Journal of Gerontology: MEDICAL SCIENCES
2000, Vol. 55A, No. 11, M698–M702

Determination of Prefracture Physical Function in Community-Dwelling People Who Fracture Their Hip

Thomas James Overend,[1] Bert Mackenzie Chesworth,[2] Marina Sandrin,[2] Shelley Stroud,[3] Robert John Petrella,[4] and Richard McCalden[3]

[1]School of Physical Therapy and [4]Faculty of Medicine, University of Western Ontario, London, ON, Canada.
[2]Physiotherapy Department and [3]Division of Orthopedics, London Health Sciences Centre, ON, Canada.

Background. Prefracture physical function must be accurately determined to set appropriate and attainable goals for rehabilitation following hip fracture. This is especially important for people who were living independently prior to their fracture. This study determines reliability and internal consistency of a prefracture physical function questionnaire (PFPFQ) completed by both patients and knowledgeable informants (KIs).

Methods. A 20-item PFPFQ, including ambulation, transfers, balance, and self-care domains, was developed using focus groups. Community-dwelling patients with a hip fracture ($N = 40$, 77.9 ± 8 years) completed the PFPFQ on two occasions during postoperative acute care. Forty KIs were identified by the patients and also completed the PFPFQ on two occasions via telephone interview. Day-to-day reliability of the patients and KIs [intraclass correlation coefficients (ICC)], and internal consistency [Kuder-Richardson coefficient (KR)] of the PFPFQ were determined.

Results. Intrarater reliability was high with ICCs (95% confidence interval) of 0.94 (0.89, 0.96) for patients and 0.96 (0.93, 0.98) for KIs. Interrater reliability on occasion 1 had an ICC of 0.81 (0.69, 0.88). Internal consistency of the patient responses on the first occasion was high (KR coefficient = 0.896).

Conclusions. The PFPFQ is a reliable and internally consistent instrument for determining prefracture physical function in community-dwelling people who fracture their hip. In situations where patients with a hip fracture are unable to provide this necessary information, KIs can provide reliable estimates of prefracture function to assist in setting appropriate rehabilitation goals.

FRACTURE of the hip is a significant medical problem in older people because this injury is associated with significant levels of morbidity and disability. As few as one third of patients over the age of 65 with a hip fracture injury ever return to their previous level of mobility (1–3).

Physical function is an integral component of achieving and maintaining independence in activities of daily living (ADLs) and is a major contributor to the overall health status of older adults. One out of six community dwellers who fracture a hip will be discharged directly to institutional or long-term care (4). Recovery of prefracture physical function is perhaps most important for those older adults who lived in the community at the time of their fracture and desire to return there following rehabilitation. Returning these people to community-living environments not only meets their goals, but also postpones the high human and economic costs associated with institutionalization.

We suggest that determination of prefracture physical function in persons who fracture their hip is essential for setting appropriate rehabilitation goals. Currently, there is a lack of validated, reliable, clinically friendly tools available to assess premorbid physical function in the hip fracture population. Obtaining such information is frequently complicated by the inability of postsurgical patients to provide accurate and reliable responses, either because of preexisting or operatively induced confusion, or by the denial of their true health status (5,6).

Secondary respondents, also known as patient proxies, patient surrogates, significant others, or knowledgeable informants (KIs), are defined as those people (relatives, friends, neighbors, or health care providers) capable of providing accurate information about a patient. Such respondents have been used in other patient populations (palliative, chronic care, arthritic, Alzheimer's) as a secondary source of information about a patient's functional status (7–10). Patient and proxy ratings of presurgical health and functional status in elderly patients with hip fracture have been previously compared (11). The proxies were given a shortened version of the questionnaire that the patients completed. Agreement was higher for physical function compared with more subjective areas such as general ratings of health. Overall, proxies tended to underrate patients' performance.

The purpose of our study was to assess (i) the day-to-day reliability and internal consistency of a short, clinically friendly, prefracture physical function questionnaire (PFPFQ) in older patients receiving surgery for a hip fracture, and (ii) quantify the relationship between patients' and KIs' ratings of prefracture physical function.

