

Operationalizing a Frailty Index from a Standardized Comprehensive Geriatric Assessment

David M. Jones, MSc, MD,* Xiaowei Song, PhD,[†] and Kenneth Rockwood, MD[†]

OBJECTIVES: To construct and validate a frailty index (FI) that is clinically sensible and practical for geriatricians by basing it on a routinely used comprehensive geriatric assessment (CGA) instrument.

DESIGN: Secondary analysis of a 3-month randomized, controlled trial of a specialized mobile geriatric assessment team.

SETTING: Rural Nova Scotia. Participants were seen in their homes.

PARTICIPANTS: Frail older adults, of whom 92 were in the intervention group and 77 in the control group.

MEASUREMENTS: A standard CGA form that accounts for impairment, disability, and comorbidity burden was scored and summed as a frailty index (FI-CGA). The FI-CGA was stratified to describe three levels of frailty. Patients were followed for up to 12 months to determine how well the index predicted adverse outcomes (institutionalization or mortality, whichever came first).

RESULTS: The three levels of frailty were mild (FI-CGA 0–7), moderate (FI-CGA 7–13), and severe (FI-CGA >13). Demographic and social traits were similar across groups, but greater frailty was associated with worse function ($r = 0.55$) and mental status ($r = 0.33$). Those with moderate and severe frailty had a greater risk of adverse outcomes than those with mild frailty (unadjusted hazard ratio = 1.9 and 5.5, respectively). There was no difference between frailty groups in mean 3-month goal-attainment scaling scores. Intrarater reliability was 0.95.

CONCLUSION: The FI-CGA is a valid, reliable, and sensible clinical measure of frailty that permits risk stratification of future adverse outcomes. *J Am Geriatr Soc* 52:1929–1933, 2004.

Key words: frailty; controlled trial; comprehensive geriatric assessment

The construct of frailty appears to have a firm biological basis,¹ and frailty is easily recognized clinically. Still, operationalization of frailty remains a challenge.^{1–5} Frailty has been measured by constructing a frailty index (FI), based on the accumulation of a variety of functional deficits (such as comorbid illness, poor health attitudes, signs of disease, and self-reported disabilities).^{1,6,7} The FI, replicated in several studies, showed a high correlation with mortality^{1,6–8} but left open the question of whether or how it might be calculated from routinely collected data. In consequence, although the idea that frailty might be precisely measurable seems intriguing, how to achieve such precision through clinical assessment remains to be elucidated.

Useful clinical measures demonstrate validity, reliability,⁹ and sensibility.¹⁰ Sensibility means that the test must seem rational (make sense) to clinicians and be reasonably easy to use. An earlier attempt to develop a frailty scale from information on cognition and function¹¹ was criticized as perhaps being too easy to use, in that it simply rearranged known disability and impairment data. Although it too demonstrated a dose response in relation to mortality and to institutionalization, its prespecified patterns of impairment did not always conform to what was observed in patients. Moreover, although linked with disability, in that many frail people are disabled, frailty is not the same as disability, as evidenced by the fact that not all disabled people are frail.

At the heart of geriatric medicine is comprehensive geriatric assessment (CGA). The goal, in a secondary analysis of data from the Mobile Geriatric Assessment Team (MGAT) trial,^{12,13} was to construct an FI, investigate its construct validity, assay its predictive validity, test its interrater reliability, and consider its sensibility. These objectives allowed the matter of whether the frailty status of older adults can be summarized in a way that is sensible and practical for geriatricians to apply to be addressed.

From the *Faculty of Medicine, Memorial University, St. John's, Newfoundland, Canada; and [†]Division of Geriatric Medicine, Dalhousie University, Halifax, Nova Scotia, Canada.

The MGAT study was supported by Health Canada through NHRDP Grant 6603-1417-302. Kenneth Rockwood is supported by an Investigator Award from the Canadian Institutes for Health Research and by the Dalhousie Medical Research Foundation as the Kathryn Allen Weldon Professor of Alzheimer Research.

