Lund University

Predictive Validity and Cut-Off Scores in Four Diagnostic Tests for Falls – A Study in Frail Older People at Home

Olsson Möller, Ulrika; Kristensson, Jimmie; Midlöv, Patrik; Ekdahl, Charlotte; Jakobsson, Ulf

Published in:
Physical & Occupational Therapy in Geriatrics

DOI:
10.3109/02703181.2012.694586

2012

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Physical & Occupational Therapy in Geriatrics, 30(3):189–201, 2012

Available online at http://informahealthcare.com/potg
DOI: 10.3109/02703181.2012.694586

informa
healthcare

Predictive Validity and Cut-Off Scores in Four Diagnostic Tests for Falls – A Study in Frail Older People at Home

U. Olsson Möller, PT, MSc, PhD Student[1,2], J. Kristensson, RN, PhD, Senior Lecturer[3], P. Midlöv, MD, PhD, Associate Professor[1,4], C. Ekdahl, PT, PhD, Professor[3], & U. Jakobsson, RN, PhD, Associate Professor[1,4]

[1]Center for Primary Health Care Research, Lund University, Malmö, Sweden, [2]Vårdalinstitutet, The Swedish Institute for Health Sciences, Lund University, Lund, Sweden, [3]Department of Health Sciences, Lund University, Lund, Sweden, [4]Department of Clinical Sciences in Malmö, Lund University, Malmö, Sweden

ABSTRACT. No study has investigated the predictive validity and cut-off scores in diagnostic tests for falls used in in-home assessment in frail older people. The objective was to investigate the predictive validity for falls in the Downton Fall Risk Index (DFRI), Timed Up and Go (TUG) and Romberg test (RT) used in in-home assessment of frail older people (65+ years). Data on the diagnostic tests were collected at baseline ($N = 153$) and fall frequency were collected at six- and twelve-month follow-ups. The optimal cut-offs were 3 p in DFRI and 12 s in TUG. However, the validity indexes were generally low and only 40–50% were correctly classified. The RT showed low sensitivity. To increase the predictive validity for falls in this context, the use of DFRI and/or TUG as a part of a comprehensive fall-risk assessment tool, should be investigated in future studies.

KEYWORDS. Falls, predictive validity, cut-off score, aged, aged 80+, frail

INTRODUCTION

The predictive validity of diagnostic tests for falls in older people has been investigated using comprehensive fall-risk assessment tools and functional mobility assessment tools in a variety of clinical settings (Perell et al., 2001; Persad, Cook & Giordani, 2010; Scott, Votova, Scanlan & Close, 2007). However, few studies have investigated the predictive validity of fall-risk assessment tools used as a part of in-home assessments. This setting places special demands on functional mobility assessment tools as lack of space and equipment precludes many of them being used in this setting. In Sweden, in the 21st century there has been a substantial

Address Correspondence to: U. Olsson Möller, PT, MSc, PhD Student, Department of Health Sciences, Lund University, P.O. Box 157, SE-221 00, Lund, Sweden (E-mail: ulrika.olsson_moller@med.lu.se).

(Received 15 February 2012; accepted 14 May 2012)

reduction in the number of beds in special accommodations, instead frail older people are to a greater extent receiving healthcare at home (National Board of Health and Welfare, 2011), thus investigating the predictive validity of diagnostic tests for falls in this context is important.

There has been a substantial amount of research over the last two decades into the prediction of falls in older people because of the severity of associated outcomes for the individual after a fall, such as fractures and disability as well as high healthcare costs for society (World Health Organization [WHO], 2007). The causes of falls in older people are often multifactorial and over 400 risk factors for falls has been described (National Institute for Clinical Excellence [NICE], 2004), many of which are not immediately identifiable. These risk factors occur in different areas such as impaired functional and balance ability, sensory deficits, and drugs. Risk-assessment tools are needed for healthcare providers to be able to identify future fallers and introduce implementation of effective interventions to prevent falls. The direct cost of fall injuries in Sweden during 2006 in people aged 65 years and older was SEK[1] 4.8 billion (530 million Euro) and the majority of the costs concern injuries in older people sustained from falls in their homes (National Board of Health and Welfare, 2011). The risk of falling increases with increased level of frailty (WHO, 2007) and older people receiving care at home consume more hospital care than people in special accommodation (Condelius, Edberg, Hallberg, & Jakobsson, 2010; Karlsson, Edberg, Westergren & Hallberg, 2008). Being frail implies the presence of many risk factors for falls, such as dizziness, muscle weakness, and impaired vision, and these factors often interact. It is, therefore, important to be able to identify and intervene early in frail older people living at home by eliminating as many risk factors as possible, thus minimizing the risk of falls. Interventions in frail older people should be multifactorial and include exercise (Gillespie et al., 2009).