Methods

Sample

We recruited 40 (29 women) community-dwelling older adults ($\overline{X} \pm SD$, 77.9 ± 8.0 years; range, 64–93 years) ad-

mitted to the Victoria Campus of the London Health Sciences Centre between July 1996 and February 1997 for surgical repair of a hip fracture. Sample size was determined based on detecting a reliability coefficient of >0.60 using the method of Donner and Eliasziw (12). This represents *substantial* reliability as defined by Landis and Koch (13). Patients were included if they were older than 60 years of age, had no comorbidities likely to interfere with rehabilitation, and could identify a person (KI) who they believed was most knowledgeable about their prefracture physical function. Patients were excluded if they were participating in any other studies, were admitted from an institutional setting, had clinically significant comorbidities likely to interfere with their rehabilitation, or had sufficient cognitive impairment precluding provision of informed consent, self-report of prefracture physical function, and identification of a KI. Subjects were also excluded if they had knowledge of discharge plans to an institutional setting at the time of admission. For purposes of this study, institutional settings were defined as nursing homes or other such long-term care facilities. Retirement communities, in which people live independently with no on-site medical services, were not defined as institutional settings. The study was approved by an ethical review board, and all subjects provided written informed consent before participation.

Questionnaire

We developed a 20-item PFPFQ (Appendix) covering activities in four functional domains (ambulation, transfers, balance, and self-care). The individual questionnaire items were derived from concepts of physical function that have been used in the literature to quantify aspects of physical function in older adults. Specifically, we reviewed items from the Activities-Specific Balance Confidence (ABC) Scale (14), the Falls Efficacy Scale (15), the Life-Space Diary (16), and the Functional Independence Measure (17). Response categories were dichotomous (yes/no answers) to facilitate the use of the questionnaire in a clinical environment. The number of positive responses was summed to provide a total score out of 20. Following initial development of the 20 items, the PFPFQ was evaluated qualitatively for content validity using three focus groups: (i) patients after hip fracture rehabilitation, (ii) their KIs, and (iii) members of the hip fracture care team. It has been recommended that focus groups be a routine part of the development of a survey research instrument (18).

Procedure

The PFPFQ was administered to the patients on 2 different days between postoperative day 2 and day 5. The PFPFQ was also administered to KIs via telephone interviews on two separate occasions within 1 week. A standard script was developed for the telephone interviewers to explain the nature of the study and to request permission to administer the KI-specific PFPFQ. Both versions of the PFPFQ were worded to assess *capacity* to perform activities ("Could you . . .?" or "Could Mr./Mrs. . . .?") rather than *performance* ("Do you . . .?" or "Does Mr./Mrs. . . .?"). Capacity can be assessed via questionnaire, but performance is best assessed by direct observation.

Data Analysis

Intrarater reliability for patients and KIs and the relationship between patient and KI ratings of prefracture physical function were determined with the intraclass correlation coefficient [(ICC) 2, 1] (19). The ICCs were interpreted on the basis of subjective categories described by Fleiss: ICCs of 0.00 to 0.40 were considered "poor"; 0.40 to 0.75, "fair to good"; and greater than 0.75, "excellent" (20).

Correlation analysis was used to examine the extent of systematic bias between test occasions and between raters. Correlations between the variables of interest were determined using a Pearson product moment correlation coefficient (r). Internal consistency was evaluated for patient responses on occasion 1, using the analog to Cronbach's alpha for dichotomous responses—the Kuder-Richardson (KR) coefficient (21). Item-by-item analyses were also carried out. To control for the effects of rater bias and overall prevalence of "yes" and "no" responses (22), we used prevalence-adjusted and bias-adjusted kappa (PABAK) scores (23). All statistical analyses were performed with a computer-based statistical package (SigmaStat, version 2.0, Jandel Scientific, San Rafael, CA). Statistical significance was set at $p < .05$ for all tests.

