

Validity of an early risk score for older adults

¹PD St John, ²PR Montgomery

¹Associate Professor, Section of Geriatric Medicine, Department of Medicine, and the Centre on Aging, University of Manitoba, Winnipeg, Manitoba; ²Consulting Geriatrician, Vancouver Island Health Authority, Victoria, British Columbia, Canada

ABSTRACT

Objectives: To determine if a risk score developed in hospitalised older adults in the UK in 1962 is correlated with other measures of health and if this risk score predicts death or institutionalisation in community-living older adults.

Methods: A total of 1,735 older adults residing in the community in 1991 were followed over five years. We replicated the original risk index, a composite score of cognitive status, disability and continence. Other measures included age, gender, education, self-rated health (SRH), life satisfaction (LS) and frailty. Death and nursing home (NH) admission were determined five years later.

Results: The risk score was strongly associated with frailty, SRH and LS. The index predicted mortality and NH use. The adjusted odds ratio (95% confidence interval) from multinomial logistic regression models was 0.74 (0.69, 0.79) for death and 0.74 (0.67, 0.83) for NH.

Conclusions: A risk score devised to measure inpatient rehabilitation also predicts outcomes in community-dwelling older adults. Cognition and function predict a variety of adverse outcomes in a variety of settings in different historic eras.

KEYWORDS comprehensive geriatric assessment, disability, frailty, institutionalisation, mortality, risk index, risk score

DECLARATION OF INTERESTS No conflict of interest declared.

BACKGROUND

Prognostication has been an important part of medical care since the time of Hippocrates. Prognostic indices have been developed to predict the risk of adverse events in general populations of older persons,^{1–3} in hospitalised older adults^{2,4} and in those with specific disease states.^{5,6} Many of these instruments have been developed in recent decades. However, risk scores were used much earlier in the history of geriatric medicine. To evaluate an early geriatric unit, Arnold and Exton-Smith⁷ devised a score that incorporated activities of daily living (ADL), instrumental activities of daily living (IADL), continence and cognition (Table 1). They used this scale both to predict outcomes and to evaluate the success of rehabilitation. It was administered in hospital, and repeated after discharge. This score was a strong predictor of discharge to home three months later, and sensitive to clinically relevant changes in patient status.

This risk score represents an early frailty measure, and is remarkably similar to more recent measures of frailty, which also try to quantitate risk and vulnerability of older people. The importance of cognitive and functional status was recognised in this early score. Over the intervening decades, a large body of literature has developed that has documented the importance of these factors, both as prognosticators and as outcomes. Yet they are not universally incorporated into clinical decisions or administrative decisions, which increasingly focus on technical and biological measures and interventions. The original scale itself was never widely disseminated.

Correspondence to PD St John
Section of Geriatrics
University of Manitoba
GE 547 Health Sciences Centre
820 Sherbrook Street
Winnipeg, MB R3A 1R9, Canada

tel. +1 204 787 3365
e-mail pstjohn@hsc.mb.ca

TABLE 1 Risk score devised by Arnold and Exton-Smith

Mental state				
Normal, 4	Slight impairment, 3	Moderate impairment, 2 ;	Gross impairment, 1	Coma, 0
Incontinence				
Not, 2	Moderate, 1	Severe, 0		
Physical state				
<i>Mobility</i>	<i>Daily living activities</i>			
Out of doors	Feeding			
Climb stairs	Dressing			
Indoors	Washing			
Get in and out of chair	Shaving (men)/			
Get in and out of bed	Attention to hair (women)			
	Cooking			
	Cleaning			
Performed unaided, 2	With help, 1	Unable to do it/ never did it, 0		

This scale was devised prior to modern methods of scale development and validation,⁸ but is easily performed and measures clinically important factors. It could be used to measure risk for clinical and administrative reasons in a variety of settings, including community populations. One of the criticisms of current risk instruments is that they have not been validated in populations different from the derivation cohort.⁹ We therefore sought to validate the scale, as closely as possible, in a very different population at a very different

time, using an existing data set that included similar variables. Reinforcing the importance of measuring physical and cognitive functioning in all care settings and at various historic periods is needed to ensure that they are not forgotten in the modern era, which focuses increasingly on technical measures. Specifically, the objectives were:

1. To determine if a modified risk score developed in hospitalised older adults in the UK in 1962 is correlated with measures of health and frailty;
2. To determine if this risk score predicts death or institutionalisation in community-dwelling older adults over a five-year period; and
3. To determine if a very straightforward modification considering age and gender further increases the predictive accuracy for death or institutionalisation.