A diagnostic test distinguishes between those at risk and those not at risk, and is assessed according to its ability to diagnose the outcome correctly (Bewick, Cheek & Ball, 2004). Investigating an instrument's optimal cut-off score helps to guide correct use of the instrument, e.g., to identify people at risk of a future fall. This knowledge is useful when targeting people most likely to benefit from preventive interventions. In Senior Alert (Senior Alert, www.lj.se/senioralert), a national quality register in Sweden aimed at developing systematic, preventive healthcare for older people, the Downton Fall Risk Index (DFRI) (Downton, 1993) is used to identify people at risk of future falls. DFRI is a questionnaire that includes risk factors for falls, such as medications and impairment, and was developed to establish fall risk in older people, with $\geq$3 as the cut-off score (range 0–11, higher score indicating a higher fall risk) (Downton, 1993), and is usable for different professions in the care for older people. This cut-off score used in institutionalized older people has shown 81–100% sensitivity and 9–40% specificity (Downton, 1993; Nyberg & Gustafson, 1996; Rosendahl et al., 2003). However, to our knowledge, the predictive validity of DFRI has not been investigated in frail older people living at home.

When a person has been identified as being at risk of falling, e.g., after assessment using the DFRI, there should be a more comprehensive evaluation to identify specific risk factors (NICE, 2004) including an assessment of physical performance.

[1] 1 EURO = 9.04 SEK 2012–12-16.

Physical assessment instruments used in the healthcare of older people, usually by physical therapists, are, for example, Timed Up and Go (TUG) (Podsiadlo & Richardson, 1991) and the Romberg test (RT). These tests can be used in in-home assessments (McMichael, Vander Bilt, Lavery, Rodriquez & Ganguli, 2008). TUG is a functional mobility assessment tool and measures the time it takes to rise from a chair, walk a distance of 3 meters, turn, walk back, and sit down again (Podsiadlo, 1991). In a systematic review of TUG and risk of falls in older adults (aged 65+ years and without neurological disease) the cut-off values varied from 10 to 32.6 s (Beauchet et al., 2011). Sensitivity and specificity varied from 63–87% (Shumway-Cook, Brauer & Woollacott, 2000; Trueblood, Hodson-Chennault, McCubbin & Youngclarke, 2001). In a study of 455 community-dwelling frail older people with a 12-month follow-up, and the cut-off set at ≥16 s, 53% sensitivity and 63% specificity were registered (Shimada et al., 2009). A recent systematic review of TUG (Rydwik, Bergland, Forsén & Frändin, 2011) stated that, because of the many different cut-offs and inconclusive results when investigating the predictive validity more research is needed, especially in frail older people.

The Romberg test (RT) is a neurological test for assessing pathology in the proprioceptive pathway (Khasnis & Gokula, 2003) and is used to assess static balance performance. Its predictive validity for falls was investigated in an in-home assessment with 358 older people and showed 24% sensitivity and 89% specificity (McMichael et al., 2008). In the present study a modified version of RT, i.e., the Semitandem Romberg test (SRT) and Tandem Romberg test (TRT) were used (Guralnik et al., 1994). In a study in 58 older people (aged 65+ years) with dizziness ≥16 s on TRT almost doubled the risk of falls (Hansson, Månsson, Ringsberg & Håkansson, 2008). To our knowledge, no study has yet investigated the predictive validity of TUG, RT, SRT, and TRT in in-home assessments of frail older people.

It may be possible to use the instruments mentioned above as a part of a broad screening procedure (DFRI) or as a part of comprehensive in-home assessments (TUG, RT, SRT, TRT) in frail older people at home. However, since their predictive validity for falls differs and various cut-off levels are described, investigation of the predictive validity in these instruments when used in frail older people at home is needed.

AIM

The objectives of the study was to investigate the predictive validity for falls and to identify optimal cut-off scores in DFRI, TUG, RT, SRT, and TRT used in in-home assessments in frail older people (aged 65+ years).

METHOD

Sample

Data were extracted from a study involving 153 people aged 65 years and older (Kristensson, Ekwall, Jakobsson, Midlöv, & Hallberg, 2010). They were participating in a randomized controlled trial (RCT) aimed at evaluating the effects of a healthcare model using case managers for frail community-dwelling older people.

The RCT was a collaboration between municipal health and social care, primary care, and a university hospital, and was conducted in a medium-sized municipality in southern Sweden (approximately 30,000 inhabitants) living in both urban and rural areas. The inclusion criteria were: 65 years or older; living in the municipality where the study was conducted; needing help with at least two activities of daily living (ADL); admitted to hospital at least twice or with at least four contacts with outpatient or primary healthcare during the previous 12 months. The participants had to be able to communicate verbally and have no cognitive impairments, i.e., ≥25 in Mini Mental State Examination (MMSE) (Folstein, Folstein & McHugh, 1975).

Ethics

The Regional Ethics Review Board in Lund approved the study (LU 342–06) and informed written consent was obtained from all respondents.

Measurements

The data collection started in 2006 and was expanded in 2008 with, e.g., TUG and Romberg tests. Thus 150 participants responded to DFRI but only 85 participants performed TUG, RT, SRT, and TRT. The data were collected by means of personal interviews and examinations in in-home assessments. To analyse the predictive ability of the instruments, baseline data concerning DFRI, TUG, RT, SRT, and TRT were used and data on falls were collected at the six- and twelve-month follow-ups.

Falls

Data on falls were collected by asking "Have you had a fall in the last 3 months?". When there was a positive response, follow-up questions about frequency, location, fall injuries, and medical treatment were asked. A fall was defined as "an unexpected event in which the participant comes to rest on the ground, floor, or lower level" (Lamb, Jørstad-Stein, Hauer & Becker, 2005).