Address correspondence to Kenneth Rockwood, Center for Health Care of the Elderly, 1421–5955 Veterans' Memorial Lane, Halifax, NS B3H 2E1, Canada. E-mail: kenneth.rockwood@dal.ca

METHODS

Study Setting, Design, and Subjects

The MGAT study was a randomized, controlled trial in rural, community-dwelling frail elders in three counties around Halifax, Nova Scotia.^{12,13} The intervention consisted of multidisciplinary specialized geriatric assessment and management by the MGAT team for 3 months. The control group received usual care. Goal Attainment Scaling (GAS),¹⁴ a measure of achieving individualized and clinically relevant goals, adapted for use in a specialized geriatric service,¹⁵ was the primary outcome, calculated as the extent to which goals were met at 3 months. Secondary measures included another global index of function and cognition (the Geriatric Status Scale),⁹ physical and daily functioning (the Barthel Index and the Lawton-Brody Physical Self-Maintenance Scale and instrumental activity of daily living/activity of daily living (IADL/ADL) scale,^{16,17} and cognition (the Mini-Mental State Examination (MMSE)).^{18,19}

To be eligible, a patient had to be frail, which was defined as “a vulnerable state of health, arising from the complex interaction of medical and social problems, resulting in a decreased ability to respond to stress, and associated with a decline in functional performance.”¹² Of 265 older patients referred to the study, 54 refused or withdrew before baseline, 27 did not meet inclusion criteria, and two died before assessment. Of the 182 enrolled, 95 were randomized to intervention and 87 to the control group. Two in the intervention and four in the control group refused assessment from all raters, whereas one in the intervention and four in the control refused assessment by at least one rater, and two in the control group had missing values, leaving 169 (92 intervention and 77 control) patients with complete data for the evaluation of validity and interrater reliability.

Each of three raters, i.e., the geriatric nurse assessor who examined patients in their own homes, the geriatrician who led the multidisciplinary team, and a second geriatrician who reviewed each of the cases and all of the forms, completed the CGA separately.^{12,13} Each rater was blind to the other ratings, from which the interrater reliability was calculated. Where disagreements existed, discussion between raters and team members led to a consensus rating of the existence of the problem, and degree of severity was used to score and construct the FI-CGA, following discussion.

Constructing the FI-CGA

Following the method used in population-based samples,^{1,6–8} the baseline FI-CGA was calculated as a count of the impairments identified at the baseline CGA. The standardized CGA used to calculate the FI comprises assessments in 10 standard domains: cognitive status (rated as no cognitive impairment = no problem; cognitive impairment, no dementia = mild problem; delirium or dementia = severe problem), mood and motivation (each rated separately and then combined so that the highest level of specificity was scored for the domain), communication (vision, hearing, speech), mobility, balance (each of the latter two scored at the highest level of independence with aids where used), bowel function, bladder function, IADLs and ADLs (rated as no impairment = no problem, IADL impairment = mild problem, ADL impairment = major prob-

lem), nutrition, and social resources (scored as a problem if there was need for additional help).²⁰ Problems in each domain were scored as 0 (no problem), 1 (a minor problem), or 2 (a major problem). For evaluating the contribution of each domain, the mode of the three ratings determined the value for each subject.

The CGA also records the number of comorbidities.²¹ Balancing the superior performance of weighted versus unweighted indices²² against the need to retain simplicity for clinical use, the number of comorbidities was divided by two to achieve appropriate weighting between impairments (e.g., reduced glomerular filtration rate due to small vessel disease) and their consequences (chronic renal failure). (In this example, as in others, there would be important double-counting of disease, because antecedent and consequent items—here, diabetes mellitus and chronic renal failure—would typically be included in a list of comorbidities.) Simple addition of the count of CGA-identified problems, plus the comorbidity count of active diagnoses yields the FI-CGA score.