METHODS

Sample

Manitoba is a province in the prairie region of Canada. It had a stable population of 1,091,940 in 1991, with 13.4% elderly.¹⁰ Data used in this study are from the Manitoba Study of Health and Aging (MSHA), a prospective cohort study conducted in conjunction with the Canadian Study of Health and Aging (CSHA).¹¹ The MSHA enriched the CSHA sampling frame, including participants from all regions of the province, and added extra measures. The MSHA received ethics approval from the Research Ethics Committee of the University of Manitoba and adhered to the Helsinki Declaration. Informed consent was obtained from the participants or the appropriate proxy.

Persons over the age of 65 who were living in the community were randomly selected according to health region and age group from a list provided by Manitoba Health. Since healthcare is universal in Manitoba, this is one of the most complete listings of residents available. There was an over-sampling of the oldest (>85 years). The original survey was conducted in 1991/92, and follow-up was conducted in 1996/97. Initially, 2,890 people were selected. Of these, 443 refused to participate, 480 were not eligible (had died, entered a nursing home or were too ill), 162 could not be located and 54 did not complete the screening questionnaire. There were ten participants who were missing data necessary to recreate the risk score, and six whose time 2 status was unknown. This resulted in a sample of 1,741 for cross-sectional analyses and 1,735 for prospective analyses.

Measures

Trained interviewers visited the participants in their homes. Age, gender, living arrangement, education and self-rated health (SRH)^{12,13} were all self-reported. After the ADL questions, participants were asked: 'How much do your health troubles stand in the way of doing the things you want to do? Not at all; a little; a great deal.'

Life satisfaction (LS) was assessed using the Terrible–Delightful Scale.¹⁴ There were 12 questions, each rated on a seven-point Likert-like scale, scored from 0 (terrible) to 7 (delightful). We considered two items: health ('The present state of your general health [relatively free of common chronic illnesses]'), and general LS ('Now, how do you feel about your life as a whole right now? Is your life generally dissatisfying, satisfying, etc.?').

We also considered two measures of frailty – the brief measure of frailty¹⁵ and the Frailty Index.^{16,17} Both of these measures were defined according to the original definitions. The brief measure of frailty is a measure of disability and cognition, which is graded from 0 (no frailty) to 4 (severe frailty). The Frailty Index is a tally of self-reported health complaints, diseases, risk factors and impairments. It measures 40 items, each scored 0 to 1. These are summed and divided by the number of deficits considered.

Risk score

We attempted to replicate the original risk score as closely as possible in the MSHA. The original score used direct observation of which ADLs the participant was able to do, whereas the MSHA used self-report. The original score included one item which the MSHA did not: the ability to transfer from a chair. The original score used a three-point measure of urinary incontinence, where we used a two-point item. Finally, the measure of cognition used in the original risk score was not reported, where the MSHA used the modified Mini-mental State Examination (3MS)¹⁸ from which the Mini-mental State Examination (MMSE)¹⁹ could be derived. The scale was devised before the MMSE was used, and there were few reliable and valid cognitive measures, although some were being devised at the same time.^{20,21} Otherwise, the score could be replicated very closely.

To measure cognitive status, we considered the MMSE. There were no comatose participants. We considered severe impairment to be less than 20; moderate impairment to be an MMSE of 20–23; mild impairment to be an MMSE of 24–26; and normal to be greater than 26. We considered different categories, which gave similar results. Continence was measured with the question: 'Do you ever lose control of your bladder?' and was scored yes/no. Functional status was measured using the Older Americans Resource Survey (OARS),²² with some additional items unique to the MSHA.

