Maturitas 146 (2021) 26–33

Contents lists available at ScienceDirect

Maturitas

journal homepage: www.elsevier.com/locate/maturitas

Physical performance and risk of hip fracture in community-dwelling elderly people in China: A 4-year longitudinal cohort study

Bi-Xia Zhong [a], Hai-Li Zhong [b], Guan-Qun Zhou [a], Wen-Qi Xu [a], Ying Lu [c], Qian Zhao [a,c,*]

[a] School of Public Health, Guangzhou Medical University, GuangZhou, PR China
[b] School of Public Health, Sun Yat-sen University, GuangZhou, PR China
[c] Department of Biomedical Data Science, Stanford University, Stanford, CA, USA

ARTICLE INFO

Keywords:
Hip fracture
Physical performance
Elderly

ABSTRACT

Background: This study aims to evaluate the association between the risk of hip fracture and score on the Short Physical Performance Battery (SPPB) and handgrip strength in community-dwelling elderly people in China.
Methods: A total of 5,958 community-dwelling Chinese people aged 60 years or more from the China Health and Retirement Longitudinal Study (CHARLS) were surveyed in 2011 (baseline) and followed through to 2016. Score on the SPPB (which comprises tests of balance, walking speed, and repeated chair stands) and handgrip strength were determined at baseline. Binary logistic regression models were used to estimate the risk ratio (RR) and 95 % CI.
Result: During an average of approximately 4 years of follow-up, 180 (3.0 %) participants experienced incident hip fracture. After multivariate adjustment, the overall SPPB score and repeated chair stands alone distinguished a gradient of hip fracture risks. The risk of hip fracture was 1.65-fold higher in poor SPPB performers (score 0−6) than in good SPPB performers (score 10−12). Participants unable to complete repeated chair stands, and those who took ≥16.7 s or 13.7–16.6 s to complete them, had a higher risk than those who took ≤ 11.1 s to complete them, with RRs of 2.45, 2.12, and 1.93, respectively. Participants unable to complete the balance tests had a higher hip fracture risk than those with scores of 4, with an RR of 2.16. Walking speed and handgrip strength were not associated with increased hip fracture risk.
Conclusion: Among community-dwelling elderly Chinese people, overall SPPB score, as well as performance on repeated chair stands and balance tests within the SPPB, were significantly and independently associated with increased hip fracture risk. These indicators could be used to predict hip fracture in clinical settings.

1. Introduction

Hip fracture in the elderly is a major public health problem. The number of hip fractures is predicted to increase to 7.3–21.3 million globally and 5.9 million in China by 2050 [1,2]. Hip fractures cause morbidity, disability, and loss of independence worldwide [3,4], and a rising rate of fractures will markedly increase the associated economic burden in the coming decades [5]. Most hip fractures are a direct result of a fall [6], and poor neuromuscular function increases the risk of a fall in older adults [7,8]. Although many different physical performance tests are considered sensitive to indicators of the risk of falling, there is no gold standard tool to assess the risk of a fall in elderly patients [9]. Several scales and tests have been described and validated in the literature [9]. The Berg balance scale is one of the balance-evaluating tools with a good sensitivity and specificity in predicting falls [9], but its accuracy depends on the setting of application [10]. Physical performance tests have been combined to produce batteries of tests for addressing fall risk in the elderly [11–13]. In many cases, the loss of balance during walking or an inability to maintain the body's center of gravity over its base of support increases the risk of a fall in older adults. Moreover, lower limb neuromuscular function plays an important role in maintaining both static and dynamic balance [14]. In this regard, the Established Populations for Epidemiological Studies in the Elderly (EPESE) short physical performance battery (SPPB) is one of the combinations that is most often used to assess the lower extremity strength in older adults. SPPB consists of three tests (walking speed test, balance tests and repeated chair stands test) that provide information on several motor domains such as the static and dynamic balance, coordination,

* Corresponding author at: School of Public Health, Guangzhou Medical University, Xinzao Town, Panyu District, GuangZhou, PR China.
E-mail address: zhaoqian121@126.com (Q. Zhao).

https://doi.org/10.1016/j.maturitas.2021.01.003
Received 9 March 2020; Received in revised form 6 January 2021; Accepted 17 January 2021
Available online 22 January 2021

and strength of lower limbs [15]. Furthermore, the SPPB score is a simple and inexpensive indicator, which has been widely shown to predict outcomes such as falls, death, disability, hospitalization and nursing home admission [16–23]. Despite the association between fall and fracture risks, a limited number of studies have evaluated the association between physical performance and fracture risk [24]. However, what has not been clearly demonstrated is whether the SPPB is associated with hip fracture risk in the elderly. Therefore, the aim of the present study is to evaluate the association between hip fracture and the SPPB and handgrip strength in the China Health and Retirement Longitudinal Study (CHARLS) cohort.