Because the index was constructed post hoc, and given that clinical sensibility requires the index to discriminate between levels of frailty, how to grade severity was explored. Computed FIs have shown a dose response across a range of impairments.^{1,6–8} A rules-based FI suggested that three grades of frailty (mild, moderate, severe) could be identified.¹¹ Natural cutpoints within the distribution of the FI-CGA scores were evaluated by observing the frequency distribution of the patients at various FI-CGA levels to identify the points of separation; then whether these demonstrated a dose response for adverse outcomes was assessed. Because there is no biological referent standard against which a clinical index can be compared, predictive validity (i.e., demonstrating a dose response in relation to relevant and nonarbitrary outcomes) is the highest form of criterion validation available.⁹ The outcomes that were selected were time to death or institutionalization, whichever came first. These were selected as traditional, dichotomous, easily distinguishable, relevant, and unequivocal.

To test the construct validity of the FI-CGA, it was examined in relation to other health status measurements at baseline. In general, frail people are more likely to be disabled.^{2,3} Consequently, the cutpoints were also tested against increasing levels of functional impairment.

A tripartite hierarchy from earlier analyses^{2,3,24} (of dependence in basic and intermediate self-care, and complex self-management) was used to combine IADL/ADL items scored as independent, assisted, or dependent. For these analyses only, the FI-CGA was recomputed to exclude functional disability items from the count, and mild moderate and severe frailty scores were adjusted accordingly.

Analysis

Scores on traditional geriatric health status measures (e.g., Barthel Index, MMSE, Physical Self-Maintenance Scale) were correlated (Pearson correlation) with the FI-CGA. The predictive validity of the FI-CGA for 12-month adverse outcomes (death or institutionalization, whichever occurred first) was evaluated by constructing Kaplan-Meier survival curves and by calculating hazard rates for institutionalization or death. Interrater reliability was tested by

Table 1. Characteristics of Patients by Degree of Frailty

Characteristic	Mild Frailty (n = 29)	Moderate Frailty (n = 89)	Severe Frailty (n = 42)
Female, %	55	60	55
Married, %	45	47	47
Living alone, %	55	60	14
Intervention, %	55(n = 20)	45	59
Age, mean \pm SD	82.7 \pm 7.9	81.8 \pm 6.4	81.7 \pm 8.6
Poor self-rated health, %	4 (n = 23)	4 (n = 71)	32 (n = 25)
Physical Self-Maintenance Scale score, mean \pm SD	7.4 \pm 1.8	9.3 \pm 3.5	13.4 \pm 4.7
Mini-Mental State Examination score, mean \pm SD	24.0 \pm 6.1	23.8 \pm 0.5	19.2 \pm 8.4
Barthel Index score, mean \pm SD	94.8 \pm 8.7	88.7 \pm 14.2	68.1 \pm 20.7
Geriatric Status Scale score, mean \pm SD	2.1 \pm 0.5	2.5 \pm 0.7	3.3 \pm 0.9
Spitzer quality of life index, mean \pm SD	11.8 \pm 1.6	11.2 \pm 9.2	7.9 \pm 2.0
Geriatric Status Scale total score at baseline, mean \pm SD	35.4 \pm 0.86	35.0 \pm 0.76	34.1 \pm 0.91
Number of goals, mean \pm SD	4.3 \pm 1.3	5.2 \pm 1.5	7.2 \pm 2.5

Note: n = the total number of subjects of the category unless n is specified in the cells. SD = Standard deviation.

comparing the scores of the three raters using an intraclass correlation coefficient. The hazard ratios (HRs) were calculated from a proportional hazards model. To identify the possible confounding effects of demographic factors and intervention status, the odds ratios of the FI-CGA were adjusted for age, sex, marital status, and treatment group. Receiver operating characteristic (ROC) curves²⁵ were used to compare the performance of the FI-CGA with other scores in predicting the adverse outcomes.