Participants were asked about their ability to perform basic ADLs (eating, dressing, grooming, getting in and out of bed, taking a bath or shower and ability to use the bathroom), IADLs (using the telephone, getting to places out of walking distance, going shopping, preparing meals, doing housework, taking medication[s], managing money) and ambulation. These were all initially coded as: without

any help; with help from a device; with some help from a person; with help from a device and a person; and unable to do. To replicate the scale, we used items matching those of Arnold and Exton-Smith (Table 1). Since the weather can occasionally be inclement, the MSHA included two items on the ability to go outdoors: 'Can you go out of doors in good weather?' and 'Can you go out of doors in any weather?' We considered the ability to go in out in good weather for these analyses, since it is likely more generalisable to other regions.

For all the items in the original risk score, we used the corresponding item from the OARS. Since age and gender remained significantly associated with mortality and institutionalisation, we modified the score by adding a point for male gender, and by adding a point for each five-year age grouping after the age of 65.

Outcomes

We used the status five years later. Death was ascertained by proxy, administrative data or vital statistics data. Nursing home (NH) placement was measured at time 2 by proxy report or administrative data. Some participants were admitted to a NH and then died within the five-year interval. We considered these individuals within the deaths, and not a NH.

Analysis

For bivariate analyses, we used chi-square tests for categorical variables. For continuous variables, we used Student's t-tests (assuming unequal variance). To determine if the risk score predicted death or institutionalisation over the five-year period, we constructed multinomial logistic regression models. Here, we adjusted for age, gender and education. The outcomes were dead at time 2 vs in a NH at time 2 vs alive and in the community at time 2. To determine the predictive accuracy, we constructed receiver operator characteristic (ROC) curves and used the original risk score as well as the score that considered age and gender. We then calculated the area under the curve (AUC) for the ROC curve using a c-statistic. We also considered each item individually in sensitivity analyses. Data manipulation and graphs were done in SPSS 8 (SPSS Inc., Chicago) and data analysis was done in Stata 10 (Statacorp, College Station).

RESULTS

The mean score was 24.3 (SD 2.0). The baseline characteristics of the sample are shown in Table 2. Those with lower scores on the risk score were older and had less education. The score was strongly related to other measures of health and well-being: those with lower scores were more likely to report poor SRH, have lower levels of LS with health, lower levels of LS overall, and report more troubles with daily living. The risk score was correlated with the brief frailty measure (Spearman's correlation coefficient of 0.64, $p < 0.001$), and the Frailty

TABLE 2 Characteristics of sample at time 1

	Risk score at time 1			
	<20 (N=72)	20-24 (N=611)	>24 (N=1,058)	Total sample (N=1,741)
Age (mean years)	83.6*	80.4*	75.5*	77.6
Gender (% women)	69.4%	59.4%	57.8%	58.8%
Education (mean years)	5.4*	7.8*	9.6*	9.3
Living arrangement (% living alone)	45.8%	49.8%	38.4%*	42.7%
Self-rated health (% poor)	35.2%*	31.9%*	20.1%*	24.9%
Health troubles				
A great deal	78.6%*	37.7%*	10.6%*	22.9%
A little	10.0%*	34.6%*	35.3%*	34.0%
Not at all	11.4%*	27.7%*	54.1%*	43.1%
LS health (mean)	3.9*	4.1*	4.9*	4.8
LS overall (mean)	4.7*	4.8*	5.3*	5.2
Frail (% frail by brief frailty measure)	94.4%*	46.7%*	4.9%*	22.4%
Frailty Index (mean)	0.47*	0.31*	0.18*	0.20
Five-year mortality (% dead)	60.0%*	35.2%*	15.0%*	23.9%
Five-year institutionalisation (% in NH)	21.4%*	11.5%*	3.3%*	6.9%

* denotes $p < 0.05$.

