers; physical functioning; gender differences; aging

Requests for offprints should be sent to: / Les demandes de tirés-à-part doivent être adressées à :

Julie C. Shaver or Diane E. Allan
Centre on Aging
Sedgewick Building, Room A104
University of Victoria
PO Box 1700, Station CSC
Victoria, BC V8W 2Y2
(jshaver@uvic.ca)

Introduction

Gerontological research often uses self-reports by older adults to assess their physical, functional, and mental health. However, when individuals are unable to provide health information for themselves, caregiver respondents may be interviewed as replacements. These responses are then treated as equivalent to those from the individuals themselves. This practice of substituting caregiver responses for those of others generates concern about response validity. Invalid assessments of health and functioning can have serious implications for older adults. For example, assessments of functional capacity provided by older adults and their caregivers can have an impact on the receipt of services, as functional impairment is often used as a criterion for home care services. As well, activities of daily living (ADL) measures are used for a variety of purposes in policy and research (e.g., for projecting future health care needs).

Despite the importance of obtaining accurate and reliable assessments of health and functioning among older adults, literature assessing the validity of caregiver ratings of another's health remains limited. As well, empirical research that compares older adults' self-reports of physical and functional health with the reports of caregivers show somewhat mixed results. In addition, the degree to which both physical (e.g., disability) and mental (e.g., cognitive impairment) health influences the agreement between care-receiver and caregiver reports warrants further investigation. There also remains a need to examine the gender of care-receivers to assess whether it is a determinant of concordance between care-receiver and caregiver reports.

The current study examined the level of agreement between elderly care-receivers and their caregivers on the senior's functional capacity. Several care-receiver characteristics that may affect care-receiver and caregiver agreement were also examined. For example, both the gender and age of care-receivers were explored to determine if they have an impact on agreement. As well, this study looked at the physical and mental health status of care-receivers in an attempt to determine whether such characteristics influence care-receiver and caregiver agreement.

Review of the Literature

Some research suggests that care-receiver and caregiver ratings of functional ability are not completely concordant. Caregiver proxies are, in general, more likely to rate older adults as having impairment and disability than are the individuals themselves (Dorman, Waddell, Slattery, Dennis, & Sandercock,

1997; Duncan et al., 2002; Ellis et al., 2003; Long, Sudha, & Mutran, 1998; Magaziner, 1992; Neumann, Araki, & Gutterman, 2000; Shaw, McColl, & Bond, 2001). For example, Rubenstein, Schairer, Wieland, & Kane (1984) looked at functional status assessments given by elderly individuals and significant others. They found that care-receivers rated themselves significantly higher (i.e., more independent) than did significant others, with most discrepancies found in ambulation, bathing, dressing, and grooming. Similarly, Elam et al. (1991) found significant differences in responses between care-receivers and caregivers for walking and telephone use. Uriel, Wood-Dauphinee, Korner-Bitensky, Gayton, & Hanley (1998) found moderate concordance for physical functioning but concluded that overall, proxies were not able to provide information that corresponded with subjects.

Yet, other studies report that while caregivers are more likely to rate older adults as requiring assistance with functional tasks, the level of agreement between care-receivers and caregivers remains relatively high. For example, Epstein, Hall, Tognetti, Son, & Conant (1989) found subject and proxy mean responses to be generally similar for functional status. Rothman, Hendrick, Bulcroft, Hickam, & Rubenstein (1991) found care-receiver-generated and caregiver-generated (mostly spouses) functional ability scores to be highly correlated, although caregivers were more likely to rate care-receivers as slightly more impaired than the care-receivers themselves. Magaziner, Bassett, Hebel, & Gruber-Baldini (1996) found excellent agreement for instrumental activities of daily living (IADL) but also noted that proxies tend to over-report dependency for these activities. Similarly, Ostbye, Tyas, McDowell, & Koval (1997) found that the overall level of agreement between care-receiver and caregiver reports of functional ability were reasonably high for older adults without dementia, although caregivers were more likely to rate older adults as requiring assistance with ADLs than were care-receivers themselves. Santos-Eggimann, Zobel, & Bérod (1999) found informal caregivers to be very acceptable proxies for older adults, with the two exceptions of incontinence and laundry, where caregivers reported dependence with these activities more often than did care-receivers. Shaw et al. (2001) reported high levels of agreement on functional ability questions, while Andresen, Vahle, & Lollar (2001) found fair to good agreement on ADL and IADL activities, with the exceptions of moving a wheelchair and managing money. A study by Yip, Wilber, Myrtle, & Grazman (2001) found physical functioning to have high agreement between subjects and proxies. Despite the fact that proxies

tended to rate subjects as more impaired than subjects rate themselves, Duncan et al. (2002) reported excellent agreement between proxies and subjects on ADL and IADL domains.

