Published in final edited form as:
J Aging Health. 2012 April ; 24(3): 367–383. doi:10.1177/0898264311424208.

Comparing Reports From Hip-Fracture Patients and Their Proxies: Implications on Evaluating Sex Differences in Disability and Depressive Symptoms

Michelle Shardell, PhD[1], **Dawn E. Alley, PhD**[1], **Ram R. Miller, MDCM**[1], **Gregory E. Hicks, PhD**[2], and **Jay Magaziner, PhD**[1]

[1]University of Maryland School of Medicine, Baltimore, MD, USA

[2]University of Delaware, Newark, DE, USA

Abstract

Objective—This study compared sex differences in disability and depressive symptoms using reports from hip fracture patients and their proxies.

Method—Hip fracture patients (49 men, 183 women) aged ≥65 years and proxies were interviewed 1 year postfracture. Outcomes were Center for Epidemiological Studies Depression (CES-D) and number of dependencies in performing activities of daily living and instrumental activities of daily living.

Results—Mean ADL sex differences (men minus women) were 0.40 (p = .37) using proxy reports and 0.70 (p = .08) using patient self-reports. Mean CES-D sex differences were −3.60 (p = .02) using proxy reports and −1.26 (p = .38) using patient self-reports. Discrepancies between patients and proxies were smallest for proxies who have lived with the patient ≥ 1 year.

Discussion—Patients and proxies produced conflicting conclusions about sex differences. Results suggest that ideal proxies to recruit are those who have been cohabitating with the patients for an extended length of time.

Keywords

proxy respondents; disability; depressive symptoms; hip fracture; sex differences

Introduction

Disability and depressive symptoms are important outcomes to consider when assessing disease consequences and treatment effectiveness among older adults. However, both disability and depressive symptoms are subjective constructs that rely on respondent perceptions for assessment. When a patient is unable to self-report disability and depressive symptoms in the research setting, proxies (e.g., relatives or caregivers) are commonly recruited to provide assessments. Proxy reports can help to minimize selection bias and preserve studies' external validity because the sickest, most cognitively impaired patients would otherwise be excluded from analysis (Corder, Woodbury, & Manton, 1996; Long,

---

Corresponding Author: Michelle Shardell, PhD, Department of Epidemiology and Public Health, University of Maryland School of Medicine, 660 West Redwood Street, Baltimore, MD 21201, USA, mshardel@epi.umaryland.edu.

Declaration of Conflicting Interests: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Sudha, & Mutran, 1998; Snow, Cook, Lin, Morgan, & Magaziner, 2005). However, although older adults and their proxies have been shown to hold similar conceptualizations of health (Lawrence, 1995), issues of bias remain even when proxy assessments are collected because of proxy-patient dyadic disagreement. Specifically, proxies tend to overreport older adults' disability (Dorevitch et al., 1992; Epstein, Hall, Tognetti, Son, & Conant, 1989; Lum, Lin, & Kane, 2005; Magaziner, Bassett, Hebel, & Gruber-Baldini, 1996; Magaziner, Simonsick, Kashner, & Hebel, 1988; Rothman, Hedrick, Bulcroft, Hickam, & Rubenstein, 1991; Santos-Eggimann, Zobel, & Berod, 1999; Todorov & Kirchner, 2000; Weinberger et al., 1992) and depressive symptoms (Bassett, Magaziner, & Hebel, 1990; Mackenzie, Robiner, & Knopman, 1989; Magaziner, Zimmerman, Gruber-Baldini, Hebel, & Fox, 1997; Teri & Wagner, 1991). In general, large discrepancies have been observed when measuring private, typically unobservable characteristics such as depressive symptoms (Epstein et al., 1989; Hung, Pickard, Witt, & Lambert, 2007; Snow et al., 2005).

Evaluation of proxy assessments is particularly relevant for hip fracture patients. Hip fracture is an increasingly common condition among older adults in the United States (Agency for Health Care Policy and Research, 1999). However, many older hip-fracture patients are unable to self-report disability and depressive symptoms due to cognitive impairment; thus hip fracture studies are highly dependent on proxy reports (Magaziner et al., 1988, 1997).