Downton Fall Risk Index (DFRI)

DFRI includes 11 fall risk items, i.e., history of falls in the preceding 12 months; use of tranquilizers or sedatives; diuretics; antihypertensive drugs; antiparkinsonian drugs; antidepressants; visual impairment; hearing impairment; limb impairment; cognitive impairment, and walking ability. A total score of 0–11 is achieved and a cut-off score ≥3 indicates a high risk of falls (Downton, 1993). In this study, the questions in DFRI were assessed as follows: the history of falls was reported by the participant; medication use by the participants was determined from drug prescriptions; visual impairment was classified as moderate to severe impairment or blindness, and needing glasses on a daily basis; hearing impairment was scored as moderate to severe impairment or deafness, or needing hearing aid; limb impairment was scored if there were signs of extremity paresis or muscle weakness; cognitive impairment was determined according to an MMSE score ≤23 and walking ability classed as walking safely or unsafely with or without a walking aid.

Timed Up and Go (TUG)

TUG assesses functional mobility, i.e., it reflects the balance and gait manoeuvres used in everyday life and is measured as the time in seconds it takes to rise from a chair, walk a distance of 3 meters, turn, walk back, and sit down again (Podsiadlo & Richardson, 1991). The test has shown strong inter- and intra-rater reliability (Podsiadlo & Richardson, 1991, Shumway-Cook et al., 2000) and a moderate association with ADL, balance, and mobility (Rydwik et al., 2011). In this study, the participants were able to practice the TUG test once before execution and were reassured about receiving support from the researcher if they lost their balance. In this study, a 44–46-cm-high arm chair was used and the participants were verbally instructed to rise, walk across a line on the floor 3 meters away, turn, return, and sit down again, as fast but safely as possible. The participant was given a practical trial before being timed and the timing started when their back left the backrest of the chair and stopped when they were seated again. The participants were able to use any normally used walking aid and efforts were made to ensure that, if used, did not affect the time required, e.g., there was enough space to turn around with a walker.

Romberg Test (RT), Semitandem Romberg Test (SRT) and Tandem Romberg Test (TRT)

Romberg test (RT) was used in the beginning to detect neurosyphilis, but today it is used as a part of a routine neurological examination to assess pathology in the proprioceptive pathway (Khasnis & Gokula, 2003), i.e., static balance impairment. The RT and TRT have shown moderate to good test-retest reliability (Steffen & Seney, 2008). Tests of static balance in this study included RT, SRT, and TRT (Guralnik et al., 1994). In this study, the participants were instructed to stand with their feet together (RT), semitandem (SRT) or tandem (TRT) with arms crossed and eyes closed. The time (in seconds) was measured (maximum 30 or 60 s) from the closing of the eyes until test failure, i.e., they opened their eyes, they made compensatory movements, i.e., moved arms or feet to achieve stability, or they lost their balance/started to fall. The best of three trial times was used and the participants were allowed a pretest practice run.

Statistical Analyses

Descriptive statistics was used to describe the sample at baseline. The Pearson chi-square test, Fischer's exact test, the Mann–Whitney U-test and Student's t-test were used for group comparisons between fallers and nonfallers. A p-value < 0.05 was considered statistically significant. There were no difference in fall rates between the intervention ($n = 80$) and the control group ($n = 73$) at the six-month ($p = 0.642$) as well as twelve-month ($p = 0.176$) follow-ups and, therefore, both groups were included in this study. The sensitivity, specificity, positive likelihood ratio (LR+), and negative likelihood ratio (LR–), Youden's index, and the percent of correctly classified cases were calculated for different cut-off values on each scale (i.e., in seconds or points). Sensitivity refers to the number of fallers correctly identified by the test (i.e., true positives) and specificity the number of nonfallers correctly identified by the test (i.e., true negatives) (Bewick et al., 2004). LR+ refers to the increase in the odds of having a positive test result if the outcome is positive (i.e., having a fall) and LR– to an increase in the odds of having a negative result if the outcome is positive (Bewick et al., 2004). LR+ > 10 constitutes strong evidence that the test can identify fallers and LR– <0.1 that the test can rule out a high risk of falling (Moe-Nilssen, Nordin & Lundin-Olsson, 2008). To determine the appropriate threshold, Youden's index was used, range –1 to +1, J = +1 indicates a perfect test (Bewick et al., 2004). This test explores the cut-off value for which sensitivity and specificity are maximized. Statistical analyses were conducted using the software program SPSS 18.0 (International Business Machines Corp, Armonk, NY, USA).

RESULTS

In the total sample at baseline ($N = 153$) 67% were women and the mean age was 81.5 years (SD 6.3, range 66–94 years). As shown in Table 1, at least one fall was reported by 25.0% ($n = 38$) of whom 68.4% were injured and 31.6% needed medical care as a result. Most of the falls, 70.3%, occurred inside and 62.7% reported a fear of falling (Table 1). No instrument showed a statistically significant difference between fallers and nonfallers at the six- or twelve-month follow-up (Table 2).