Other studies also suggest that care-receiver and caregiver ratings of functional status are comparable. For example, Long et al. (1998) found high levels of agreement between care-receiver and caregiver ratings of functional status, although there was some disagreement on IADL items. Likewise, Kiyak, Teri, & Borson (1994) found no significant differences in care-receiver and caregiver reports of functional ability for healthy non-demented older adults, while Epstein et al. (1989) found strong correlations between ratings by healthy older adult and caregiver (close family member or friend), for functional health.

One factor that may influence levels of agreement is gender. The Medical Research Council Cognitive Function and Ageing Study (2000a) found that gender of care-receivers did not influence agreement with caregiver responses for the health-oriented questions in their study. The Medical Research Council Cognitive Function and Ageing Study (2000b) found better agreement between female care-receivers and female caregivers. Gender is known to have an impact on various facets of health. For example, there is extensive research evidence indicating that there are gender differences in health status. In studies of self-rated health, males are more likely to report excellent health than females (Wilcox, Kasl, & Idler, 1996). Research also indicates that more females than males suffer mild to moderate kinds of disability at all ages (see Verbrugge, 1989). Further, health status can affect the level of agreement between care-receiver and caregiver ratings of functional ability. Indeed, Clipp and Elder (1987) found a divergence between care-receiver and caregiver ratings of functional health status. They discovered that this discrepancy is most apparent when physical decline is in its early stages, when changes in ability tend to be discounted by care-receivers but are recognized by caregivers.

Despite gender differences in self-rated health, no studies have specifically examined whether these types of differences have an impact on the level of agreement between care-receiver and caregiver pairs. Given this fact, there remains a need to explore these issues to determine the extent of their impact on levels of agreement between care-receivers and caregivers.

Another factor that may influence agreement is mental health (e.g., cognitive impairment). Studies that examine care-receiver and caregiver reports using samples comprising older adults with

some form of cognitive impairment generally find that care-receivers with dementia typically rate themselves as having higher functional ability than do caregivers. Furthermore, this discrepancy increases as level of cognitive impairment increases. For example, Kiyak et al. (1994) conducted a longitudinal study and found that individuals with Alzheimer's disease reported significantly higher levels of functional ability than did their family member caregivers. They found that ratings of functional status by healthy older adults (e.g., no cognitive impairment) and their families were more concordant. Interestingly, this study also found that care-receivers with mild to moderate Alzheimer's disease were aware of and able to report declines in cognitive and functional status, but they reported lower levels of decline than did family members. Along similar lines, Karagiozis, Gray, Sacco, Shapiro, & Kawas (1998) found that older adults with greater cognitive impairment (i.e., dementia) showed poor judgment of their functional status and typically rated their functional abilities as higher than did their caregivers. Further, this tendency increased with the severity of the dementia, although fewer rated their skills higher than caregivers for money management, shopping, and hobbies. This study also determined that spouses were better informants for care-receivers with dementia than "other" informants (e.g., children, siblings, friends, paid caregivers). Magaziner et al. (1996) found a tendency toward less agreement between respondents and proxies when the respondent had cognitive limitations. Another study by Ostbye et al. (1997) examined the level of agreement between care-receiver and caregiver reports of physical functioning and found that the highest levels of disagreement were for pairs where the care-receiver was suffering from dementia, with agreement decreasing as severity of dementia increased. Thus, several studies demonstrate that cognitive impairment can have an effect on care-receiver and caregiver agreement; when cognitive impairment increases, agreement decreases.

It is also possible that age influences care-receiver and caregiver levels of agreement. Increasing age is positively correlated with greater levels of physical and mental impairment (see Christensen et al., 1994). Thus, one would expect less care-receiver and caregiver concordance as care-receivers get older and are more apt to be experiencing physical and cognitive decline. On the other hand, it is also possible that as age increases, "older old" adults become more accepting of their conditions and are more willing to acknowledge their physical and mental limitations than "younger old" adults.

Methods

Data

Data were drawn from the 1991–1992 community sample component of the Manitoba Study of Health and Aging (MSHA). The MSHA was an extension of the Canadian Study of Health and Aging (CSHA), a national study conducted to examine various aspects of dementia (CSHA Working Group, 1994). A list of names of those persons over the age of 65 and living in Manitoba was obtained from the Manitoba Health Services Commission. Persons living in mental health hospitals, correctional institutions, other institutions, personal care homes, the northern Norman and Thompson health regions (except those in Flin Flon and The Pas), and members of the military and Royal Canadian Mounted Police were excluded. The sample was stratified by health region and age group. A total of 2,890 persons were contacted between February 1991 and November 1992. Overall, 1,763 persons participated in a personal interview, which took, on average, just over one hour to complete. The refusal rate for the interviews was 20 per cent.