This reliance on proxy reports may be particularly important in understanding postfracture sex differences. Although men who fracture a hip die earlier than women (Hawkes, Wehren, Orwig, Hebel, & Magaziner, 2006; Penrod et al., 2008; Wehren et al., 2003), men and women tend to experience similar postfracture physical functioning (Hawkes et al., 2006; Samuelsson et al., 2009). Accurately quantifying sex differences in postfracture disability and depressive symptoms is important in light of this apparent contradiction. Such differences may merit tailoring treatment decisions, such as type of pharmacological preventative and postfracture surgical treatment, based on sex (Hawkes et al., 2006; Dy, Lamont, Ton, & Lane, 2011; Endo, Aharonoff, Zuckerman, Egol, & Koval, 2005; Samuelsson et al., 2009). Indeed, sex differences have been observed at the time of fracture. Male hip-fracture patients are younger and sicker (Bacon, 1996; Hawkes et al., 2006) and more likely to be married or cohabitating (Endo et al., 2005) than female hip-fracture patients. As a result, spouses are more likely to serve as proxies for male than for female patients, whereas offspring are more likely to serve as proxies for female than for male patients. These differences in proxies may lead to biased results when using proxy reports to evaluate sex differences, because spouse proxies (Ostbye, Tyas, McDowell, & Koval, 1997) and proxies cohabitating with a patient (McLoughlin, Cooney, Holmes, & Levy, 1996) tend to provide reports that are more concordant with patients than other types of proxies.

The primary objective of this study was to assess postfracture sex differences in disability and depressive symptoms using both patient self-reports and proxy reports, and to determine whether the two raters' reports produce conflicting conclusions. The secondary objective was to compare rater conclusions according to subgroups of proxy respondents to identify ideal types of proxies to recruit in studies. We hypothesized that conclusions about sex differences would differ most between proxies and patients where the proxy was not living with the patient, such as most offspring or unrelated caregivers, because these proxies tend to spend less time observing the patient compared to cohabitating proxies (e.g., spouses).

Method

Study Sample

This analysis used data from a validation substudy conducted as part of a larger prospective cohort study to examine the sequelae of hip fracture during the year postfracture (Magaziner et al., 1997, 2000). Among 674 patients enrolled in the prospective study, 328 patients were eligible for the proxy substudy. Among eligible patients, 233 patients completed an interview 12 months postfracture and provided the name of a proxy who completed the proxy interview at 12 months postfracture. The final sample for analysis consisted of 232 patients (49 men, 183 women) and patient-identified proxies (patient–proxy dyads). One patient was excluded from analyses due to missing values for all outcome variables examined in this study. Patients were recruited from eight Baltimore, Maryland hospitals from January 1990 to June 1991. Patient eligibility criteria included being community-dwelling and at least 65 years of age at the time of fracture. Eligible patients were asked to identify the person most knowledgeable about their health and general abilities. Patients were interviewed in person at their residence, and proxies were interviewed via telephone. The study was approved by the Institutional Review Board at the University of Maryland School of Medicine. Informed consent was obtained from all participants. Additional details regarding the study sample can be found elsewhere (Magaziner et al., 1997, 2000).

Measures

Study outcomes included depressive symptoms as measured by the Center for Epidemiological Studies–Depression (CES-D) scale and disability as measured by activities of daily living (ADLs) and instrumental activities of daily living (IADLs). The CES-D is a 20-item scale consisting of questions about feelings and behaviors. For each question, patients and proxies were asked how often within the past week the patient felt a certain way (*rarely, sometimes, occasionally, most of the time*). Each item was worth 0 to 3 points, for a total of 0 to 60 points, where higher scores indicate more severe depressive symptomatology (Radloff, 1977).

For each ADL, patients and proxies were asked whether in the past week the patient performed the task independently, performed the task with human and/or equipment assistance, did not perform the task for health reasons, or did not perform the task for nonhealth reasons. Each task was dichotomized as dependent (requiring assistance or unable to perform the task for health reasons) versus independent. Patients who reported not performing the task for non-health reasons were classified as not dependent, because this reason does not indicate disability. The ADL summary scale was calculated as the number of ADL dependencies (range: 0-10), where higher scores indicate more ADL dependence. The scale consisted of 10 tasks (walking 10 ft, transferring to/from bed, putting on socks and shoes on both feet, transferring to/from toilet, eating, putting on shirt, buttoning shirt, grooming, bathing, putting on pants) that tap into mobility and domains considered by Katz (Katz, Ford, Moskowitz, Jackson, & Jaffe, 1963): eating, toileting, bathing, dressing, and transferring.

Patients and proxies were asked about the patients' need for assistance in performing each IADL task during the last 2 weeks. As with ADLs, each task was dichotomized as dependent versus independent. The IADL summary scale was calculated as the number of IADL dependencies (range: 0-7), thus higher scores indicate more IADL dependence. The scale consisted of seven tasks adapted from the Older Americans Resources and Services instrument (Fillenbaum, 1988): using the telephone, handling money, taking medications, traveling to places out of walking distance, shopping, preparing meals, and doing housework. Proxies completed assessments of patient CES-D, ADLs, and IADLs from the proxy viewpoint, the “proxy-proxy” perspective (Pickard & Knight, 2005). That is, proxies reported based on their own assessment of the patient rather than what they thought the patient self-reported.