RESULTS

Demographic information on the patients is provided in Table 1. Ninety-three percent lived in their own home or apartment at the time of their hip fracture, and 7% lived in a retirement community. A greater proportion of women (59%) lived alone than men (18%) ($\chi^2 = 5.23$, $p < .022$). Sixty-three percent used no ambulatory aids, and the balance of participants used canes (10%), walkers (12%), or both canes and walkers (15%).

The most common relationship of KI to the patient was daughter ($n = 15$), followed by wife ($n = 8$). Other KIs included friends ($n = 5$), sons ($n = 5$), husbands ($n = 2$), and brother, daughter-in-law, niece, partner, and granddaughter ($n = 1$, each).

Mean scores on the PFPFQ for the patients were 17.7 (occasion 1) and 18.0 (occasion 2). For the KIs, corresponding values were 17.4 (occasion 1) and 17.6 (occasion 2). These scores indicate a high level of prefracture function and reflect the fact that older adults who lived in the community at the time of their hip fracture were recruited for the study.

High correlations were observed between the PFPFQ scores on occasion 1 and occasion 2 for both patients ($r = .941$, $p < .001$) and KIs ($r = .959$, $p < .001$), and the correlation between patients and KIs on occasion 1 was .823, $p < .001$. Data plots comparing KI evaluations on the two test-

Table 1. Age and Living Status for the Subjects (Mean, *SD*)

Group	Age (y)
Women ($n = 29$)	78.2 (9.1)
Men ($n = 11$)	76.8 (4.1)
Total ($N = 40$)	77.9 (8.0)
Live alone?	
Yes ($n = 19$)	80.1 (9.0)
No ($n = 21$)	75.8 (6.6)

ing occasions revealed clustering of questionnaire scores at the two extremes of function, with greater dispersion at lower levels of estimated function, suggesting that KIs were less reliable when judging function in persons with lower functional ability. No systematic bias was noted in these data. Intrarater reliability was excellent, with ICCs [95% confidence interval (CI)] of 0.94 (0.89, 0.96) for patients and 0.96 (0.93, 0.98) for their KIs. Internal consistency of the patient responses on the first day was high (KR coefficient = 0.896). The interrater relationship between KIs and patients on the first testing occasion was also excellent, with an ICC (95% CI) of 0.81 (0.69, 0.88).

Item-by-item comparisons between the patients' responses on occasion 1 and occasion 2 indicated high PABAK values for all items [range, 0.70 (item 5) to 1.0]. Similar PABAK values were obtained for the item-by-item comparisons between the KI responses on both occasions [range, 0.60 (item 7) to 1.0]. For the comparison of items between patient and KI on occasion 1, PABAK values were similar [range, 0.59 (item 7) to 1.0], with the exception of item 14 (0.33). This question dealt with the ability to dress the lower body while standing.

Discussion

There are two main results of this study. The first is that the PFPFQ was shown to be a reliable and internally consistent instrument for determining prefracture physical function in community-dwelling people who fracture their hip. The second is that KIs were shown to provide reliable and concurrently valid estimates of pre-fracture function when compared with patient estimates. The results of this study will be of value to health care personnel involved in the rehabilitation of people who fracture their hip.

The accurate assessment of prefracture physical function is an important prerequisite for effective, goal-oriented rehabilitation. The "gold standard" for such assessment is direct observation, as utilized by Marottoli and colleagues in a large, prospective study of community-dwelling people (24). Over a 6-year period, 120 members in a cohort of 2,806 people sustained a hip fracture. Direct observation of the cohort was performed every 3 years, with phone interviews in the intervening years. This design provided good estimates of prefracture function. However, even a carefully designed prospective study such as this cannot control declines in subject function between the last functional assessment and the hip fracture (there was a mean interval of 6 months between most recent assessment and fracture) (24).