A screening measure, the Modified, Mini-Mental State Exam (3MS) (Teng & Chiu, 1987), was administered to all 1,763 respondents. Those scoring below 78 on the 3MS test, along with those unable to complete the test, and a sample of those who scored 78 or above, went on to the clinical phase of the study. Based on the results of four additional tests, individuals were classified as exhibiting no cognitive loss, cognitive loss but no dementia, or some form of dementia. In total, 211 individuals (64%) received the clinical assessment.

Those diagnosed with cognitive impairment or dementia were eligible for the next phase of the study involving caregivers. The purpose of these interviews was to identify problems encountered by caregivers, and what resources they require and use to assist them with their tasks. In all cases the primary caregiver was identified and interviewed. Eighty-three per cent of those eligible completed the caregiver interview. As well, 256 individuals and their caregivers who had screened negative for cognitive loss or dementia were included in the caregiver sub-study. A total of 388 care-receiver and caregiver pairs were completed.

Care-Receiver Characteristics

Table 1 presents sample characteristics for care-receivers and caregivers. Females comprised 59 per cent of the sample. Just over half of the respondents reported greater than a primary school education.

Almost half (47%) of the respondents reported living alone, while 67 per cent reported living in detached or semi-detached housing. Further, 56 per cent of respondents reported living in an urban area.

When compared to others their age, half the respondents reported their health as good. Only seven per cent perceived their health as poor or bad. The mean number of chronic conditions reported by respondents was five (SD=3). Thirteen per cent of care-receivers were depressed as measured by the CES-D scale (Radloff, 1977).

Caregiver Characteristics

The average age of caregivers was 58 (SD=14), ranging from 26 to 91 years. About two-thirds (67%) of caregivers were female. Just over one-third (36%) were spouses, while almost one-third (31%) were daughters and 17 per cent were sons. The mean number of years of education for caregivers was 11 (SD=4).

Measurement

The dependent variable in this study was functional ability, as measured by activities of daily living (ADL). In order to examine the association between care-receiver and caregiver responses across individual ADL items, the kappa statistic (k) was used. All ADLs were defined as follows: (a) absence of in-person assistance=0, and (b) receipt of in-person assistance=1. For the purposes of discussion, ADLs are classified as instrumental ADLs (IADL) and advanced ADLs (AADLs).¹ This typology was originally developed by Wolinsky, Callahan, Fitzgerald, & Johnson (1993). IADLs include light housework, heavy housework, yard work, and shopping. AADLs include using the telephone, taking medication, preparing meals, daily spending, and managing long-term finances.

Independent variables include gender, age, and physical and mental health status. Gender was dichotomized into male (1) and female (2). Age was separated into three categories: 65–74 (1), 75–84 (2), and 85 and over (3). One feature of the MSHA is the categorization of subjects into cognitive and physical impairment groups. For the 388 care-receiver and caregiver pairs used in this paper, 54 subjects had dementia, 89 had cognitive loss but no dementia, 184 reported no cognitive loss and no disability, and 72 reported no cognitive loss and disability. Disability was defined as requiring in-person assistance with three or more ADLs.

Table 1: Care-receiver and caregiver characteristics

Care-Receiver	<i>n</i>	
Age		
65–74	133	34%
75–84	173	44%
85+	82	21%
Female	230	59%
Primary school education only	195	52%
Average monthly household income	310	\$1,463 (SD = \$1,070)
Living arrangement		
Live alone	178	47%
Live in detached/semi-detached house	252	67%
Live in urban area	216	56%
Perceived health		
Excellent	59	16%
Good	187	49%
Fair	106	28%
Poor/bad	28	7%
Chronic conditions	381	Mean = 5 (SD = 3)
Depression	49	13%
Caregiver		
Age	388	Mean = 58 years (SD = 14)
Female	260	67%
Education	385	Mean = 11 years (SD = 4)
Relationship to care-receiver		
Spouse	138	36%
Daughter	122	31%
Son	67	17%

Analysis Techniques

The kappa statistic (*k*) (Cohen, 1968) – a test that measures agreement between groups corrected for chance (i.e., tests for level of agreement that is greater than agreement expected by chance alone) – was used to determine the strength of agreement between care-receivers and caregivers. Kappa ranges from –1 to 1, with –1 indicating total disagreement, 0 indicating agreement due to chance, and 1 indicating total agreement. Typically, scores greater than .75 suggest excellent agreement, scores from .41 to .74 suggest fair to good agreement, and scores less than .41 suggest poor agreement (Fleiss, 1973). If there is very little variability in the raw scores (i.e., if there

is a high level of agreement), kappa cannot be calculated because there is not enough variability. In addition, since a low kappa rating can occur even though there may be 100 per cent agreement (*k* is sensitive to the number of observations made and the distribution of those observations around the diagonal), percentages were also reported. Care-receiver/caregiver ratings were determined to be in agreement if both persons rated the care-receiver as totally independent (absence of in-person assistance) or dependent with each task considered. Otherwise, ratings were determined to be discordant. Non-response was so low (<0.5%) for both care-receivers and caregivers that further analyses to compare

the rate of missing values among the groups were not conducted.