According to DFRI in this sample of frail older people, 78% at baseline had an increased risk of falling, i.e., ≥3 (Table 2). At the six-month follow-up (Table 3), a cut-off of ≥2 was suggested by Youden's index with sensitivity of 100% but specificity of only 8% resulting in only 34% being correctly classified. Youden's index suggested a cut-off value of ≥3 at the twelve-month follow-up with 79% sensitivity and 24% specificity giving 39% correctly classified (Table 3).

The mean time for TUG was 15.8 s with a range from 5.6 to 36.4 s (Table 2). Youden's index suggested a cut-off of ≥12–13 which resulted in 67% sensitivity and 50% specificity at the six-month follow-up and 78% sensitivity and 37% specificity

TABLE 1. Baseline Characteristics of the Participants ($N = 153$)

Age, mean (SD)	81.5 (6.3)
Men/women, (%)	33.3/66.7
Falls in the past 3 months, *n* (%)	38 (25.0)
1 fall[1], *n* (%)	20 (55.6)
2 falls, *n* (%)	10 (27.8)
≥3 falls, *n* (%)	6 (16.6)
Inside[2], *n* (%)	26 (70.3)
Outside, *n* (%)	9 (24.3)
Both inside and outside, *n* (%)	2 (5.4)
Fall related injury, *n* (%)	26 (68.4)
Medical attendance due to fall related injury, *n* (%)	12 (31.6)
Fear of falling[3,4], *n* (%)	94 (62.7)
ADL-staircase[5] score, median (q1–q3)	2 (1–3)
Walking aid, *n* (%)	121 (80.7)

[1]Number of falls, missing data $n = 2$.
[2]Location of falls, missing data $n = 1$.
[3]Single item question: "Are you afraid of falling?" (yes/no).
[4]Missing data $n = 3$.
[5]Sonn & Åsberg, 1991.

TABLE 2. Baseline Characteristics of the Diagnostic Tests and Differences in the Diagnostic Tests Between Falls and no Falls at the Six- and Twelve-Month Follow-Ups

Parameters	Baseline $N = 153$ All	Six-Month Follow-Up ($n = 120$) Falls	Non Falls	P	Twelve-Month Follow-Up ($n = 106$) Falls	Non Falls	P
Timed Up and Go (sec)	$n = 85$	$n = 15$	$n = 54$		$n = 18$	$n = 43$	
Range	5.6–36.4	10.0–36.4	5.6–29.7		7.0–31.5	5.6–36.4	
Mean (SD)	15.8 (6.2)	16.9 (7.7)	15.0 (5.6)	0.297[a]	15.9 (5.9)	15.1 (6.1)	0.654[a]
Romberg test (sec)	$n = 85$	$n = 15$	$n = 54$		$n = 18$	$n = 43$	
<15 sec, n (%)	12 (14)	2 (13)	7 (13)	1.000[b]	4 (22)	4 (9)	0.220[b]
<30 sec, n (%)	18 (21)	3 (20)	10 (18)	1.000[b]	4 (22)	8 (19)	0.736[b]
Semitandem Romberg test (sec)	$n = 85$	$n = 15$	$n = 54$		$n = 18$	$n = 43$	
<15 sec, n (%)	25 (29)	3 (20)	15 (28)	0.743[b]	4 (22)	12 (28)	0.757[b]
<30 sec, n (%)	29 (34)	3 (20)	17 (32)	0.526[b]	6 (33)	12 (28)	0.672[c]
Tandem Romberg test (sec)	$n = 85$	$n = 15$	$n = 54$		$n = 18$	$n = 43$	
<15 sec, n (%)	78 (92)	14 (93)	49 (91)	1.000[b]	17 (94)	38 (88)	0.660[b]
<30 sec, n (%)	81 (95)	14 (93)	52 (96)	0.527[b]	17 (94)	41 (95)	1.000[b]
Downton Fall Risk Index	$n = 150$	$n = 31$	$n = 85$		$n = 29$	$n = 75$	
Median (q1–q3)	4 (3–5)	4 (3–5)	3 (3–4.5)	0.828[d]	3 (3–4)	4 (3–5)	0.297[d]
≥3, n (%)	114 (78)	24 (77)	66 (78)	0.979[c]	23 (79)	57 (76)	0.719[c]

[a]Student's T-test.
[b]Fisher's Exact Test.
[c]Pearson Chi-Square.
[d]Mann-Whitney U Test.

at the twelve-month follow-up. Using this cut-off resulted in 50–54% being correctly classified (Table 4).