Proxy characteristics considered in subgroup analyses were self-rated health (poor or fair vs. good or excellent), relationship with patient (spouse, other relative, not related), and living arrangement with patient (not living together, living < 1 year together, living ≥1 year together). These characteristics were chosen because previous studies found relationships between these factors and proxy–patient discrepancy or agreement (Magaziner et al., 1988, 1996; Ostbye et al., 1997). Proxies' self-rated health was assessed because it may relate to psychological distress and caregiver burden, both of which are associated with proxy–patient discrepancies (Long et al., 1998; Rothman et al., 1991).

Data on patient age and cognition and proxy age and sex were also collected. Cognition was measured using the Mini-Mental State Examination (MMSE), a scale ranging from 0 to 30, where higher scores indicate better cognition (Folstein, Folstein, & McHugh, 1975).

Statistical Analysis

Characteristics of proxies and patients described above were compared between male and female patients using independent two-sample t tests or Wilcoxon rank-sum tests for continuous variables. Chi-square tests were used to compare categorical characteristics. Means of ADLs, IADLs, and CES-D were calculated separately by patient sex and rater. Rater-specific mean sex differences and 95% confidence intervals were calculated and assessed using two-sample t tests. Next, mean sex differences for ADLs, IADLs, and CES-D and 95% confidence intervals were calculated separately by rater within subgroups of proxy characteristics (proxy self-rated health, relationship to patient, and length of time living with patient). Statistical significance was defined as $p < .05$ or 95% confidence intervals that exclude 0.

Results

Table 1 provides patient and proxy characteristics for the study sample. Female patients were significantly older (mean age 81.2, standard deviation [*SD*] = 7.0 years) than male patients (mean age 78.9, *SD* = 7.6 years). Compared to proxies for female patients, proxies for male patients were significantly older (mean age 64.1, *SD* = 15.8 years vs. mean age 59.4, *SD* = 13.4 years), and significantly more likely to be female (85.7% vs. 74.9%), to report poor or fair health (22.4% vs. 13.1%), to be the patient's spouse (46.9% vs. 8.2%), and to have lived at least 1 year with the patient (73.5% vs. 45.0%), all $p < .05$. In addition, the minimum MMSE score for female and male patients was 10 and 19, respectively; the maximum score for both groups was 30.

Table 2 shows mean number of ADL and IADL dependencies, and mean CES-D score by patient sex and rater. The mean ADL sex difference (males minus females) was 0.40 (95% confidence interval [CI] [−0.46, 1.25]; $p = .37$) when using proxy reports, and 0.70 (95% CI [−0.07, 1.47]; $p = .08$) when using patient self-reports. The mean IADL sex differences were similar between proxy reports (mean difference = −0.12; 95% CI [−0.79, 0.56]; $p = .73$) and patient reports (mean difference = −0.13; 95% CI [−0.74, 0.48]; $p = .68$). Last, the mean CES-D sex difference was −3.60 (95% CI [−6.89, −0.31]; $p = .02$) when using proxy reports, and −1.26 (95% CI [−4.12, 1.61]; $p = .38$) when using patient self-reports.

Table 3 shows ADL, IADL, and CES-D sex differences by rater and proxy subgroups. Rater discrepancies of mean ADL sex differences were largest among proxies living with the patient for less than 1 year, where the sex difference (male patients minus female patients) was −1.96 (95% CI [−3.79, −0.13]) using patient self-reports and 0.32 (95% CI [−2.03, 2.67]) using proxy reports. In contrast, for proxies living with the patient for at least 1 year, the sex difference was 1.02 (95% CI [0.01, 2.03]) using patient self-reports and 0.55 (95% CI [−0.57, 1.66]) using proxy reports. Rater discrepancies of mean ADL sex differences were also large among proxies who rated their health as poor or fair although both raters reported greater ADL disability among males. The sex difference was 2.43 (95% CI [0.77, 4.08]) using patient self-reports and 0.92 (95% CI [−1.09, 2.94]) using proxy reports.

Rater discrepancies of mean IADL sex differences shown in Table 3 were small and the sex differences for all subgroups and both raters were all nonsignificant. The mean CES-D sex differences for all subgroups and both raters were also all nonsignificant; however, some subgroups had large rater discrepancies. Rater discrepancies of mean CES-D sex differences in Table 3 were largest among proxies not living with the patient, where the sex difference (male patients minus female patients) was 5.82 (95% CI [−1.27, 12.91]) using patient self-reports and −5.42 (95% CI [−12.74, 1.90]) using proxy reports. In contrast, for proxies living with the patient for at least 1 year, the sex difference was −1.67 (95% CI [−5.07, 1.73]) using patient self-reports and −1.79 (95% CI [−5.84, 2.26]) using proxy reports. Large rater discrepancies in CES-D sex differences were also found for all three proxy–patient relationship subgroups (spouse, other relative, unrelated), where patients and proxies reported mean sex differences in opposite directions within all three subgroups.