Results

Gender Differences

Table 2 reports percentages of care-receivers and caregivers who perceived that care-receivers were able to perform various ADLs without assistance. In general, male care-receivers were more likely than their caregivers to report they could manage ADLs without assistance. Tasks most likely to pose difficulty for care-receivers in this sample included yard work, heavy housework, shopping, and management of long-term finances. Overall, both male care-receivers and male caregivers indicated that males were more likely than females to manage five of nine ADLs without assistance. However, the five activities in which males reported greater frequency of independence than females differed between male care-receivers and their caregivers. Male care-receivers reported independence with shopping, heavy housework, yard work, managing long-term finances, light housework, and using the telephone. Caregiver ratings suggested male care-receivers were independent with shopping, heavy housework, yard work, managing long-term finances, taking medication, and daily spending. Female care-receivers were more likely than male care-receivers to report absence of assistance on only one activity – taking medication. However, caregivers reported

that female care-receivers were more likely than male care-receivers to report no need for assistance with light housework, using the telephone, and preparing meals.

The degree of agreement (k) between care-receivers and caregivers for individual ADL items is reported in Table 3. In general, there was fair to good agreement between care-receivers and caregivers, as evidenced by the k scores ranging from .40 to .75. The two activities that exhibited the most disagreement for both males and females were getting about the house and using the telephone. However, no discernable response patterns were evident across gender and ADL categories. Interestingly, no ADL activities were defined as having excellent agreement.

Differences by Gender and Age

ADL ratings were further analyzed by the age groups of the care-receivers. Table 3 reports percentages of care-receivers and caregivers who perceived that subjects were able to perform various ADLs without assistance. Across all three age categories, males were more likely than females to report not requiring assistance with ADLs. However, as age increased, there was a greater tendency for both genders to report a need for assistance. This age trend was particularly evident for IADLs and, to a lesser degree, AADLs.

Table 3 shows level of agreement (k) between care-receiver and caregiver responses by the gender and

Table 2: Percentage of care-receivers (CR) and caregivers (CG), and agreement (k), who report that care-receivers can manage ADL item independently, by gender of care-receiver

	Male ($n = 155-157$)			Female ($n = 219-225$)		
	CR	CG	k	CR	CG	k
IADL						
Light housework	96	87	.24	90	91	.45
Heavy housework	75	58	.46	57	46	.61
Yard work	72	58	.47	41	29	.39
Shopping	83	70	.41	65	58	.51
AADL						
Using telephone	98	89	.17	97	90	.23
Taking medication	92	88	.49	94	87	.53
Preparing own meals	87	77	.36	87	81	.59
Daily spending	87	83	.64	87	76	.43
Managing money	79	73	.60	68	55	.34

Table 3: Percentage of care receivers (CR) and caregivers (CG) who report care receivers can manage ADL item independently by gender and age of care receiver

	65-74 years						75-84 years						85+ years						
	Male (n = 58-60)			Female (n = 68-70)			Male (n = 71-73)			Female (n = 98-99)			Male (n = 22-23)			Female (n = 52-56)			
	CR	CG	k	CR	CG	k	CR	CG	k	CR	CG	k	CR	CG	k	CR	CG	k	
IADL																			
Light housework	95	92	.47	94	93	.41	97	82	.09	93	94	.26	91	87	.33	79	82	.55	
Heavy housework	83	72	.58	76	68	.64	78	55	.28	57	42	.64	48	35	.56	36	25	.33	
Yard work	78	77	.67	65	42	.40	76	53	.34	36	27	.32	48	30	.29	19	17	.29	
Going shopping	92	78	.37	84	72	.42	83	72	.37	65	60	.41	57	39	.32	39	38	.59	
AADL																			
Using telephone	100	93	*	97	93	.26	99	86	-.03	97	92	.33	91	87	.78	96	82	.11	
Taking medication	97	97	.48	96	91	.65	93	88	.37	99	89	.15	78	70	.55	82	80	.71	
Preparing own meals	90	83	.57	93	87	.53	86	79	.28	93	87	.45	83	57	.24	68	63	.65	
Daily spending	95	98	.49	97	87	.33	85	79	.63	91	80	.41	74	57	.63	66	55	.33	
Managing money	83	86	.48	78	62	.23	77	68	.66	73	56	.33	73	55	.62	44	42	.37	

*Kappa statistic unavailable, given that 100% of care-receivers report requiring no assistance in using the telephone

age group of the care-receiver. After taking chance into account, both male and female care-receivers reported fair to good agreement with caregivers on the majority of ADLs. The greatest level of agreement was found in the 65–74 age group. Males in the 75–84 age group reported the poorest levels of agreement; they exhibited agreement for only daily spending and managing long-term finances. In the age categories of 65–74 and 85 and over, the discrepancy between care-receiver and caregiver was greater for females than males. Furthermore, female care-receivers consistently exhibited poor agreement with caregivers across all age groups, while for males there was a similar discrepancy with shopping.