RESULTS

The mean \pm standard deviation value of the FI was 7.2 \pm 3.0 (median 6.7, range 1–19.2). Observing the frequency distribution of FI-CGA scores suggested three intervals corresponding to degrees of frailty: mild (FI-CGA 0–7), moderate (FI-CGA 7–13), frailty (FI-CGA \geq 13). Comparing patient characteristics and expected outcomes provided

further support for interval selection, so that 29 patients (17%) were classified as mildly frail, 98 (58%) as moderately frail, and 42 (25%) as severely frail.

Table 1 shows the characteristics of patients by grade of frailty. Patients with different levels of frailty are comparable in their social and demographic characteristics, although fewer intervention patients were classified as mildly frail. As expected, the mean values of the health measures differed considerably between each level of frailty, with worse function corresponding to each advancing level of frailty. In addition, frailer patients were identified as needing to achieve a greater number of goals.

Table 2 shows the unadjusted and adjusted HRs for the FI-CGA and each individual functional domain included in the FI-CGA. The HRs of the FI-CGA were significant in relation to the adverse outcomes, whereas only three of 11 of its individual attributes (cognition, IADL/ADL function, and social behavior) showed a significant HR. No

Table 2. Predictive Validity of the Items That Constitute the Frailty Index from a Routine Comprehensive Geriatric Assessment (FI-CGA)

Attribute	Unadjusted	Adjusted
	Hazard Ratio (95% Confidence Interval)	
Cognitive (n = 170)	1.55 (0.99–2.44) [†]	1.75 (1.08–2.84) [†]
Mood (n = 171)	1.18 (0.76–1.83)	1.30 (0.81–2.08)
Communication (n = 169)	1.17 (0.69–1.96)	1.21 (0.70–2.07)
Mobility (n = 170)	1.26 (0.77–2.05)	1.21 (0.73–2.00)
Balance (n = 169)	1.27 (0.82–1.98)	1.25 (0.80–1.95)
Bowel (n = 170)	1.28 (0.68–2.43)	1.34 (0.70–2.57)
Bladder (n = 169)	1.30 (0.75–2.25)	1.26 (0.72–2.22)
Nutrition (n = 170)	1.15 (0.63–2.08)	1.18 (0.63–2.19)
Instrumental activities of daily living/ activities of daily living (n = 171)	1.75 (1.05–3.09) [†]	1.84 (1.04–3.24) [†]
Social (n = 170)	1.91 (1.05–3.47) [†]	1.95 (1.07–3.55) [†]
Comorbidity (n = 182)	0.94 (0.57–1.42)	0.90 (0.61–1.46)
FI-CGA (n = 169)	1.12 (1.01–1.24) [†]	1.23 (1.01–1.45) [†]

* Adjusted for age, sex, marriage status, and status of intervention.

[†] P < .05.

significant differences were found between the adjusted HRs and the unadjusted ones for any attribute, and none of the adjusting factors (age, sex, marital status, and intervention status) were found to be significantly related to the outcomes (HR = 1.04–1.56, $P = .15$ –.66). When tested in a multivariate logistic regression model consisting of the 11 attributes of the FI-CGA, the correlation coefficient (r) of any single attribute in relation to the adverse outcome was not significant ($P = .15$ –.89).

The FI-CGA was notionally correlated ($r = 0.33$) with the MMSE and moderately correlated ($r = \sim 0.55$) with each of the measures of function and with the comorbidity index ($r = 0.57$). Another way to conceptualize degrees of frailty in frail samples is to consider the relative degree of functional dependence. Dependence in complex self-care was high in each group but least common in those who were mildly frail (72%), of whom most were independent in basic self-care (62%). By contrast, almost all patients rated as moderately or severely frail showed dependence in complex self-management (93% and 100%, respectively). A readily clinically recognizable difference between those with moderate and severe frailty was that 17% of the former, compared with 41% of the latter, showed dependence in personal ADLs. People with moderate frailty were more likely (9%) to require only assistance with their personal ADLs than were those classified as showing severe frailty, of whom 41% required only assistance with most personal ADLs.