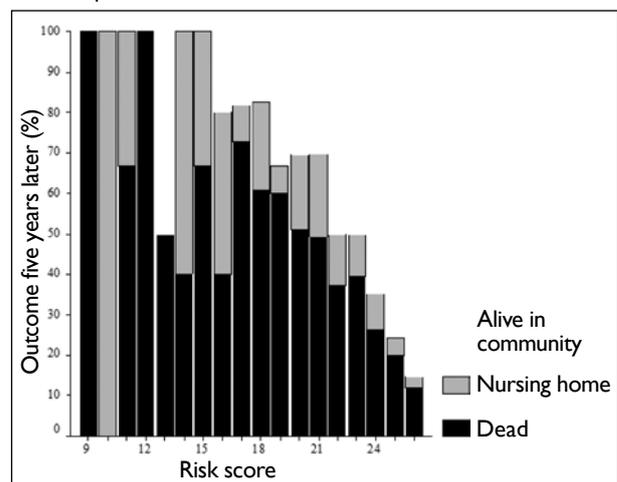


FIGURE 1 The risk score predicts death and NH placement over a five-year interval in community-dwelling older adults.

Index (Pearson's correlation coefficient of 0.64, $p < 0.001$). At the time 2 follow-up, 414 participants were dead, 120 were in a NH, and 1,201 were living in the community. Those with lower risk scores were much more likely to die or be living in a NH. This effect appeared to be a gradient across the range of the score (Figure 1). This effect persisted after adjustment for age, gender and

TABLE 3 Results of multinomial logistic regression model. The risk of dying or residing in a NH five years later is shown, compared with being alive and in the community

	Status at time 2 (reference = alive in community)	
	Nursing home	Dead
Risk score, unadjusted (per point)	0.64 (0.59, 0.69)	0.67 (0.63, 0.71)
Risk score, adjusted (per point)	0.74 (0.67, 0.83)	0.74 (0.69, 0.79)
Age (per year)	1.14 (1.11, 1.17)	1.09 (1.07, 1.11)
Gender (ref=male)	1.54 (0.98, 2.45)	0.51 (0.40, 0.66)
Education (per year)	0.95 (0.89, 1.01)	0.99 (0.96, 1.02)

education (Table 3). Age and gender remained strongly predictive for death and institutionalisation as well. Furthermore, each item of the scale predicted both of these outcomes. Self-rated health, LS with health, health troubles interfering with ADLs, and both measures of frailty were also associated with death and NH admission.

We constructed ROC curves. The risk score was fairly accurate in predicting death over the five-year period, with an AUC (95% CI) of 0.69 (0.66, 0.72). This was superior to either the AUC for either the brief frailty measure (AUC=0.62) or the Frailty Index (AUC=0.63). This accuracy was increased when age and gender were considered in the risk score: the AUC for this modified score was 0.74 (0.71, 0.77). The risk score also predicted admission to a NH over the five-year interval (here, we included all 211 participants who had been admitted into a NH, whether they were dead or alive at time 2). The AUC for the original score was 0.76 (0.72, 0.80); and 0.81 (0.78, 0.84) for the score modified by age and gender.

DISCUSSION

We have found that a risk prediction score developed in 1962 for the assessment of older adults in hospital is associated with broad measures of health and frailty in older adults in the community in the 1990s. Furthermore, this simple instrument predicts death and institutionalisation with surprising accuracy. Indeed, the predictive accuracy is comparable with many modern risk measures, and superior to measures of frailty in predicting death. In spite of very different populations at very different times, functional status and cognitive status strongly predict adverse outcomes.

There are both strengths and limitations to our approach. First, we were unable to exactly replicate the score, since one item was missing from our data set (the ability to transfer from a chair). Furthermore, the measurement of cognitive status was not clear in the original paper, and we therefore used categories of the MMSE. Second, we lacked the data to recreate more modern risk instruments, which often measure medical conditions or

biomedical factors. We therefore cannot compare the risk instrument with more current risk indices. However, we were able to compare it with more recent measures of frailty, compared with which it performs well.

Strengths of our approach are that we had other reliable and valid measures of health and well-being which were gathered by trained interviewers. We also had a sample drawn from a representative population-based sampling frame of an entire province. This broadens the generalisability of the original score. There was also complete data collection, and few participants had missing data at time 2.