Differences by Gender and Physical and Mental Health Status

Given the four sub-groups described earlier, it was possible to examine care-receiver and caregiver differences in levels of agreement on ADLs by gender and physical and mental health status of the care-receiver. Table 4 presents the percentages of care-receivers and caregivers who perceived that care-receivers were able to perform various ADLs without assistance. Overall, male care-receivers across all health status groups were more likely than female care-receivers to report not requiring ADL assistance. Those with a disability were also the most likely to indicate a need for assistance with IADLs, while those with dementia were the most likely to report requiring assistance with AADLs. Given the nature of these tasks (i.e., greater physical demand for IADLs and greater cognitive demand for AADLs), these patterns are not surprising.

In the case of male care-receivers, the largest discrepancy with caregivers was evident among those in the no disability group. These individuals and their caregivers disagreed on six of seven ADLs. Managing long-term finances was the only ADL for which there was agreement. Unlike male care-receivers, females did not appear to differ in agreement or disagreement by physical and mental health status. However, there was a consistent pattern of poor agreement across health status groups for yard work, using the telephone, and managing long-term finances.

Table 4 also displays the level of agreement corrected for chance (k) for individual ADLs by gender and physical and mental health status of the care-receiver. As a result of the small cell sizes at this level of analysis, and the equal row and column prerequisite of k , this statistic could not be calculated for some activities. Overall, the agreement between care-receivers and caregivers was mixed, with greatest

agreement between care-receivers and caregivers found in the dementia, cognitive impairment, and no cognitive impairment/disability groups. Interestingly, the care-receivers who were classified as having no mental and physical disabilities reported the greatest disagreement with caregivers. The individual ADLs for which disagreement was consistently reported across health status group and gender include yard work, shopping, using the telephone, and managing long-term finances.

Discussion

The literature suggests that self-reports and caregiver reports of functional ability are, in some instances, not interchangeable. The mixed results reported in the literature suggest a need to further examine care-receiver and caregiver reports of functional ability. The purpose of this study was to determine the level of agreement between elderly care-receivers and caregiver proxy respondents on functional ability, and to assess whether (1) care-receiver's gender affects level of agreement between care-receivers and caregivers, and whether (2) agreement varies by age and health status sub-group. It should be noted that the *accuracy* of self-versus-proxy reports cannot be validated, as we did not have an objective verification of independent functional status. Thus, we did not know, for example, whether care-receivers overstated their ability or the caregivers under-reported care-receiver ability. Without this "gold standard," researchers need to be aware that combining proxy and subject responses in surveys may lead to biased results. Despite the fact that we were not able to know for certain whose perception (care-receiver or caregiver) represents true functional status (i.e., whose opinion depicts the "gold standard"), we *could* discuss the levels of agreement by physical and mental health status, gender, and age, and offer some explanations for any observed discrepancies. As well, it should be noted again that when there is a high level of agreement (i.e., very little variability in the raw scores) between care-receivers and caregivers, kappa cannot be calculated, because there is not enough variability.

This study sought to determine whether gender is a determinant of concordance for care-receiver and caregiver ratings of receipt or absence of assistance with ADLs. To satisfy this objective, this study looked at gender differences in the percentages of receipt or absence of assistance on individual ADL items and differences by gender in the degree of agreement between care-receiver and caregiver. Degree of agreement was calculated using kappa (k).

Table 4: Percentage of care-receivers (CR) and caregivers (CG), and agreement (k), who report care-receivers can manage ADL item independently by gender and cognitive impairment status of care-receiver

	Dementia						Cognitive impairment						Disability						No disability								
	Male (n=18-19)			Female (n=29-31)			Male (n=42-43)			Female (n=32-33)			Male (n=9)			Female (n=63)			Male (n=86)			Female (n=95-97)					
	CR	CG	k	CR	CG	k	CR	CG	k	CR	CG	k	CR	CG	k	CR	CG	k	CR	CG	k	CR	CG	k			
IADL																											
Light housework	89	74	.50	71	61	.64	93	86	.37	82	91	.37	87	81	-.17	78	89	-.17	87	81	.09	100	90	*	100	98	*
Heavy housework	47	42	.26	30	13	.53	63	49	.54	48	30	.51	32	22	.61	22	11	.61	32	22	.44	93	71	.24	86	76	.44
Yard work	42	37	.23	24	21	.31	62	50	.57	28	13	.35	8	13	-.13	11	11	-.13	8	13	.23	91	72	.27	73	48	.19
Going shopping	56	39	.68	45	13	.31	71	67	.33	31	50	.50	44	46	.40	67	33	.40	44	46	.52	95	81	.14	95	84	.02
AADL																											
Using telephone	89	68	.11	90	68	.19	98	81	.19	94	85	-.10	100	100	*	100	100	*	97	87	.37	100	96	*	100	100	*
Taking medication	58	53	.47	60	43	.42	96	86	.46	100	91	*	89	89	-.13	89	89	-.13	97	89	.42	99	98	-.02	100	99	*
Preparing own meals	74	58	.43	61	29	.41	79	69	.51	82	91	.37	67	56	.31	67	56	.31	81	77	.71	97	87	-.06	100	96	*
Daily spending	53	37	.69	68	13	.13	81	79	.63	73	79	.34	78	89	.61	78	89	.61	83	88	.46	99	95	-.02	100	97	*
Managing money	44	28	.65	33	10	.18	72	67	.56	50	50	.50	56	56	.10	56	56	.10	60	48	.37	92	87	.51	89	74	-.12