2. Methods

2.1. Study participants and design

The data set for this study originated from a follow-up survey of the CHARLS [25], a national longitudinal survey of older persons conducted in 2011, 2013 and 2015 and sponsored by Peking University. The national baseline survey began in June 2011 and ended in March 2012 and was conducted in 28 provinces, 150 countries/districts, and 450 villages/urban communities across the country. The CHARLS sample is representative of people aged 45 and over living in households. All samples were drawn in four stages: county-level, neighborhood-level, household-level and respondent-level [25].

The CHARLS survey provided information on demographics, anthropometrics, physical measurements (balance tests, walking speed, repeated chair stands test, handgrip strength), health status, and functioning, including hip fracture [25].

Participants aged 60 years or older at baseline were included in this study. The exclusion criteria were as follows: (1) age less than 60 years or unknown; (2) previous history of hip fracture; (3) inability to understand or answer the study questionnaires; and (4) inability to understand physical performance measures or unwillingness to participate.

2.1.1. Hip fracture

The primary outcome of interest in this study was hip fracture. Hip fractures were measured by self-reported responses to the question "Have you fractured your hip since the last interview?" [26]. The interviewer would introduce the location of the hip bone in detail to let the respondent understand the specific definition of a hip fracture before the interview and ensure the accuracy of the respondent's answers. Hip fractures were categorized as either "Yes" or "No". The deadline for follow-up was January 2016.

2.1.2. Physical performance measures

2.1.2.1. Short physical performance battery. The SPPB has been shown to have high reliability, validity, and responsiveness in measuring physical function in an older community-dwelling population [27]. It includes three kinds of tests: walking speed (performed on a 2.5-meter course), balance tests (semi-tandem, full-tandem, and side-by-side stands), and repeated chair stands [15,19,20,28]. For each test, step-by-step demonstrations were provided to the interviewees. Each physical performance measurement was assigned a score ranging from 0 (worst performance) to 4 (best performance). A summary performance score ranging from 0 to 12 was then calculated that indicated the level of physical performance. The SPPB score was divided into three categories: 0–6 (poor performers), 7–9 (fair performers), and 10–12 (good performers) [15]. This scale's ability to predict the risk of incident physical disability [15,23], falling [16], rehospitalization [20–22], nursing home admission [29] and mortality [18] has been validated.

2.1.2.1.1. Balance tests. The balance tests included semi-tandem stands, full-tandem stands, and side-by-side stands. First, each participant was asked to stand with the side of the heel of one foot touching the big toe of the other foot for 10 s without stepping out of place or grabbing hold of anything. The participants unable to hold this position for full 10 s were asked to do the side-by-side stand for 10 s. The participants who were able to hold the semi-tandem position for the full 10 s without stepping out of place or grabbing hold of anything went on to the full-tandem stand, with the heel of one foot in front of and touching the toes of the other foot for approximately 30 or 60 s, depending on the participants' age. If the participant was 70 years old or older, a 30-second full-tandem balance measurement was carried out. If the participant was younger than 70 years old, a 60-second full-tandem balance task was given. Scores were assigned based on how long the position could be held, as follows: side-by-side 10 s and semi-tandem $<$10 s (1 point), semi-tandem 10 s and full-tandem 0−2 s (2 points), full-tandem 3−9 s (3 points), full-tandem 10 s or longer (4 points), and unable to perform the tests (0 points) [19].

2.1.2.1.2. Walking speed. Participants older than 60 years walked two laps on a 2.5-meter course at a normal pace, and the time was measured. Walking sticks and other forms of aid were allowed if necessary. Walking speed was calculated, and the faster speed of the two laps was used for statistical analysis. The criteria for the scores were as follows: $\leq$0.43 m/s (1 point), 0.44 to 0.60 m/s (2 points), 0.61 to 0.77 m/s (3 points), $\geq$0.78 m/s (4 points), and unable to perform the test (0 points) [19].