Discussion

Results from this study indicated that conclusions about sex differences in ADL disability and depressive symptoms differed by rater. We found ADL disability to be marginally significantly greater in men compared to women using patient data, but no statistically significant sex difference in ADLs using proxy data. This discrepancy was due to proxies overreporting ADL disability in females relative to patient self-reports.

Depressive symptoms, as measured by the CES-D, were significantly greater in women compared to men using proxy data, but not significantly different using patient data. This result occurred because, in contrast to previous studies (Bassett et al., 1990; Mackenzie et al., 1989, 1997; Teri & Wagner, 1991), proxies in this study underreported depressive symptoms in both sexes, and at a greater magnitude in male patients than in female patients. Proxies overreported IADL disability at equal magnitudes in both sexes, resulting in no discrepancy between patient and proxy reports regarding IADL sex differences.

We found that proxy–patient living arrangement affected the magnitude and direction of discrepancies between raters when assessing sex differences. Most notably, rater discrepancies for ADLs and CES-D sex differences were smallest for proxy–patient dyads living together for at least 1 year compared to other living arrangements. This result suggests that the ideal proxy, when available, is one who has been cohabitating with the patient for at least 1 year. One potential reason for the large CES-D discrepancy is that male patients may be reluctant to express feelings of despair to noncohabitants, leading to large discrepancies between male patients and their noncohabitating proxies; but these feelings may be expressed to or perceived by cohabitants via nonverbal cues (Kessler, Brown, & Broman, 1981; Mirowsky & Ross, 1995).

Rater discrepancies in sex differences may differ by outcome examined. ADL and IADL scales are designed to assess disability, which is not a characteristic inherent in individuals. Rather, disability is the interaction of individuals with their environment (Verbrugge & Jette, 1994). Thus rater discrepancies may be due to differences in how the proxies and patients interact with their environment (e.g., different environments, different ability expectations, etc.; Albrecht & Devlieger, 1999). These differences may be gender-role related, where, for example, male patients who do not regularly do housework may not report being unable to do housework. Also, given that the patients have been in the process of recovering from an acute injury for 12 months, patients may compare their current abilities to those during an earlier phase of recovery rather than to prefracture abilities, thus leading to higher levels of self-efficacy than the proxy's perceived efficacy of the patient (Seeman, Unger, McAvay, & Mendes de Leon, 1999). This last explanation is an example of response-shift, in which adaptation to conditions (e.g., disability following hip fracture) leads to different perceptions of conditions, and therefore changes in responses (Daltroy, Larson, Eaton, Phillips, & Liang, 1999; Gibbons, 1999). Shifts in patient response may be due to changes in standards, definitions, or values. For example, in the absence of objective confirmatory information of improved independence, a patient who previously reported requiring assistance in doing housework may now report requiring no assistance. This change in response may reflect compensations or environmental modifications being employed for doing housework, or housework ceasing to be a priority. Such changes in perspective may not be observable to a proxy and may differ by sex (e.g., men and women may perceive changes in their ability to do housework differently), thus leading to dyadic disagreement that differs by sex. Indeed, previous work in community-dwelling older adults has found that women have worse perceived declines in function than do men, even after adjusting for functional performance (Daltroy et al., 1999; Gibbons, 1999).

An additional origin for proxy–patient discrepancies that is particularly relevant for hip fracture patients is caregiver burden. Perceived caregiver burden and psychological distress of proxies have been associated with over-reporting of disability (Long et al., 1998; Rothman et al., 1991). Spouse proxies (Ostbye et al., 1997) and proxies living with the patient (McLoughlin et al., 1996) may perceive less burden than offspring proxies because, unlike offspring, spouses and cohabitants generally need not commute between two households or juggle caregiving activities with other duties (e.g., work, raising a family). In this sample, proxies for men were more likely to be a spouse or a cohabitant of the patient than proxies for women. These differences in proxy characteristics correspond to previous reports of sex differences in care giver characteristics, with men receiving more hours of home care than women (Katz, Kabeto, & Langa, 2000).

Last, discrepancies between patients and proxies may be due to the perspective from which proxies provided reports. Pickard and Knight (Pickard & Knight, 2005) described two proxy perspectives: the “proxy-proxy” perspective, where proxies report their own assessment of the patients, and the “proxy-patient” perspective, where proxies report how they think the patients would respond. Although this study used the proxy-proxy perspective, which is expected to be less consistent with patient reports than the proxy–patient perspective, a recent randomized trial of cancer patients found that perspective did not significantly affect dyadic disagreement (Gundy & Aaronson, 2008). However, a similar study on perspective has not been performed among hip fracture patients, and it is unknown whether perspective affects proxy reporting differently between male and female patients.