In this sample of frail older people, 86% and 79% were able to perform the RT for 15 and 30 s, respectively, and about two-thirds completed 30 s in SRT

TABLE 3. Sensitivity, Specificity, LR +, LR–, Youden's Index and Percentage Correctly Classified in Downton Fall Risk Index (DFRI) Predicting Falls (Score 0–11)

DFRI	Sensitivity	Specificity	LR+	LR–	Youden's Index	Correctly Classified (%)
			Six-month follow-up			
≥1	1.000	0.012	1.021	–	0.012	27.6
≥2	1.000	0.082	1.089	–	0.082	33.8
≥3	0.774	0.224	0.997	1.009	–0.002	37.1
≥4	0.516	0.518	1.070	0.934	0.034	51.7
≥5	0.290	0.753	1.174	0.943	0.043	62.9
			Twelve-month follow-up			
≥1	1.000	0.013	1.014	–	0.013	28.8
≥2	0.965	0.053	1.019	0.660	0.018	30.8
≥3	0.793	0.240	1.043	0.862	0.033	39.4
≥4	0.414	0.467	0.777	1.255	–0.119	45.2
≥5	0.138	0.707	0.471	1.219	–0.155	54.8

TABLE 4. Sensitivity, Specificity, LR+, LR–, Youden's Index and Percentage Correctly Classified in Timed Up and Go (TUG) Predicting Falls (in seconds)

TUG	Sensitivity	Specificity	LR+	LR–	Youden's Index	Correctly Classified (%)
			Six-month follow-up			
≥10	1.000	0.130	1.149	–	0.130	31.9
≥11	0.933	0.222	1.999	0.302	0.155	37.7
≥12	0.733	0.315	1.070	0.848	0.048	40.6
≥13	0.667	0.500	1.334	0.666	0.167	53.6
≥14	0.467	0.537	1.009	0.992	0.004	52.2
≥15	0.400	0.556	0.901	1.079	–0.044	52.2
≥16	0.400	0.611	1.028	0.982	0.011	56.5
			Twelve-month follow-up			
≥10	0.889	0.116	1.006	0.957	0.005	35.5
≥11	0.833	0.209	1.053	0.799	0.042	40.3
≥12	0.778	0.372	1.239	0.597	0.150	50.0
≥13	0.611	0.465	1.142	0.836	0.076	50.0
≥14	0.556	0.558	1.258	0.796	0.114	56.4
≥15	0.500	0.558	1.131	0.896	0.058	54.1
≥16	0.500	0.628	1.344	0.796	0.128	59.0

(Table 2). Only 8% were able to perform the TRT for more than 15 s (Table 2). Different cut-offs at 0, 15, 30, and 60 s were suggested by Youden's index (Table 5). The sensitivity for RT and SRT were low, between 0% and 39%. TRT showed an acceptable sensitivity of 67–94% but using the suggested cut-offs at the six- and twelve-month follow-ups, 0 and ≥15 s, respectively, only resulted in approximately 40% being correctly classified (Table 5).

DISCUSSION

The objective of this study was to investigate the predictive validity for falls in diagnostic tests used in in-home assessment of frail older people. No instrument used in this study showed a strong predictive validity. A cut-off of ≥12 s for TUG had the highest predictive validity according to the validity indexes, i.e., LR+, LR–, and Youden's index, with the same sensitivity as DFRI at the twelve-month follow-up, i.e., 80%. However, specificity, LR+, LR–, Youden's index, and the number correctly classified were generally low in all tests. The low validity of TUG is not so surprising since it measures only one aspect, i.e., functional mobility, of the multifactorial spectra of fall risks in older people. DFRI, however, assesses different risk factors for falls, which should increase the predictive validity. That it did not do so may indicate that not all aspects of fall risks are covered in DFRI. A cut-off of 3 was suggested in DFRI as the optimal at the twelve-month follow-up in this study (Table 3), and 79% sensitivity and 24% specificity at this cut-off were in congruence with other studies in patients in a geriatric stroke rehabilitation unit (Nyberg & Gustafson, 1996) and older people living in special accommodation (Rosendahl et al., 2003).

A study on community-dwelling frail older people suggested 16 s as a cut-off for the TUG with 53% sensitivity and 63% specificity (Shimada et al., 2009), i.e., 4 s more than in this study. The reason might be that people with cognitive

TABLE 5. Sensitivity, Specificity, LR+, LR−, Youden's Index and Percentage Correctly Classified in Romberg Test (RT), Semitandem Romberg Test (SRT) and Tandem Romberg Test (TRT) Predicting Falls (in seconds)

RT	Sensitivity	Specificity	LR+	LR−	Youden's Index	Correctly Classified (%)
			Six-month follow-up			
			RT			
0	–	0.926	–	–	–0.074	72.5
≥15	0.133	0.870	1.023	0.996	0.003	71.0
≥30	0.200	0.815	1.081	0.982	0.015	68.1
			SRT			
0	0.133	0.759	0.552	1.142	–0.108	62.3
≥15	0.200	0.722	0.719	1.108	–0.078	60.9
≥30	0.200	0.685	0.635	1.168	–0.115	57.9
≥60	0.200	0.593	0.491	1.349	–0.207	50.7
			TRT			
0	0.733	0.315	1.070	0.848	0.048	40.6
≥15	0.933	0.092	1.028	0.728	0.025	27.5
≥30	0.933	0.037	0.969	1.811	–0.030	23.2
≥60	0.933	–	–	–	–0.067	20.3
			Twelve-month follow-up			
			RT			
0	0.056	0.953	1.191	0.990	0.009	68.8
≥15	0.222	0.907	2.387	0.858	0.129	70.5
≥30	0.222	0.814	1.194	0.956	0.036	63.9
			SRT			
0	0.222	0.767	0.953	1.014	–0.011	60.6
≥15	0.222	0.721	0.796	1.079	–0.057	57.4
≥30	0.333	0.704	1.125	0.947	0.037	60.6
≥60	0.389	0.674	1.193	0.906	0.063	59.0
			TRT			
0	0.667	0.349	1.024	0.954	0.016	44.3
≥15	0.944	0.116	1.068	0.483	0.060	36.1
≥30	0.944	0.046	0.990	1.217	–0.010	31.1
≥60	0.944	–	–	–	–0.056	27.9