2.1.2.1.3. Repeated chair stand test. A standard chair was used with a height of 47 cm measured from the floor to the seat. The participants kept their arms folded across their chest and sat on the chair. When they were ready, they stood up straight and then sat down again as fast as they could five times without stopping or using their arms to push themselves off the chair. The total time to complete the five stands was recorded, and the scores were given as follows: $\geq$16.7 s (1 point), 13.7–16.6 s (2 points), 11.2–13.6 s (3 points), $\leq$11.1 s (4 points), and unable to perform the test (0 points) [19].

2.1.2.2. Handgrip strength. Handgrip strength was measured for each hand twice with a dynamometer (YuejianTM WL-1000 dynamometer). The participants stood and held the dynamometer at a right angle and squeezed the meter as hard as possible for several seconds and then let go. Two measurements were recorded for each hand. A maximum of four measurements was used for statistical analysis. The handgrip strength was organized into quartiles according to the sample distribution.

2.1.3. Covariate assessment at baseline

2.1.3.1. Demographic variables. Demographic variables included age, gender and education. Education levels included no formal education, primary school or below, middle school, and high school or above.

2.1.3.2. Personal health-related behaviors. Personal health-related behaviors included smoking and alcohol drinking. The three categories for cigarette smoking were never smoked, formerly smoked, and currently smoke. The frequency of alcohol consumption was reported as never drank, once a month, or more than once a month.

2.1.3.3. Physical measurement. Physical measurements included body weight and height. Body weight and height were measured with an OmronTM HN-286 scale and SecaTM213 stadiometer, respectively. Body mass index (BMI) was calculated as weight in kilograms divided by the height squared in meters.

2.1.3.4. Self-reported chronic diseases. Self-reported chronic diseases were recorded based on the responses to the question "Have you been diagnosed with XX by a doctor?" and included hypertension, dyslipidemia, diabetes or high blood sugar, chronic lung diseases (excluding tumors or cancer), heart attack, stroke, kidney diseases (excluding tumors or cancer), memory-related diseases (included Alzheimer's

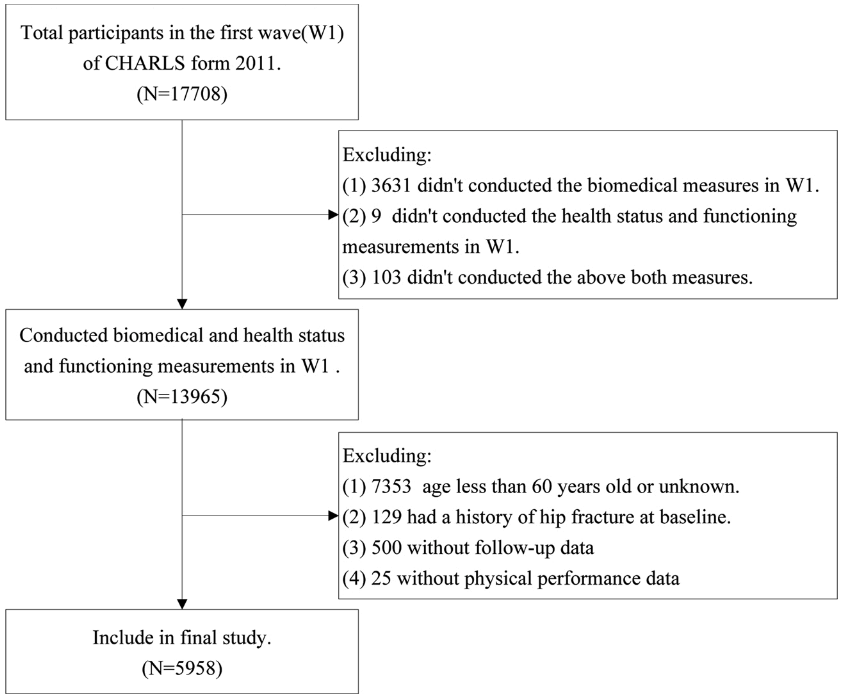

Fig. 1. Flowchart showing the selection of the elder participants enrolled in the study of hip fracture in China.

disease, brain atrophy, and Parkinson's disease), and arthritis or rheumatism. These diseases were chosen based on a longitudinal pilot study. The chronic diseases were categorized as "Yes" or "No".