Higher levels of frailty are associated with increasing risk of an adverse outcome of death or institutionalization (Figure 1). In general, the risk of an adverse outcome is highest for those in whom frailty is severe, but the Kaplan-Meier curves for mild and moderate frailty cross during the first 5 months, suggesting that the hazards are not proportional. In consequence, a proportional hazards model was not estimated. The unadjusted HRs (compared with mild

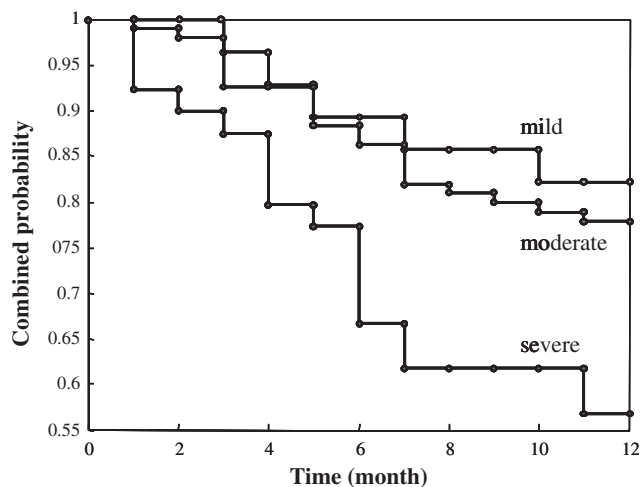


Figure 1. Kaplan-Meier curves for differing grades of frailty for time to an adverse outcome. The patients were followed for up to 12 months, and their time to death or time to institutionalization (whichever came first) was recorded. Patients were divided into three groups based on their comprehensive geriatric assessment frailty index (mild, moderate, and severe frailty). The proportions of people who survived in the community are plotted against time.

frailty) of moderate and severe frailty were 1.9 (95% confidence interval (CI) = 1.7–2.1) and 5.5 (95% CI = 3.6–7.4), respectively.

The interrater reliability of the FI-CGA, using three raters, was 0.95 for the baseline assessment. At 3 months, the value was 0.96. The area under the ROC curve for the FI-CGA with the IADL/ADL items was 5% higher than that for the disability score. Without the IADL items, it was 3% higher than that for the disability score.

DISCUSSION

Routinely collected data from a standardized CGA was used to construct an FI in a secondary analysis from a controlled trial of a specialized geriatric intervention. The need to translate from recent work on developing a relative fitness/FI in population studies to possible use in a clinical setting motivated this study. The resulting FI-CGA demonstrated appropriate construct validity. A higher FI-CGA score was associated with a higher risk of death or institutionalization. The measure showed good interrater reliability. In consequence, it appears that the lessons from earlier analyses—that frailty can be measured precisely in population studies—can be extended to clinical settings. Combining items from a standard CGA appears to yield a workable FI, giving the possibility of deriving an FI from routinely collected data. In this way, these analyses demonstrate that an index-based approach to measuring frailty is clinically sensible and discriminates between groups of patients.

These data must be interpreted with caution. A clinical trials sample poses special challenges of generalizability. Moreover, in the MGAT trial, all the patients were frail, so there was no fit unimpaired referent group, but discriminating grades of frailty is not a trivial task and is a common clinical challenge for geriatricians.^{1–3} Although three levels of frailty could be distinguished, these less persuasively predicted grades of adverse outcomes—for example, the stratification of mild, moderate, and severe frailty showed discriminant validity for adverse outcomes, but only modestly for the mild and moderate categories, and only after 6 months. In addition to having no unimpaired referent group, this observation raises the questions of what items to include and how best to combine them. Each item in the CGA was included as clinically necessary²⁰ and after a validation exercise that focused on the assets and deficits that interact dynamically to give rise to frailty.²⁶ It is also not clear whether other weighting schemes, such as an artificial neural network, might produce better discrimination, which seems likely,²² but that would have to be balanced against the clinical transparency of this approach. The rating of a problem as none, minor, or major introduces some arbitrariness and raises the question of major/minor to whom? In the MGAT trial, this question was answered by focusing on the patient. Problems were evaluated as major or minor from the patient/caregiver point of view. This is arguably the usual standard in geriatric medicine, with its pragmatic emphasis on comfort and function. In the MGAT trial, through the use of GAS as the primary outcome, patient-centeredness was the standard by which the intervention was evaluated. Given that the GAS scores at 3 months favored the intervention over usual care in each frail group (mild = 50.0 ± 6.6 vs 39.3 ± 3.9 , moderate = 47.6 ± 5.4