It is neither practical nor economically feasible to assess a huge sample of elderly people on a regular basis in order to have accurate knowledge of physical function prior to the hip fractures that will inevitably occur to some of the sample. Therefore, prefracture function must be determined postfracture by an indirect method, either from the patient or from a secondary respondent. Only two previous studies have investigated indirect assessment of prefracture function in people with a hip fracture. Magaziner and colleagues reported that patient proxies overestimated patient disability relative to the patients themselves (11). This effect was particularly true for instrumental ADLs and physical health measures. Proxy responses were more accurate if proxies were older than 65 years and if proxies lived with the patients. Magaziner and colleagues noted that poorer agreement was obtained with behaviors that were less concrete, less observable, or more private (11).

A second study from the same research group compared proxy ratings of functional status with subject self-reports in five domains (physical, instrumental, affective, cognitive, social) (25). Direct observation of performance was also carried out in subsets of the physical and instrumental domains. An ICC of 0.65 was reported for agreement between proxy and self-reports in the physical function domain. Once again, greater agreement between proxy and self-reports was noted for functions that were more observable and less private. Overall, proxies tended again to report more disability than subjects, with the highest levels of bias present in instrumental and physical function domains (25).

We chose to use dichotomous responses in the PFPFQ to enhance its clinical utility. Both previous studies reported that dichotomous responses yielded higher agreements than questions asking about quantity of participation (11,25). Kaufert and colleagues noted that the advantage of dichotomous measures of performance is that they can easily be summarized as numerical scores (7). The disadvantage is that they mask a more complex decision process involved in the final rating of an individual as able or unable to perform a given activity. In that regard, higher concordance has been reported for dichotomous versus trichotomous scales in basic mobility and self-help functions (7).

Previous studies of agreement between patient and proxy reports of physical function and health status have consistently reported the following: (i) level of agreement is dependent upon concreteness, visibility, and subjectivity of the functions assessed; and (ii) proxies tend to underestimate functional capacity (26–30). Previous studies have also consistently suggested that direct observation is more accurate than self-report, which is, in turn, more accurate than proxy report (5,26,31,32).

Several reasons have been proposed for the existence of a reporting bias in the responses of both patients and proxies (26,33). Patients may overrate their health status to avoid being identified as a burden to caregivers. Conversely, they may underrate their health status to delay discharge planning. Patients may have a shifted time frame and respond based on their function before their illness or injury. Cognitive deficits may also influence accuracy of responses. Proxy respondents, including health professionals and significant others, may underestimate the health status of patients to justify their roles as caregivers or to gain sympathy. Significant others may underrate a family member's function in an attempt to encourage physicians to recommend institutional care. Proxies may also underrate the patient's capabilities simply by assuming the role of "evaluator."

The results of our study suggest that KIs do not underestimate patient function, at least as measured on the PFPFQ and in this sample of community-dwelling people with a hip fracture. This may be due to the following: (i) the dichotomous nature of the items on the PFPFQ, (ii) the generally concrete and observable nature of the items, (iii) the generally high physical function of our patients, or (iv) the fact

that we excluded patients with cognitive impairments from the study. It has been reported that poor cognitive status of the patient and poorer physical status both tend to predict low agreement (11), and the KIs in this study were less reliable when judging patients of lower function. Thus, it is important to recognize that our results cannot be generalized to the broader population of people who fracture their hip. It is also important to further test the utility of the PFPFQ in a wider sample of people who fracture their hip, including people with lower function in institutions, and people at any functional level who have cognitive deficits. It is in populations like these, where people may not be able to give accurate information on their prefracture physical function, that the PFPFQ, when completed by a KI, would have particular value. Further, the ceiling effect noted in this study of community-dwelling people who fracture their hip would likely be less evident in other hip fracture populations with lower functional ability at the time of injury. However, the tendency of KIs to be less reliable when judging patients of lower function in this study may affect reliability and concurrent validity of the PFPFQ to a greater extent in hip fracture populations with less functional ability.

This study has shown that the PFPFQ is a reliable instrument with good internal consistency for determining prefracture physical function in community-dwelling people who fracture their hip. Importantly, the PFPFQ is also user friendly for clinicians, as both direct and telephone administration take only between 5 to 10 minutes to complete. Furthermore, we have shown that KIs can be used to provide reliable and concurrently valid estimates of prefracture physical function for community-dwelling patients with a hip fracture. By providing reliable and consistent estimates of prefracture physical function, the PFPFQ may be helpful to those involved in the acute care and rehabilitation of community-dwelling older people with hip fracture.