*Kappa statistic cannot be calculated because there is not enough variability in raw scores.

Gender

In overall percentages of agreement, care-receivers were more likely to report requiring no assistance with ADLs than were caregivers. This finding is consistent with research that indicates that care-receivers tend to report higher levels of functioning than caregivers (see Ellis et al., 2003; Magaziner, 1992; Neumann et al., 2000; Rubenstein et al., 1984). Interestingly, caregivers were more likely to report that males do not require assistance with items such as heavy housework, yard work, and managing long-term finances, while they reported that females did not require assistance with items such as light housework, telephone use, and meal preparation. For this cohort of respondents it is likely that females would be socialized to perform the more expressive tasks such as shopping, meal preparation, and housework, while males would typically perform the more instrumental tasks such as yard work and managing long-term finances. Thus, the caregivers (in many cases the spouse) may have based their responses on a socialized division of labour rather than on the care-receiver's ability to do the task. Hence, this factor could have contributed to the observed disagreement between care-receiver and caregiver responses.

Kappa (*k*), which distinguishes whether agreement between care-receivers and caregivers is better than that predicted by chance alone, displayed no consistent patterns. The kappa values revealed that, in general, there was fair to good agreement between care-receivers and caregivers for just over half the items. This finding suggests substantial disagreement between care-receivers and caregivers, contrary to other studies that found a similarity between care-receiver and caregiver reports (Duncan et al., 2002; Kiyak et al., 1994; Long et al., 1998; Ostbye et al., 1997; Rothman et al., 1991; Santos-Eggimann et al., 1999; and Shaw et al., 2001).

Gender and Age

Across all age categories there was a tendency for male care-receivers to be less likely than female care-receivers to report requiring assistance. In line with Christensen et al. (1994), this study also discovered an age trend in which older age corresponds with increased likelihood of requiring assistance, particularly for IADLs.

The kappa values revealed one pattern amid the gender and age data: women of all ages exhibited poor agreement for long-term finances, while men of all ages disagreed on shopping. Again, we are left to wonder whether this type of task segregation is a result of caregivers (mostly spouses) reporting less independence because the care-receiver cannot per-

form the activity, or if the subject simply does not perform the activity (i.e., never learned how, or the task is completed by the spouse). This confusion suggests that survey questions about functional ability need to be refined. For example, Magaziner (1997) points to the need to ask questions about what the care-receiver actually does, versus a proxy's impression of the care-receiver's ability. Thus, a more appropriate ADL question should first determine whether the care-receiver performs the task at all. Furthermore, there is a need to ask clear and objective questions about ADL abilities in general, such as asking whether the care-receiver can button his shirt as opposed to whether he can dress himself (Magaziner, 1992).

In general, fair to good agreement was found between care-receiver and caregiver reports; however, the 75–84 group exhibited much lower kappa values (i.e., less agreement) than the two other age groups. For example, this age group showed fair to good agreement on only 6 of 18 ADL items. It is plausible that persons in this age group represent a transitory group. That is, they may exhibit greater disagreement because they are at an age where they are just beginning to experience declines in their physical and mental capacity. Those individuals in the 65–74 age group may not have witnessed much decline in their physical and mental health status yet, while those over the age of 85 may have come to terms with their limitations. As well, there may be higher agreement for those in older than 85 because their disability is more observable. Thus, it is the middle age group (75–84 years) that find themselves in disagreement with caregivers, as they are likely going through a transition in their health status and may not recognize, want to admit, or be willing to accept, such declines.