vs 39.2 ± 3.9 , severe = 45.3 ± 5.2 vs 38.0 ± 3.4) and that the interrater reliability was high, these judgments can be made in a clinical setting.

Functional impairment is generally seen as integral to frailty.^{11,18,27–29} These data suggest that the relationship between frailty and functional decline (correlation ~ 0.55) follows the hierarchical pattern suggested by population studies^{11,22,23} but that frailty is more than just disability, as evidenced here by the combination of impairment, disability, and disease that constitute the index. Still, the ROC analyses showed that the current combination has only a slightly better performance than the disability measure, although there may be other reasons to prefer an FI. Other analyses have shown that the FI has additional important properties, such as a power law relationship to mortality.^{8,30} The CGA also provides a more comprehensive plan for clinical intervention than does a disability inventory. Still, the relationship between frailty and disability requires further disentanglement, and this is an active area of investigation. As argued elsewhere,³⁰ frailty appears to operate analogously to macroscopic state variables (such as temperature measured by the summary of the average kinetic energy of the component molecules) in physical systems. It may be that disability has similar attributes; indeed, it might be that other items (such as muscle wasting or inflammatory markers) are important for this reason too, although this proposition also requires further evaluation.

Although moderately and severely frail older adults are at a greater risk of adverse outcomes, frailty did not affect their ability to achieve individualized goals. The 3-month mean GAS scores for the mildly, moderately, and severely frail groups were 44.6 ± 7.41 , 44.9 ± 6.5 , and 44.8 ± 7.7 , respectively. Given that the randomization was not balanced with respect to the grade of frailty, the effect of the intervention on the secondary outcomes was also recalculated by grade of frailty.

How to best measure frailty remains unresolved. The merit of the approach offered is that it can be done readily from data that a geriatrician would need to gather to decide on interventions. As such, it requires no special instrumentation, beyond the clinical skills needed to do a CGA. In this context, whether the FI-CGA is a sensible measure can also be addressed. There is no easy numeric test for the assessment of sensibility, but the FI-CGA is self-evidently a sensible measure by virtue of its complete account of factors theorized to influence frailty. It is easy to use in the setting of specialized geriatric assessment, is readily calculated, and provides a means of risk stratification. In conclusion, it should be studied prospectively in other clinical settings.