Acknowledgments

This project was funded by the London Health Sciences Centre, Health Services Research Fund. The authors would like to thank Karolyn Smith, Guido Wisoztki, and Brenda Coleman for their assistance in data collection with the knowledgeable informants. Preliminary reports of the data were presented in abstract form at the American Physical Therapy Association and Canadian Physiotherapy Association annual meetings in 1998.

Address correspondence to Tom Overend, PhD, BScPT, School of Physical Therapy, University of Western Ontario, London, ON, Canada N6G 1H1. E-mail: toverend@julian.uwo.ca

References

1. Jette AM, Cleary PD, Harris BA, Campion EW. Functional recovery after hip fracture. *Arch Phys Med Rehabil.* 1987;68:735–740.
2. Mossey JM, Mutran E, Knott K, Craik R. Determinants of recovery 12 months after hip fracture. The importance of psychosocial factors. *Am J Public Health.* 1989;79:279–286.
3. Magaziner J, Simonsick EM, Kashner TM, Hebel JR, Kenzora JE. Predictors of functional recovery one year following hospital discharge for hip fracture. A prospective study. *J Gerontol Med Sci.* 1990;45:M101–M107.
4. Jaglal SB. Osteoporotic fractures: Incidence and impact. In: Williams JI, Badley EM, eds. *Patterns of Health Care in Ontario: Arthritis and Related Conditions.* Toronto, Canada: Institute for Evaluative Sciences; 1998.
5. Kuriansky JB, Gurland BJ, Fleiss JL. The assessment of self-care capacity in geriatric psychiatric patients by objective and subjective methods. *J Clin Psych.* 1976;32:95–102.
6. Carp FM, Carp A. It may not be the answer, it may be the question. *Res Aging.* 1981;3:85–100.
7. Kaufert JM, Green S, Dunt DR, Corkhill R, Creese AL, Locker BDS. Assessing functional status among elderly patients: a comparison of questionnaire and service provider ratings. *Med Care.* 1979;27:807–817.
8. Muhlenkamp AF, Joyner JA. Arthritis patients' self reported affective states and their caregivers' perceptions. *Nurs Res.* 1986;35:24–27.
9. Epstein AM, Hall JA, Tognetti J, Son LH, Conant L. Using proxies to evaluate quality of life. *Med Care.* 1989;27:591–598.
10. Tepper S, Beatty P, DeJong G. Outcomes in traumatic brain injury: self-report versus report of significant others. *Brain Inj.* 1996;10:575–581.
11. Magaziner J, Simonsick E, Kashner T, Hebel J. Patient-proxy response comparability on measures of patient health and functional status. *J Clin Epidemiol.* 1988;41:1065–1074.
12. Donner A, Eliasziw M. Sample size requirements for reliability studies. *Statistics Med.* 1987;6:441–448.
13. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–174.
14. Powell LE, Myers AM. The activities-specific balance confidence (ABC) scale. *J Gerontol Med Sci.* 1995;50A:M28–M34.
15. Tinetti ME, Richman D, Powell L. Falls efficacy as a measure of fear of falling. *J Gerontol Psych Sci.* 1990;45:P239–P243.
16. May D, Nayak USL, Isaacs B. The life-space diary: a measure of mobility in old people at home. *Int Rehabil Med.* 1985;7:182–186.
17. Hamilton BB, Laughlin JA, Granger CV, Kayton RM. Interrater agreement of the seven level Functional Independence Measure (FIM). *Arch Phys Med Rehabil.* 1991;72:572–580.
18. Fowler FJ. *Survey Research Methods.* 2nd ed. Newbury Park, CA: Sage Publications; 1993.
19. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull.* 1979;86:420–428.
20. Fleiss JL. *The Design and Analysis of Clinical Experiments.* Toronto, Canada: John Wiley; 1986.
21. Bruning JL, Kintz BL. *Computational Handbook of Statistics.* 3rd ed. Glenview, IL: Scott, Foresman; 1987.
22. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol.* 1990;43:543–548.
23. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol.* 1993;46:423–429.
24. Marottoli RA, Berkman LF, Cooney LM. Decline in physical function following hip fracture. *J Am Geriatr Soc.* 1992;40:861–866.
25. 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. *Am J Epidemiol.* 1997;146:418–428.
26. Rubenstein LZ, Schairer C, Wieland GD, Kane R. Systematic biases in functional status assessment of elderly adults: effects of different data sources. *J Gerontol.* 1984;39:686–691.
27. Rothman ML, Hendrick SC, Bulcroft KA, Hickam DH, Rubenstein LZ. The validity of proxy-generated scores as measures of patient health status. *Med Care.* 1991;29:115–124.
28. Spranger M, Aaronson N. The role of health care providers and significant others in evaluating the quality of life in patients with chronic disease: a review. *J Clin Epidemiol.* 1992;45:743–760.
29. Østbye T, Tyas S, McDowell I, Koval K. Reported activities of daily living: agreement between elderly subjects with and without dementia and their caregivers. *Age Ageing.* 1997;26:99–106.
30. Santos-Eggiman B, Zobel F, Clerc Bérod A. Functional status of elderly home care users: do subjects, informal and professional caregivers agree? *J Clin Epidemiol.* 1999;52:181–186.
31. Elam JT, Graney MJ, Beaver T, El Derwi D, Applegate WB, Miller ST. Comparison of subjective ratings of function with observed functional ability of frail older persons. *Am J Public Health.* 1991;81:1127–1130.
32. Dorevitch MI, Cossar RM, Bailey FJ, et al. The accuracy of self and informant ratings of physical functional capacity in the elderly. *J Clin Epidemiol.* 1992;45:791–798.
33. Pierre U, Wood-Dauphinee S, Korner-Bitensky N, Gayton D, Hanley J. Proxy use of the Canadian SF-36 in rating health status of the disabled elderly. *J Clin Epidemiol.* 1998;51:983–990.