This study has multiple strengths. First, the inclusion of male and female hip-fracture patients and patient-identified proxies allowed direct proxy–patient comparisons. Second, we not only examined differences in patient–proxy reporting differences but we also examined how these differences affect conclusions about an important clinical issue: sex differences in disability and depression after fracture. Third, we were able to examine reporting differences across subgroups, allowing us to identify optimal proxy characteristics for recruitment in future studies. Despite these strengths, the following limitations should be kept in mind. First, the study sample was small, leading to difficulty in detecting proxy–patient dyadic disagreement in mean sex differences, especially within proxy subgroups.

More important, proxy–patient dyadic disagreement can only be assessed among patients willing and able to respond to interview questions. Comparisons cannot be made with patients who require a proxy, that is, those who are too sick or too cognitively impaired to respond to interview questions. Also, we treat patient data as the gold standard here as is often done in proxy–patient validation studies (i.e., patient perceptions of disability and depressive symptoms are the targets of evaluation); however, it is possible that patients may misreport disability or depressive symptoms, for example, due to cognitive impairment (Snow et al., 2005). Last, the study only included patients from the greater Baltimore area, resulting in limited generalizability.

Consistent with previous research, we found that proxy data should not be used interchangeably with study participant data (Kane et al., 2005). The innovation of this study was showing that discrepancies between patient and proxy reports may lead to qualitatively different scientific conclusions about postfracture sex differences, depending on the rater. These results are clinically significant because they demonstrate that substituting proxy data for missing patient data can lead to inaccurate findings. A consequence of this problem is that proxy–patient dyadic disagreement can impede identifying important targets of intervention to improve recovery from hip fracture among older men and women.

The current practice of substituting proxy data for missing patient data is equivalent to making the implausible assumption of perfect correlation between patient and proxy responses (Snow et al., 2005). An open methodologic question is how to account for differential proxy–patient disagreement, in particular, how to use proxy characteristic data to correct for discrepant proxy responses in statistical analyses when patient perceptions are considered the gold standard. When proxy responses are collected, one solution is to recruit a small proxy validation sample for patients who also provide self-reports. The validation sample can be used to perform a sensitivity analysis by reanalyzing data, where proxy responses are adjusted by the average amount of proxy–patient disagreement estimated from the validation sample (Magaziner, 1992). Such an approach would improve the transparency and reproducibility of research reports by providing readers a characterization of potential bias from proxies. Also, a recent report suggested performing a sensitivity analysis by adapting statistical methods that were originally designed to handle missing data to handle discrepancies from proxy data (Shardell, Hicks, Miller, Langenberg, & Magaziner, 2010).

In conclusion, this study showed that findings about sex differences in patient disability and depressive symptoms differed by rater. When proxy reports are needed in research studies, proxies living with study patients for at least 1 year may provide more accurate reports than other proxies for both sexes. If such a proxy is unavailable, then sensitivity analyses accounting for proxy characteristics, ideally using a validation sample, are recommended. Additional research identifying other targets of intervention prone to differential proxy–patient dyadic disagreement is warranted.

Acknowledgments

Funding: The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Institutes of Health grant numbers K12 HD043489, K12 HD055931, K23 AG027746, K25 AG034216, R37 AG009901, R01 AG09902.

References

Agency for Health Care Policy and Research. Hospital Inpatient Statistics, 1996. Washington, DC: US Department of Health and Human Services; 1999.

Albrecht GL, Devlieger PJ. The disability paradox: High quality of life against all odds. Social Science and Medicine. 1999; 48:977–988. [PubMed: 10390038]

Bacon WE. Secular trends in hip fracture occurrence and survival: Age and sex differences. Journal of Aging and Health. 1996; 8:538–553. [PubMed: 10182385]

Bassett SS, Magaziner J, Hebel JR. Reliability of proxy response on mental health indices for aged, community-dwelling women. Psychology and Aging. 1990; 5:127–132. [PubMed: 2317291]

Corder LS, Woodbury MA, Manton KG. Proxy response patterns among the aged: Effects on estimates of health status and medical care utilization from the 1982-1984 long-term care surveys. Journal of Clinical Epidemiology. 1996; 49:173–182. [PubMed: 8606318]

Daltroy LH, Larson MG, Eaton HM, Phillips CB, Liang MH. Discrepancies between self-reported and observed physical function in the elderly: The influence of response shift and other factors. Social Science and Medicine. 1999; 48:1549–1561. [PubMed: 10400256]

Dorevitch MI, Cossar RM, Bailey FJ, Bisset T, Lewis SJ, Wise LA. The accuracy of self and informant ratings of physical functional capacity in the elderly. Journal of Clinical Epidemiology. 1992; 45:791–798. [PubMed: 1619459]