impairment were included in the study by Shimada et al. (2009) and their mean score on the TUG was higher, i.e., they were more frail. The cut-off in this study agrees with a study on 198 community-dwelling older people (mean age 78.1 SD 8.2) without cognitive deficits (Trueblood et al., 2001). This may indicate that a different cut-off should be used in people with cognitive impairment. However, more research is needed to evaluate appropriate cut-off values for different groups of older people. A recent systematic review (Rydwik et al., 2011) states that more research is needed, especially in frail older people, to establish the TUG predictive validity for falls.

The sensitivity of the RT and SRT tests were low. In this study, a cut-off of 15 s in RT resulted in 22% sensitivity and 91% specificity which accords with another study using the Romberg test in in-home assessment (McMichael et al., 2008). In the TRT, a cut-off of ≥15 gave 94% sensitivity and 12% specificity. A statistically significant association between fall risk and TRT in older people has been described (Muir, Berg, Chesworth, Klar & Speechley, 2010), but the cut-off with the highest predictive validity for falls has, to our knowledge, not been investigated. However, since only a few percent were able to perform TRT in this study, neither RT, SRT, nor TRT can, according to the results of this study, be recommended for used in in-home assessments to predict falls in frail older people.

When using the suggested cut-offs in this study, i.e., ≥12 s for TUG and ≥3 for DFRI, 72–78% of the participants were at risk of future falls. This is in agreement with another study in older people living in special accommodation where 73% scored ≥3 on DFRI (Rosendahl et al., 2003). This means that frail older people at home have the same need for fall preventive interventions as older people in special accommodations. The health status of frail older people is complex and changes rapidly and the screening of older people in special accommodation for fall risk every third month has been recommended (Rosendahl et al., 2003). Since frail older people more often live at home today, this screening might also be recommended for people in this context. The high number of people at risk in this context and the complexity of fall risk in frail older people imply the need for implementation of an extensive preventive approach in the care of older people at home. The effectiveness of such an intervention should be evaluated in future studies.

The guideline for prevention of falls in older people (Panel on Prevention of Falls in Older Persons, American Geriatrics Society & British Geriatric Society, 2011) recommend that all individuals who report difficulties with gait or balance should be offered a multifactorial fall-risk assessment. With more frail older people living at home, assessing fall risk in an in-home context should be both a part of the evaluation on the older person's first contact with community healthcare and of continuous re-assessments. Although an acceptable sensitivity was shown in this study for DFRI and TUG, low specificity and only 40–50% correctly classified cases indicate that a more comprehensive assessment instrument covering different aspects of risk factors for falls may be needed. Risk models for falls have been suggested in previous studies (Pluijm et al., 2006; Stalenhoef, Diederiks, Knottnerus, Kester, & Crebolder, 2002) including, e.g., fall history, dizziness, and grip strength and the predictive validity with DFRI and/or TUG as a part of a comprehensive risk model in in-home assessment of frail older people should be investigated in future studies.

Methodological Considerations

The low number of fallers at the six- ($n = 15$) and twelve- ($n = 18$) month follow-ups probably meant low power in the analysis resulting in lack of statistically significant differences in TUG, RT, SRT, and TRT. However, the main aim for this study was to investigate the predictive validity and cut-off scores for four different diagnostic tests for falls. Performing the TUG in an in-home assessment may jeopardize the reliability of the test, since standardization of the test can be difficult to achieve. However, a normal chair height is about 45 cm and frail older people usually have a chair with armrests since the ability to rise without using arms is often limited. In this study, no participants were unable to perform the test due to lack of space. As falls were self-reported, there was a risk of recall bias. However, the participants in this study were asked the same questions in interviews every third month and were, therefore, well aware that they would need to report eventual falls. The risk of recall bias is higher in DFRI questions where the patients were asked about falls in the past year. Recall could affect the total test score and thereby lead to failure to identify people at risk for falling. According to the inclusion criteria used in this study, the meaning of the term frail is open to question. There is no agreed definition of frailty but it is characterized by, e.g., reduced mobility and an increased risk of falling (Fried et al., 2001). The high mean age and the large proportion of people at risk of falls indicate that the majority of the sample may be described as frail. The large proportion of people at risk of falling at baseline has probably affected the results in this study, i.e., low validity indexes and the sample indicate that the results cannot be generalized to healthier older people (aged 65+ years) living at home but probably to older people at home with home-help care.

CONCLUSION

The predictive validity of the diagnostic tests used in this study with frail older people at home was generally low. It is suggested that a cut-off of ≥ 3 on the DFRI and ≥ 12 s on the TUG predicts falls in frail older people (aged 65+ years) living at home with an acceptable sensitivity of 80%. The RT and SRT showed low sensitivity and the TRT could only be performed by a few percent in this sample. A comprehensive assessment tool usable in in-home assessment in frail older people may show a better predictive validity for falls and should be investigated in future studies. When using the suggested cut-offs in this study, almost 80% of the participants had an increased risk of falling, which implies a need for fall-prevention interventions in frail older people living at home.