REFERENCES

- Rockwood K, Mitnitski AB, MacKnight C. Some mathematical models of frailty and their clinical implications. *Rev Clin Gerontol* 2002;12:109–117.
- Fried LP, Darer J, Walston J. Frailty. In: Cassel CK, Leipzig RM, Cohen HJ et al., eds *Geriatric Medicine: An Evidence-Based Approach*, 4th Ed. New York: Springer-Verlag, 2003, pp 1067–1076.
- Hogan DB, MacKnight C, Bergman H et al. Models, definitions, and criteria of frailty. 2003;15(3 Suppl):1–29.
- Chin A, Paw MJM, Dekker JM et al. How to select a frail elderly population? A comparison of three working definitions. *J Clin Epidemiol* 2000;52:1015–1021.
- Morley JE, Perry HM III, Miller DK. There's something about frailty. *J Gerontol A Biol Sci Med Sci* 2002;57A:M698–M704.
- Mitnitski AB, Song X, Rockwood K. The estimation of relative fitness and frailty in community dwelling older adults using self-report. *Data. J Gerontol A Biol Sci Med Sci* 2004;59A:M627–M632.
- Mitnitski AB, Graham JE, Mogilner AJ et al. Frailty, fitness and late-life mortality in relation to chronological and biological age. *BMC Geriatr* 2002;2:1.
- Mitnitski AB, Mogilner AJ, MacKnight C et al. The mortality rate as a function of accumulated deficits in a frailty index. *Mech Ageing Dev* 2002;123:1457–1460.
- Streiner DL, Norman GR *Health Measurement Scales. A Practical Guide to Their Development and Use*. Oxford: Oxford University Press, 1995.
- Feinstein AR *Clinometrics*. New Haven: Yale University Press, 1987.
- Rockwood K, Stadnyk K, MacKnight C et al. A brief clinical instrument to classify frailty in elderly people. *Lancet* 1999;353:205–206.
- Rockwood K, Stadnyk K, Carver D et al. A clinimetric evaluation of specialized geriatric care for rural dwelling, frail older people. *J Am Geriatr Soc* 2000;48:1080–1085.
- Rockwood K, Howlett SE, Stadnyk K et al. Responsiveness of goal attainment scaling in a randomized controlled trial of comprehensive geriatric assessment. *J Clin Epidemiol* 2003;56:736–743.
- Kiresuk TJ, Sherman RE. Goal attainment scaling: A general method for evaluating comprehensive community mental health programs. *Community Mental Health J* 1968;4:443–453.
- Stolee P, Rockwood K, Fox RA et al. The use of goal attainment scaling in a geriatric care setting. *J Am Geriatr Soc* 1992;40:574–578.
- Lawton MP, Brody E. Assessment of older people: Self-maintaining and instrumental activities of daily living. *Gerontologist* 1969;9:179–186.
- Collin C, Wade DT, Davies S et al. The Barthel ADL Index: A reliability study. *Int Disabil Stud* 1988;10:61–63.
- Folstein MF, Folstein SE, McHugh PR. 'Mini-mental state' A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189–198.
- Cockrell JR, Folstein MF. Mini-Mental State Examination (MMSE). *Psychopharmacol Bull* 1988;24:689–692.
- Rockwood K, Silivius JL, Fox RA. Comprehensive geriatric assessment. *Postgrad Med* 1998;12:189–198.
- de Groot V, Beckerman H, Lankhorst GJ et al. How to measure comorbidity. A critical review of available methods. *J Clin Epidemiol* 2003;56:221–229.
- Song X, Mitnitski A, MacKnight C et al. Assessment of individual risk of death using self-report data: An artificial neural network compared with a frailty index. *J Am Geriatr Soc* 2004;52:1180–1184.
- Rockwood K, Howlett SE, MacKnight C et al. Prevalence, attributes and outcomes of fitness and frailty in community-dwelling older adults. Report from the Canadian Study of Health and Aging. *J Gerontol A Biol Sci Med Sci* in press.
- Thomas VS, Rockwood K, McDowell I. Multidimensionality in instrumental and basic activities of daily living. *J Clin Epidemiol* 1998;51:315–321.
- Metz CE. Basic principles of ROC analysis. *Semin Nucl Med* 1978;8:283–298.
- Rockwood K, Fox RA, Stolee P et al. Frailty in elderly people: An evolving concept. *Can Med Assoc J* 1994, 489–495.
- Carlson JE, Zocchi KA, Bettencourt DM et al. Measuring frailty in the hospitalized elderly: Concept of functional homeostasis. *Am J Phys Med Rehabil* 1998;77:252–257.
- Rozzini R, Frisoni GB, Franzoni S et al. Change in functional status during hospitalization in older adults: A geriatric concept of frailty. *J Am Geriatr Soc* 2000;48:1024–1025.
- Wells JL, Seabrook JA, Stolee P et al. State of the art in geriatric rehabilitation Part I & II. Review of frailty and comprehensive geriatric assessment. *Arch Phys Med Rehabil* 2003;84:890–903.
- Mitnitski AB, Rockwood K. Help available. Phenomenological models for research on aging. *Sci Aging Knowledge Environ* 2003;12:VP2.