Dy CJ, Lamont LE, Ton QV, Lane JM. Sex and gender considerations in male patients With osteoporosis. Clinical Orthopaedics and Related Research. 2011; 469:1906–1912. [PubMed: 21400003]

Endo Y, Aharonoff GB, Zuckerman JD, Egol KA, Koval KJ. Gender differences in patients with hip fracture: A greater risk of morbidity and mortality in men. Journal of Orthopaedic Trauma. 2005; 19:29–35. [PubMed: 15668581]

Epstein AM, Hall JA, Tognetti J, Son LH, Conant L Jr. Using proxies to evaluate quality of life. Can they provide valid information about patients' health status and satisfaction with medical care? Medical Care. 1989; 27:S91–98. [PubMed: 2921889]

Fillenbaum, GG. Multidimensional functional assessment of older adults: The Duke older Americans resources and Services procedures. Hillsdale, NJ: Lawrence Erlbaum; 1988.

Folstein MF, Folstein SE, McHugh PR. “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research. 1975; 12:189–198. [PubMed: 1202204]

Gibbons FX. Social comparison as a mediator of response shift. Social Science and Medicine. 1999; 48:1517–1530. [PubMed: 10400254]

Gundy CM, Aaronson NK. The influence of proxy perspective on patient-proxy agreement in the evaluation of health-related quality of life: An empirical study. Medical Care. 2008; 46:209–216. [PubMed: 18219250]

Hawkes WG, Wehren L, Orwig D, Hebel JR, Magaziner J. Gender differences in functioning after hip fracture. Journals of Gerontology Series A: Biological Sciences and Medical Science. 2006; 61:495–499.

Hung SY, Pickard AS, Witt WP, Lambert BL. Pain and depression in caregivers affected their perception of pain in stroke patients. Journal of Clinical Epidemiology. 2007; 60:963–970. [PubMed: 17689813]

Kane RL, Kane RA, Bershadsky B, Degenholtz H, Kling K, Totten A. Proxy sources for information on nursing home residents' quality of life. Journals of Gerontology Series B: Psychological Sciences and Social Sciences. 2005; 60:S318–S325.

Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged. The index of ADL: A standardized measure of biological and psychosocial function. Journal of the American Medical Association. 1963; 185:914–919. [PubMed: 14044222]

Katz SJ, Kabeto M, Langa KM. Gender disparities in the receipt of home care for elderly people with disability in the United States. Journal of the American Medical Association. 2000; 284:3022–3027. [PubMed: 11122589]

Kessler RC, Brown RL, Broman CL. Sex differences in psychiatric help-seeking: Evidence from four large-scale surveys. Journal of Health and Social Behavior. 1981; 22:49–64. [PubMed: 7240706]

Lawrence RH. The structure of physical health status. Comparing proxies and self-respondents. Journal of Aging and Health. 1995; 7:74–98. [PubMed: 10165962]

Long K, Sudha S, Mutran EJ. 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. 1998; 46:1103–1111. [PubMed: 9736103]

Lum TY, Lin WC, Kane RL. Use of proxy respondents and accuracy of minimum data set assessments of activities of daily living. Journals of Gerontology Series A: Biological Sciences and Medical Science. 2005; 60:654–659.

Mackenzie TB, Robiner WN, Knopman DS. Differences between patient and family assessments of depression in Alzheimer's disease. American Journal of Psychiatry. 1989; 146:1174–1178. [PubMed: 2764175]

Magaziner, J. The use of proxy respondents in health studies of the aged. In: Wallace, RB.; Woolson, RF., editors. The epidemiologic study of the elderly. New York, NY: Oxford University Press; 1992. p. 120-129.

Magaziner J, Bassett SS, Hebel JR, Gruber-Baldini A. 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. 1996; 143:283–292. [PubMed: 8561163]

Magaziner J, Hawkes W, Hebel JR, Zimmerman SI, Fox KM, Dolan M. Recovery from hip fracture in eight areas of function. Journals of Gerontology Series A: Biological Sciences and Medical Science. 2000; 55:M498–507.

Magaziner J, Simonsick EM, Kashner TM, Hebel JR. Patient-proxy response comparability on measures of patient health and functional status. Journal of Clinical Epidemiology. 1988; 41:1065–1074. [PubMed: 3204417]

Magaziner J, Zimmerman SI, Gruber-Baldini AL, Hebel JR, Fox KM. Proxy reporting in five areas of functional status. Comparison with self-reports and observations of performance. American Journal of Epidemiology. 1997; 146:418–428. [PubMed: 9290502]

McLoughlin DM, Cooney C, Holmes C, Levy R. Carer informants for dementia sufferers: Carer awareness of cognitive impairment in an elderly community-resident sample. Age and Ageing. 1996; 25:367–371. [PubMed: 8921141]

Mirowsky J, Ross CE. Sex-differences in distress—real or artifact. American Sociological Review. 1995; 60:449–468.