ACKNOWLEDGEMENTS

This project was a collaboration between the Faculty of Medicine at Lund University, Vårdalinstitutet, Lund University Hospital, the Primary Care and the Municipality of Eslöv, Sweden. The authors are most grateful to the Faculty of Medicine at Lund University, the Greta and Johan Kock Foundation, Vårdalinstitutet, the Governmental Funding of Clinical Research within the NHS (ALF), and the Swedish Research Council (grant no:521–2007-2940) for funding this study. They are also most grateful to the participants. The authors especially want to thank PhD students Elin Taube and Magnus Sandberg and project assistants Anneli Wallgren and Emil Ekblad for data collection, and Pat Shrimpton for revising the language.

Declaration of interest. The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

ABOUT THE AUTHORS

Ulrika Olsson Möller, RPT, is a PhD student at the Faculty of Medicine, Lund University, Sweden, and focus her thesis on falls and dizziness in older people. Jimmie Kristensson, RN, has a PhD in nursing and works as Senior Lecturer at the Faculty of Medicine, Lund University, Sweden. His research mainly focus on elderly care. Patrik Midlöv, MD, is an Associate Professor and Senior Lecturer at the Center

for Primary Health Care Research, Lund University, Sweden. His research focus on older people in general and with cardiovascular disease, in Primary Care. Charlotte Ekdahl, RPT, is a Professor in Physiotherapy at the Department of Health Sciences, Lund university, Sweden. She has authored more than 100 scientific publications primarily within the rheumatological and musculoskeletal fields. Ulf Jakobsson, RN, PhD, MSc in Statistics, is an Associate Professor at Center for Primary Health Care Research, Lund University, Sweden. His research focus on areas of caring science, with an emphasis on chronic pain, gerontology and elderly care, and research methodology.

REFERENCES

Beauchet, O., Fantino, B., Allali, G., Muir, S.W., Montero-Odasso, M., & Annweiler, C. (2011). Timed Up and Go test and risk of falls in older adults: A systematic review. *The Journal of Nutrition, Health & Aging*, *15*(10), 933–938. DOI: 10.1007/s12603-011-0062-0.

Bewick, V., Cheek, L., & Ball, J. (2004). Statistics review 13: Receiver operating characteristic curves. *Critical Care*, *8*(6), 508–512. DOI: 10.1186/cc3000.

Condelius, A., Edberg, A.K., Hallberg, I.R., & Jakobsson, U. (2010). Utilization of medical healthcare among people receiving long-term care at home or in special accommodation. *Scandinavian Journal of Caring Sciences*, *24*(2), 404–413.

Downton, J.H. (1993). *Falls in the elderly*. London: Hodder and Stoughton Limited.

Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198.

Fried, L.P., Tangen, C.M., Walston, J., Newman, A.B., Hirsch, C., Gottdiener, J., et al. (2001). Frailty in older adults: Evidence for a phenotype. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *56*(3), 146–156.

Gillespie, L.D., Robertson, M.C., Gillespie, W.J., Lamb, S.E., Gates, S., Cumming, R.G., et al. (2009). Interventions for preventing falls in older people living in the community. *Cochrane Database of Systematic Reviews*, *15*(2), CD 007146.

Guralnik, J.M., Simonsick, E.M., Ferrucci, L., Glynn, R.J., Berkman, L.F., Blazer, D.G., et al. (1994). A short physical performance battery assessing lower extremity function: Association with self-reported disability and prediction of mortality and nursing home admission. *Journal of Gerontology*, *49*(2), 85–94.

Hansson, E.E., Månsson, N.O., Ringsberg, K.A., & Håkansson, A. (2008). Falls among dizzy patients in primary healthcare: An intervention study with control group. *International Journal of Rehabilitation Research*, *31*(1), 51–57.

Karlsson, S., Edberg, A.K., Westergren, A., & Hallberg, I.R. (2008). Older people receiving public long-term care in relation to consumption of medical healthcare and informal care. *The Open Geriatric Medicine Journal*, *1*, 1–9.

Khasnis, A., & Gokula, R.M. (2003). Romberg's test. *Journal of Postgraduate Medicine*, *49*(2), 169–172.

Kristensson, J., Ekwall, A.K., Jakobsson, U., Midlöv, P., & Hallberg, I.R. (2010). Case managers for frail older people: A randomised controlled pilot study. *Scandinavian Journal of Caring Sciences*, *24*(4), 755–763. DOI: 10.1111/j.1471-6712.2010.00773.x.

Lamb, S.E., Jørstad-Stein, E.C., Hauer, K., & Becker, C. (2005). Development of a common outcome data set for fall injury prevention trials: The Prevention of Falls Network Europe consensus. *Journal of the American Geriatrics Society*, *53*(9), 1618–1622. DOI: 10.1111/j.1532-5415.2005.53455.x.