Gender and Physical and Mental Health Status

Care-receivers were more likely than caregivers to report requiring no assistance with ADLs across all four health status groups. The "healthy" group agreed on only 2 of 11 ADL items. For example, male care-receivers exhibited good agreement for managing long-term finances, while females reported fair agreement for heavy housework. Further, the disability group agreed on 7 of 17 ADL items, while the cognitive impairment group agreed on 9 of 17 items. Surprisingly, this study found the majority of functional ability ratings (10 of 18 ADL items) by care-receivers with dementia and their caregivers to be concordant. This finding contradicts Ostbye et al. (1997), who found that ratings were

most different for care-receivers with dementia and their caregiver. This difference could be a result of how dementia was defined in each study. For example, the current study looked at those with any level of dementia, whereas in the Ostbye study dementia was partitioned into mild and moderate groups. As well, the Ostbye study utilized a (larger) national sample ($n=800$), and the current study is based on data from Manitoba only ($n=388$). In addition, the Ostbye study used three-level ordinal data responses and a weighted k statistic, whereas the current study used dichotomous data responses and an unweighted k statistic. As they did in the gender-by-age findings, females consistently disagreed with caregivers about managing long-term finances, and males across health groups disagreed with caregivers about shopping.

This study also found discrepancies on reporting independence in telephone use, which may highlight a need to frame instrument questions differently. For example, instead of simply asking about "telephone use" (which is a vague query), it may be better to ask more specific questions, such as whether they can dial a telephone number or hold a conversation once on the phone.

As for study limitations, our results may be applicable only to samples such as the one indicated in this study. Since our data came from just one Canadian province, it is important to note that findings may not be generalizable to all community-dwelling older adults. Also, since sample sizes were small for some of the (male) gender and health status sub-groups, our results need to be confirmed in a larger sample.

Conclusion

These data provided an opportunity to examine ADL agreement between a sample of older adults and their caregivers. There was evidence that care-receivers were less likely to report requiring assistance with ADL activities than were caregivers. Our findings suggest substantial disagreement between care-receivers and caregivers. Further, there were no solid gender patterns in agreement. Despite not finding any consistent gender patterns, this study provided interesting findings that contribute to the literature on care-receiver and caregiver assessments of functional ability. There was considerable disagreement with IADLs and AADLs. For example, the IADL and AADL discrepancies that appeared when ADLs were examined by gender and physical and mental health status were particularly interesting and should be further examined. In addition, future studies may want to incorporate performance-based measures to verify patients' functional abilities.

Ultimately, before caregiver reports are used in place of care-receiver reports, it is important to clarify where and why these discrepancies in agreement of ADL ability occur. The variable nature of our findings highlights the need for further research to help generate a clearer understanding of the limitations of proxy-generated data so that strategies to address proxy-related reliability issues can be addressed.

Note

- 1 We also had data for personal activities of daily living (PADLs), but because of the high level of agreement for these activities, the data do not lend themselves to the kappa statistic.

References

- Andresen, E.M., Vahle, V.J., & Lollar, D. (2001). Proxy reliability: Health-related quality of life (HRQoL) measures for people with disability. *Quality of Life Research, 10*(7), 609–619.
- Canadian Study of Health and Aging Working Group. (1994). Canadian Study of Health and Aging: Study methods and prevalence of dementia. *Canadian Medical Association Journal, 150*(6), 899–913.
- Christensen, H., Jorm, A.F., Henderson, A.S., MacKinnon, A.J., Korten, A.E., & Scott, L.R. (1994). The relationship between health and cognitive functioning in a sample of elderly people in the community. *Age and Ageing, 23*, 204–212.
- Clipp, E.C., & Elder, G.H. (1987). Elderly confidants in geriatric assessment. *Comprehensive Gerontology Section B: Behavioural, Social, and Applied Sciences, 1*(1), 35–40.
- Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin, 70*(4), 213–220.
- Dorman, P.J., Waddell, F., Slattery, J., Dennis, M., & Sandercock, P. (1997). Are proxy assessments of health status after stroke with the EuroQol questionnaire feasible, accurate, and unbiased? *Stroke, 28*, 1883–1887.
- Duncan, P.W., Min Lai, S., Tyler, D., Perera, S., Reker, D.M., & Studenski, S. (2002). Evaluation of proxy responses to the Stroke Impact Scale. *Stroke, 33*, 2593–2599.
- Elam, J.T., Graney, M.J., Beaver, T., ElDerwi, D., Applegate, W.B., & Miller, S.T. (1991). Comparison of subjective ratings of function with observed functional ability of frail older persons. *American Journal of Public Health, 81*(9), 1127–1130.
- Ellis, B.H., Bannister, W.M., Cox, J.K., Fowler, B.M., Shannon, E.D., Drachman, D., Adams, R.W., & Giordano, L.A. (2003). Utilization of the Propensity Score method: An exploratory comparison of the proxy-completed to self-completed responses in the Medicare Health Outcomes Survey. *Health and Quality*