Ostbye T, Tyas S, McDowell I, Koval J. Reported activities of daily living: Agreement between elderly subjects with and without dementia and their caregivers. Age and Ageing. 1997; 26:99–106. [PubMed: 9177666]

Penrod JD, Litke A, Hawkes WG, Magaziner J, Doucette JT, Koval KJ. The association of race, gender, and comorbidity with mortality and function after hip fracture. Journals of Gerontology Series A: Biological Sciences and Medical Science. 2008; 63:867–872.

Pickard AS, Knight SJ. Proxy evaluation of health-related quality of life: A conceptual framework for understanding multiple proxy perspectives. Medical Care. 2005; 43:493–499. [PubMed: 15838415]

Radloff LS. The CES-D Scale: A self-report depression scale for research in the general population. Applied Psychological Measurement. 1977; 1:385–401.

Rothman ML, Hedrick SC, Bulcroft KA, Hickam DH, Rubenstein LZ. The validity of proxy-generated scores as measures of patient health status. Medical Care. 1991; 29:115–124. [PubMed: 1994145]

Samuelsson B, Hedstrom MI, Ponzer S, Soderqvist A, Samnegard E, Thorngren KG. Gender differences and cognitive aspects on functional outcome after hip fracture--a 2 years' follow-up of 2,134 patients. Age and Ageing. 2009; 38:686–692. [PubMed: 19767316]

Santos-Eggimann B, Zobel F, Berod AC. Functional status of elderly home care users: Do subjects, informal and professional caregivers agree? Journal of Clinical Epidemiology. 1999; 52:181–186. [PubMed: 10210234]

Seeman TE, Unger JB, McAvay G, Mendes de Leon CF. Self-efficacy beliefs and perceived declines in functional ability: MacArthur studies of successful aging. Journals of Gerontology Series B: Psychological Sciences and Social Sciences. 1999; 54:P214–222.

Shardell M, Hicks GE, Miller RR, Langenberg P, Magaziner J. Pattern-mixture models for analyzing normal outcome data with proxy respondents. Statistics in Medicine. 2010; 29:1522–1538. [PubMed: 20535763]

Snow AL, Cook KF, Lin PS, Morgan RO, Magaziner J. Proxies and other external raters: Methodological considerations. Health Services Research. 2005; 40:1676–1693. [PubMed: 16179002]

Teri L, Wagner AW. Assessment of depression in patients with Alzheimer's disease: Concordance among informants. Psychology and Aging. 1991; 6:280–285. [PubMed: 1863397]

Todorov A, Kirchner C. Bias in proxies' reports of disability: Data from the National Health Interview Survey on disability. American Journal of Public Health. 2000; 90:1248–1253. [PubMed: 10937005]

Verbrugge LM, Jette AM. The disablement process. Social Science and Medicine. 1994; 38:1–14. [PubMed: 8146699]

Wehren LE, Hawkes WG, Orwig DL, Hebel JR, Zimmerman SI, Magaziner J. Gender differences in mortality after hip fracture: The role of infection. Journal of Bone and Mineral Research. 2003; 18:2231–2237. [PubMed: 14672359]

Weinberger M, Samsa GP, Schmader K, Greenberg SM, Carr DB, Wildman DS. Comparing proxy and patients' perceptions of patients' functional status: Results from an outpatient geriatric clinic. Journal of the American Geriatrics Society. 1992; 40:585–588. [PubMed: 1587975]

Table 1

Study Sample Characteristics

	Male patients *N* = 49	Female patients *N* = 183	*p* value[a]
Patient characteristics			
Age, years mean (*SD*)	78.9 (7.6)	81.2 (7.0)	.04
Mini-Mental State Examination Score (range: 0-30), Mean (*SD*)	26.0 (3.1)	25.4 (4.7)	.33
Proxy characteristics			
Age, years number (%)			
65+	26 (53.1)	60 (32.8)	.03
50-65	14 (28.6)	74 (40.4)	
<50	9 (18.4)	49 (26.8)	
Age, years mean (*SD*)	64.1 (15.8)	59.4 (13.4)	.04
Sex number (%)			
Male	7 (14.3)	46 (25.1)	.11
Female	42 (85.7)	137 (74.9)	
Self-rated health number (%)			
Poor/fair	11 (22.4)	24 (13.1)	.10
Good/excellent	38 (77.6)	159 (86.9)	
Relationship with patient number (%)			
Spouse	23 (46.9)	15 (8.2)	<.001
Other relative	20 (40.8)	119 (65.0)	
Not related	6 (12.2)	49 (26.8)	
Years living with patient number (%)			
0	8 (16.3)	75 (41.2)	.001
< 1	5 (10.2)	25 (13.7)	
≥ 1	36 (73.5)	82 (45.0)	

[a] *p* values from two-sample *t* tests or Wilcoxon rank-sum tests (continuous variables) or chi-square tests (categorical variables).