Further validation in other data sets of other populations would be important. In particular, more recent cohorts and samples in clinical settings would help to validate the scale further, and to generalise the findings to other populations. We also noted the strong effect of simple self-reports, such as SRH on mortality. This association was also noted some time ago,¹³ yet simple SRH is often not considered in clinical settings.

Our findings are important for several reasons. First, the importance of functional and cognitive status is reinforced. These factors are very strongly associated with other measures of health, and very strong predictors of outcome. While this has been long noted, they may not always be incorporated into clinical decision-making²³ or in case mix adjustment tools which tend to focus predominantly on biomedical and physiologic data.²⁴ Second, the negative effect of accumulating deficits has been apparent for a long time, and current measures of risk and frailty have marked similarities to this older scale. The risk score of Arnold and Exton-Smith could be seen as the precursor to more contemporary frailty measures, which are in turn based on the tradition of comprehensive geriatric assessment and a third important finding is that simple self-reported measures predict outcomes. All the items we considered were easy to gather, inexpensive and should be routine in clinical practice. A central feature of geriatrics is the acceptance of complexity. Yet there is a parallel tradition of simply asking people about their health, and considering these broad general assessments.²⁵

Finally, the importance of function and cognition is seen across settings and time. Factors that predicted adverse outcomes in hospitalised older adults also predicted adverse outcomes in community-dwelling older adults. Attention to these factors across care settings has also been central to the practice of geriatrics. Indeed, one of the authors of the original scale (Exton-Smith) broadened the scope of his inpatient practice to conduct home visits.^{26,27}

We do not expect that this exact tool will be widely used. It does not have the predictive validity to identify with complete accuracy those who will die. Indeed, it is unlikely that such a tool will ever be developed, or is even desirable.

The risk score also has a ceiling effect, with the majority of the participants scoring high on the test, so the scale may not identify variation in risk among those who have no disability. However, those with no cognitive or functional issues have a very low risk of dying or institutionalisation, and identifying variation in these individuals may not be as clinically important as identifying those at high risk.

While we do not advocate using this particular scale, we do feel that considering cognition and function are very important in health and social care. For clinicians, cognition and functional status should help to guide treatment decisions. Aggressive interventions for older individuals with severe functional and/or cognitive impairment may not go well. Conversely, those with few cognitive or functional deficits may benefit from aggressive interventions.

Considering only disease stage and severity may miss important prognostic information. For public health planners, life expectancy is an important consideration in deciding on screening interventions, and early detection may be of differential benefit in those with and without functional impairment. Considering functional status and cognitive status would also help to guide

needs assessments by more accurately identifying populations at higher risk.

For healthcare administrators, the results are also important. Case-mix models that adjust only for disease-related factors will miss critically important information that could affect policy decisions, pay-for-performance models and quality assessment. For all these reasons, considering the powerful and cumulative effect of functional status and cognitive status is central to the care of older persons in hospital and in the community. This was the case in 1962, as it is today.

Funding acknowledgement: The MSHA was funded primarily by Manitoba Health, with additional funding through the CSHA by the Seniors Independence Research Program of the National Health Research and Development Program of Canada (Project No. 6606-3954-MC[S]). MSHA-2 was funded primarily by Manitoba Health's Health Communities Development Fund with additional funding provided through the CSHA by the Seniors Independence Research Program of the National Health Research and Development Program of Health Canada (Project No. 6606-3954-MC[S]).