Received October 8, 1999
Accepted January 10, 2000
Decision Editor: William B. Ershler, MD

Appendix

Prefracture Physical Function Questionnaire for Patients With a Hip Fracture

Date: _____ PIN: _____

Name: _____

Age: _____Gender: _____

Live Alone? Y / N—Residence: _____

Gait Aid: Y / N Cane: Y / N Quad Cane: Y / N Walker: Y / N

How many people came into your home to assist you in the week before your fall? _____

List the services they provided:_____

_____

_____

Before you broke your hip ..... ***(Circle Yes or No)***

Walking Domain

1. Could you walk around inside your home independently? ..... Y / N
2. Could you walk a distance of 20 feet outside your home? ..... Y / N
3. Could you walk outside your home to go across the street independently? ..... Y / N
4. Could you walk in crowded places, such as a mall, where people walk rapidly by you? ..... Y / N
5. Could you walk up a flight of stairs without help? ..... Y / N
6. Could you walk down a flight of stairs without help? ..... Y / N

Transfer Domain

7. Could you get into a chair without help? ..... Y / N
8. Could you get out of a chair without help? ..... Y / N
9. Could you get into a bathtub without help? ..... Y / N
10. Could you get out of a bathtub without help? ..... Y / N
11. Could you get on the toilet without help? ..... Y / N
12. Could you get off the toilet without help? ..... Y / N

Balance Domain

13. Could you dress yourself above the waist in standing? ..... Y / N
14. Could you dress yourself below the waist in standing? ..... Y / N
15. Could you pick up a slipper from in front of the closet? ..... Y / N
16. Could you stand on your tip-toes to reach for something above your head? ..... Y / N

Self -Care Domain

17. Could you prepare a meal for yourself? ..... Y / N
18. Could you brush your teeth without help? ..... Y / N
19. Could you comb/brush your hair without help? ..... Y / N
20. Could you wash your hands without help? ..... Y / N

TOTAL OF 'YES' ..... _____