- of *Life Outcomes*, 1(47). Retrieved March 23, 2005, from <http://www.hqlo.com/content/1/1/47>.
- Epstein, A.M., Hall, J.A., Tognetti, J., Son, L.H., & Conant, L. (1989). Using proxies to evaluate quality of life: Can they provide valid information about patients' health status and satisfaction with medical care? *Medical Care*, 27(3), S91–S98.
- Fleiss, J.L. (1973). *Statistical methods for rates and proportions*. New York: Wiley.
- Karagiozis, H., Gray, S., Sacco, J., Shapiro, M., & Kawas, C. (1998). The direct assessment of functional abilities (DAFA): A comparison to an indirect measure of instrumental activities of daily living. *Gerontologist*, 38(1), 113–121.
- Kiyak, H.A., Teri, L., & Borson, S. (1994). Physical and functional health assessment in normal aging and in Alzheimer's disease: Self-reports vs. family reports. *Gerontologist*, 34(3), 324–330.
- Long, K., Sudha, S., & Mutran, E.J. (1998). Elder-proxy agreement concerning the functional status and medical history of the older person: The impact of caregiver burden and depressive Symptomatology. *Journal of the American Geriatrics Society*, 46(9), 1103–1111.
- Magaziner, J. (1992). The use of proxy respondents in health studies of the aged. In R.B. Wallace & R.F. Woolson (Eds.), *The epidemiologic study of the elderly* (pp. 120–129). New York: Oxford University Press.
- Magaziner, J. (1997). Use of proxies to measure health and functional outcomes in effectiveness research in persons with Alzheimer disease and related disorders. *Alzheimer Disease and Associated Disorders*, 11(6), 168–174.
- Magaziner, J., Bassett, S., Hebel, J.R., & Gruber-Baldini, A. (1996). Use of proxies to measure health and functional status in epidemiologic studies of community-dwelling women aged 65 years and older. *American Journal of Epidemiology*, 143(3), 283–292.
- Medical Research Council Cognitive Function and Ageing Study. (2000a). Survey into health problems of elderly people: A comparison of self-report with proxy information. *International Journal of Epidemiology*, 29, 684–697.
- Medical Research Council Cognitive Function and Ageing Study (2000b). Survey into health problems of elderly people: Multivariate analysis of concordance between self-report and proxy information. *International Journal of Epidemiology*, 29, 698–703.
- Neumann, P.J., Araki, S.S., & Gutterman, E.M. (2000). The use of proxy respondents in studies of older adults: Lessons, challenges, and opportunities. *Journal of the American Geriatrics Society*, 48(12), 1646–1654.
- Ostbye, T., Tyas, S., McDowell, I., & Koval, J. (1997). Reported activities of daily living: Agreement between elderly subjects with and without dementia and their caregivers. *Age and Ageing*, 26, 99–106.
- Radloff, L.S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1, 385–401.
- Rothman, M.L., Hedrick, S.C., Bulcroft, K.A., Hickam, D.H., & Rubenstein, L.Z. (1991). The validity of proxy-generated scores as measures of patient health status. *Medical Care*, 29(2), 115–124.
- Rubenstein, L.Z., Schairer, C., Wieland, G.D., & Kane, R. (1984). Systematic biases in functional status assessment of elderly adults: Effects of different data sources. *Journal of Gerontology*, 39(6), 686–691.
- Santos-Eggimann, B., Zobel, F., & Bérod, A.C. (1999). Functional status of elderly home care users: Do subjects, informal and professional caregivers agree? *Journal of Clinical Epidemiology*, 52(3), 181–186.
- Shaw, C., McColl, E., & Bond, S. (2001). Functional abilities and continence: The use of proxy respondents in research involving older people. *Quality of Life Research*, 9(10), 1117–1126.
- Teng, E.I., & Chiu, H.C. (1987). The Modified Mini-Mental State (3MS) Examination. *Journal of Clinical Psychiatry* 48(8), 314–318.
- Uriel, P., Wood-Dauphinee, S., Korner-Bitensky, N., Gayton, D., & Hanley, J. (1998). Proxy use of the Canadian SF-36 in rating health status of the disabled elderly. *Journal of Clinical Epidemiology*, 51(11), 983–990.
- Verbrugge, L.M. (1989). Gender, aging, and health. In K.S. Markides (Ed.), *Aging and health: Perspectives on gender, race, ethnicity and class* (pp. 23–78). Newbury Park: Sage.
- Wilcox, V.L., Kasl, S.V., & Idler, E.L. (1996). Self-rated health and physical disability in elderly survivors of a major medical event. *Journals of Gerontology: Social Sciences*, 51B(2), S96–S104.
- Wolinsky, F.D., Callahan, C.M., Fitzgerald, J.F., & Johnson, R.J. (1993). Changes in functional status and the risks of subsequent nursing home placement and death. *Journals of Gerontology: Social Sciences*, 48(3), S94–S101.
- Yip, J.Y., Wilber, K.H., Myrtle, R.C., & Grazman, D.N. (2001). Comparison of older adult subjects and proxy responses on the SF-36 health-related quality of life instrument. *Aging and Mental Health*, 5(2), 136–142.