REFERENCES

- Cruz M, Covinsky K, Widera EW et al. Predicting 10-year mortality for older adults. *JAMA* 2013; 309:874–6. <http://dx.doi.org/10.1001/jama.2013.1184>
- Lee SJ, Lindquist K, Segal MR et al. Development and validation of a prognostic index for 4-year mortality in older adults. *JAMA* 2006; 295:801–8. Erratum in *JAMA* 2006; 295:1900.
- Carey EC, Covinsky KE, Lui LY et al. Prediction of mortality in community-living frail elderly people with long-term care needs. *J Am Geriatr Soc* 2008; 56:68–75.
- Walter LC, Brand RJ, Counsell SR et al. Development and validation of a prognostic index for 1-year mortality in older adults after hospitalization. *JAMA* 2001; 285:2987–94.
- Mitchell SL, Miller SC, Teno JM et al. The advanced dementia prognostic tool: a risk score to estimate survival in nursing home residents with advanced dementia. *J Pain Symptom Manage* 2010; 40:639–51. <http://dx.doi.org/10.1016/j.jpainsymman.2010.02.014>
- Alba AC, Agoritsas T, Jankowski M et al. Risk prediction models for mortality in ambulatory heart failure patients: a systematic review. *Circ Heart Fail* 2013; 6:881–9. <http://dx.doi.org/10.1161/CIRCHEARTFAILURE.112.000043>
- Arnold J, Exton-Smith AN. The geriatric department and the community value of hospital treatment in the elderly. *Lancet* 1962; 280:551–3.
- Sudman S, Bradburn NM. *Asking questions: a practical guide to questionnaire design*. San Francisco: Jossey Bass; 1982. p.492.
- Yourman LC, Lee SJ, Schonberg MA et al. Prognostic indices for older adults: a systematic review. *JAMA* 2012; 307:182–92. <http://dx.doi.org/10.1001/jama.2011.1966>
- Statistics Canada. *Profile of census metropolitan areas and census agglomerations, 1991 census*. Ottawa: Ministry of Industry, Science and Technology; 1992.
- Canadian Study of Health and Aging: study methods and prevalence of dementia. *CMAJ* 1994; 150:899–913.
- DeSalvo KB, Blosner N, Reynolds K et al. Mortality prediction with a single general self-rated health question. A meta-analysis. *J Gen Intern Med* 2006; 21:267–75.
- Mossey JM, Shapiro E. Self-rated health: a predictor of mortality among the elderly. *Am J Public Health* 1982; 72:800–8.
- Andrews FM, Withey FR. *Social indicators of well-being*. New York: Plenum Press; 1976.
- Rockwood K, Stadnyk K, MacKnight C et al. A brief clinical instrument to classify frailty in elderly people. *Lancet* 1999; 35:205–6.
- Mitnitski AB, Mogilner AJ, Rockwood K. Accumulation of deficits as a proxy measure of aging. *ScientificWorldJournal* 2001; 1:323–36.
- 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; 59:M627–32.
- Teng EL, Chui HC. The Modified Mini-Mental State (3MS) examination. *J Clin Psychiatry* 1987; 48:314–8.
- 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; 12:189–98.
- Isaacs B, Walkey FA. The assessment of the mental state of elderly hospital patients using a simple questionnaire. *Am J Psychiatry* 1963; 120:173–4.
- Isaacs B, Walkey FA. A simplified paired-associate test for elderly hospital patients. *Br J Psychiatry* 1964; 110:80–3.
- Fillenbaum GG. *Multidimensional functional assessment of older adults: the Duke Older Americans Resources and Services Procedures*. Hillsdale, New Jersey: Lawrence Erlbaum Associates; 1988.
- Campbell SE, Seymour DG, Primrose WR. A systematic literature review of factors affecting outcome in older medical patients admitted to hospital. *Age Ageing* 2004; 33:110–5.
- Lee SJ, Go AS, Lindquist K et al. Chronic conditions and mortality among the oldest old. *Am J Public Health* 2008; 98:1209–14. <http://dx.doi.org/10.2105/AJPH.2007.130955>
- Blazer DG. How do you feel about...? Health outcomes in late life and self-perceptions of health and well-being. *Gerontologist* 2008; 48:415–22.
- Exton-Smith AN. Investigation of the aged sick in their homes. *Br Med J* 1952; 2:182–6.
- Denham MJ. Professor Arthur Norman Exton-Smith CBE MA MD FRCP (1920–90): distinguished Geriatrician and Gerontologist. *J Med Biogr* 2009; 17:8–13. <http://dx.doi.org/10.1258/jmb.2008.008003>