Table 2

Relationships of Patient Sex and Rater With ADLs, IADLs, and CES-D

		Mean outcome by patient sex and rater		Sex difference (male minus female)[a]		
Outcome	**Rater**	**Male patients**	**Female patients**	**Mean**	**95% CI[b]**	***p* value**
ADLs[c]	Proxy	3.22	2.83	0.40	[−0.46, 1.25]	.37
	Patient	3.26	2.56	0.70	[−0.07, 1.47]	.08
IADLs[d]	Proxy	3.25	3.37	−0.12	[−0.79, 0.56]	.73
	Patient	2.75	2.88	−0.13	[−0.74, 0.48]	.68
CES-D[e]	Proxy	8.23	11.83	−3.60	[−6.89, 0.31]	.02
	Patient	11.46	12.72	−1.26	[−4.12, 1.61]	.38

[a] Mean among male patients minus mean among female patients.

[b] CI = confidence interval.

[c] ADLs = activities of daily living (range: 0-10).

[d] IADLs = instrumental activities of daily living (range: 0-7).

[e] CES-D = Center for Epidemiologic Studies Depression scale (range: 0-60).

Table 3

Sex Differences in ADLs, IADLs, and CES-D by Rater and Proxy Subgroups

		Proxy reports		Patient self-reports	
Outcome	**Proxy subgroup**	**Mean sex difference (male minus female)**[a]	**95% CI**[b]	**Mean sex difference (Male minus Female)**[a]	**95% CI**[b]
ADLs[c]	Self-rated health				
	Poor/fair	0.92	[−1.09, 2.94]	2.43	[0.77, 4.08]
	Good/excellent	0.20	[−0.77, 1.16]	0.25	[−0.62, 1.12]
	Relationship with patient				
	Spouse	1.88	[0.16, 3.61]	2.72	[1.25, 4.19]
	Other relative	0.19	[−1.11, 1.49]	−0.09	[−1.29, 1.11]
	Not related	0.48	[−1.89, 2.84]	1.00	[−0.94, 2.94]
	Years living with patient				
	0	−0.30	[−2.32, 1.72]	0.93	[−0.89, 2.75]
	<1	0.32	[−2.03, 2.67]	−1.96	[−3.79, −0.13]
	1	0.55	[−0.57, 1.66]	1.02	[0.01, 2.03]
IADLs[d]	Self-rated health				
	Poor/Fair	−0.10	[−1.70, 1.50]	0.18	[−1.39, 1.76]
	Good/excellent	−0.19	[−0.95, 0.57]	−0.26	[−0.93, 0.42]
	Relationship with patient				
	Spouse	1.26	[−0.17, 2.68]	0.83	[−0.49, 2.14]
	Other relative	−0.36	[−1.38, 0.66]	−0.46	[−1.36, 0.43]
	Not related	−0.46	[−2.30, 1.38]	0.23	[−1.56, 2.02]
	Years living with patient				
	0	0.10	[−1.45, 1.65]	−0.01	[−1.42, 1.41]
	<1	−0.97	[−3.01, 1.07]	−0.38	[−1.94, 1.18]
	≥1	−0.13	[−1.01, 0.74]	−0.23	[−1.04, 0.57]
CES-D[e]	Self-rated health				
	Poor/fair	−5.23	[−14.51, 4.05]	−4.68	[−12.16, 2.81]
	Good/excellent	−3.36	[−6.92, 0.20]	−0.54	[−3.69, 2.61]
	Relationship with patient				

		Proxy reports		Patient self-reports	
Outcome	**Proxy subgroup**	**Mean sex difference (male minus female)[a]**	**95% CI[b]**	**Mean sex difference (Male minus Female)[a]**	**95% CI[b]**
	Spouse	0.96	[−4.37, 6.29]	−1.10	[−5.71, 3.52]
	Other relative	−2.78	[−8.02, 2.45]	0.47	[−4.22, 5.15]
	Not related	−3.13	[−12.05, 5.79]	3.72	[−3.41, 10.85]
	Years living with patient				
	0	−5.42	[−12.74, 1.90]	5.82	[−1.27, 12.91]
	<1	−1.08	[−12.32, 10.16]	−3.52	[−11.04, 4.00]
	≥1	−1.79	[−5.84, 2.26]	−1.67	[−5.07, 1.73]

[a] Mean among male patients minus mean among female patients.

[b] CI = confidence interval.

[c] ADLs = activities of daily living (range: 0-10).

[d] IADLs = instrumental activities of daily living (range: 0-7).

[e] CES-D = Center for Epidemiologic Studies Depression scale (range: 0-60).