McMichael, K.A., Vander Bilt, J., Lavery, L., Rodrigues, E., & Ganguli, M. (2008). Simple balance and mobility tests can assess falls risk when cognition is impaired. *Geriatric Nursing*, *29*(5), 311–323.

Moe-Nilssen, R., Nordin, E., & Lundin-Olsson, L. (2008). Criteria for evaluation of measurement properties of clinical balance measures use in fall prevention studies. (Work Package 3 of European Community Research Network Prevention of Falls Network Europe.) *Journal of Evaluation in Clinical Practice*, *14*(2), 236–240.

Muir, S.W., Berg, K., Chesworth, B., Klar, N., & Speechley, M. (2010). Quantifying the magnitude of risk for balance impairment on falls in community-dwelling older adults: A systematic review and meta-analysis. *Journal of Clinical Epidemiology*, *63*(4), 389–406.

National Board of Health and Welfare. (2011). *Lägesrapport 2011 – Hälso- och sjukvård och socialtjänst. (Progress report 2011 – Healthcare and social service)*. Stockholm: National Board of Health and Welfare.

National Institute for Health and Clinical Excellence (NICE). (2004). *Clinical practice guideline for the assessment and prevention of falls in older people*. London: Royal College of Nursing.

Nyberg, L., & Gustafson, Y. (1996). Using the Downton Index to predict those prone to falls in stroke rehabilitation. *Stroke*, *27*(10), 1821–1824.

Panel on Prevention of Falls in Older Persons, American Geriatrics Society & British Geriatrics Society. (2011). Summary of the updated American Geriatrics Society/British Geriatrics Society clinical practice guideline for prevention of falls in older persons. *Journal of the American Geriatrics Society*, *59*(1), 148–157. DOI: 10.1111/j.1532-5415.2010.03234.x.

Perell, K.L., Nelson, A., Goldman, R.L., Luther, S.L., Prieto-Lewis, N., & Rubenstein, L.Z. (2001). Fall-risk assessment measures: An analytic review. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *56*(12), 761–766.

Persad, C.C., Cook, S., & Giordani, B. (2010). Assessing falls in the elderly: Should we use simple screening tests or a comprehensive fall risk evaluation? *European Journal of Physical and Rehabilitation Medicine*, *46*(2), 249–259.

Pluijm, S.M., Smit, J.H., Tromp, E.A., Stel, V.S., Deeg, D.J., Bouter, L.M., et al. (2006). Fall risk profile for identifying community-dwelling elderly with a high risk of recurrent falling: Results of a 3-year prospective study. *Osteoporosis International*, *17*(3), 417–425.

Podsiadlo, D., & Richardson, S. (1991). The timed "Up & Go": A test of basic functional mobility for frail elderly persons. *Journal of the American Geriatrics Society*, *39*(2), 142–148.

Rosendahl, E., Lundin-Olsson, L., Kallin, K., Jensen, J., Gustafson, Y., & Nyberg, L. (2003). Prediction of falls among older people in residential care facilities by the Downton index. *Aging Clinical and Experimental Research*, *15*(2), 142–147.

Rydwik, E., Bergland, A., Forsén, L., & Frändin, K. (2011). Psychometric properties of Timed Up and Go in elderly people: A systematic review. *Physical & Occupational Therapy in Geriatrics*, *29*(2), 102–125. DOI: 10.3109/02703181.2011.564725.

Scott, V., Votova, K., Scanlan, A., & Close, J. (2007). Multifactorial and functional mobility assessment tools for fall risk among older adults in community, home-support, long-term and acute care settings. *Age and Ageing*, *36*(2), 130–139.

Shimada, H., Suzukawa, M., Tiedemann, A., Kobayashi, K., Yoshida, H., & Suzuki, T. (2009). Which neuromuscular or cognitive test is the optimal screening tool to predict falls in frail community-dwelling older people? *Gerontology*, *55*(5), 532–538.

Shumway-Cook, A., Brauer, S., & Woollacott, M. (2000). Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go test. *Physical Therapy*, *80*(9), 896–903.

Sonn, U, Åsberg, K.H. (1991). Assessment of activities of daily living in the elderly. A study of a population of 76-year-olds in Gothenburg, Sweden. *Scandinavian Journal of Rehabilitation Medicine*, *23*(4), 193–202.

Stalenhoef, P.A., Diederiks, J.P., Knottnerus, J.S., Kester, A.D., & Crebolder, H.F. (2002). A risk model for the prediction of recurrent falls in community-dwelling elderly: A prospective cohort study. *Journal of Clinical Epidemiology*, *55*(11), 1088–1094.

Steffen, T., & Seney, M. (2008). Test-retest reliability and minimal detectable change on balance and ambulation tests, the 36-item short-form health survey, and the unified Parkinson disease rating scale in people with parkinsonism. *Physical Therapy*, *88*(6), 733–746.

Trueblood, P.R., Hodson-Chennault, N., McCubbin, A., & Youngclarke, D. (2001). Performance and impairment-based assessments among community dwelling elderly: Sensitivity and specificity. *Journal of Geriatric Physical Therapy*, *24*(1), 2–6.

World Health Organization (WHO). (2007). *Global report on falls prevention in older age*. Oslo: World Health Organization.