2.1.3.5. Falls. History of falls was recorded based on self-reported responses to the question "Have you fallen down in the last two years?" If the participants answered "yes", they were considered to have prior falls.

2.1.3.6. Cognitive function. Cognitive function was measured using the cognitive dimensions of mental intactness, episodic memory, and visuospatial abilities, including (i) the Telephone Interview for Cognitive Status (TICS-10); (ii) word recall; and (iii) figure drawing [30,31]. Consistent with a prior CHARLS publication, the overall cognitive score was calculated as a sum score of TICS-10, word recall and figure drawing and ranges from 0 to 21.

Ten items from TICS-10 were used in CHARLS: serial subtraction of 7 from 100 (up to five times), date (year, month, day), the day of the week, and season of the year. TICS-10 primarily measures the executive functions of orientation to time and attention [32,33], and the score is an aggregate number of correct answers, which ranges from 0 to 10 [30,34].

Word recall. Participants were asked to memorize and immediately recall as many words as they could in any order after interviewers read a list of 10 Chinese nouns (immediate recall). Respondents were also asked to recall as many of the original words as possible 4−10 min later (delayed recall) [30]. The episodic memory score was calculated as the average number of correct immediate and delayed word recalls and ranged from 0 to 10, with higher scores indicating better memory performance [35].

Figure drawing. Participants were shown a picture of two overlapped pentagons and asked to draw a similar figure. Participants who successfully completed the task received a score of 1, and those who failed received a score of 0 [30]. Figure drawing was designed to assess visuospatial abilities [36].

2.2. Statistical analysis

Averages and variations were described as means and standard deviations (SDs) for normal variables and as medians and interquartile ranges (IQRs) for non-normal variables. Categorical variables are shown as ratios or percentages. The χ^2 test was used to analyze the difference between categorical variables. The T test was used to analyze the difference between normal quantitative variables, and the Willcoxon test was used to compare the difference between non-normal distributed variables. The Cochran-Armitage trend test [37,38] was used to assess the linear association between physical performance at baseline and subsequent hip fracture. Binary logistic regression models were used to evaluate the association between hip fracture and physical performance measures. Since previous studies suggested that certain covariates might affect the associations between physical performance and hip fracture [24,26,39–41], two models were constructed to control potential confounders. Model 1 estimates the effects of different physical performance scores adjusting only for age. Age, gender, body mass index and significant variables from univariate analysis were further included in model 2. Significant variables included education, falls, and chronic diseases (including diabetes, chronic lung diseases, kidney diseases, arthritis or rheumatism). To explore whether cognition could modify these associations, the above-mentioned logistic regression models were applied to a subset (N = 4560) with cognitive function data. Cognitive function was further included in model 3, and interactive effects of cognitive function and physical performance on hip fracture were examined. Risk ratios (RRs) and their corresponding 95 % confidence intervals (95 % CI) are shown. R packages were used for all statistical analyses, and the significance level was 0.05 (two-tailed).

3. Results

The 2011 CHARLS baseline survey enrolled 17,708 participants. Among them, 3,743 participants were excluded from the study because they failed to produce either a physical measurement or a health report; 7,353 participants were dropped because they were less than 60 years old or their age was unknown; 129 participants had a history of hip fracture at baseline, 550 participants had no follow-up data, and 25 had

Table 1
Characteristics of surveyed participants at baseline (2011).

	All sample (n = 5958)	Non-fracture (n = 5778)	Hip fracture (n = 180)	$t / \chi^2 / Z$	*P*
Age (mean ± sd)	67.49 ± 6.69	67.46 ± 6.67	68.85 ± 7.51	−2.469	0.014*
Gender, n (%)				1.604	0.205
Male	2958 (49.65)	2877 (49.79)	81 (45.00)		
Female	3000 (50.35)	2901 (50.21)	99 (55.00)		
Education, n (%)				17.156	0.002*
No formal education (illiterate)	2225 (37.34)	2144 (37.11)	81 (45.00)		
Did not finish primary school	1209 (20.29)	1160 (20.08)	49 (27.22)		
Elementary school	1497 (25.13)	1466 (25.37)	31 (17.22)		
Middle school	686 (11.51)	675 (11.68)	11 (6.11)		
High school or above	341 (5.73)	333 (5.76)	8 (4.45)		
Cigarette smoking, n (%)				0.827	0.661
Current smoked	1817 (30.50)	1366 (23.65)	39 (21.79)		
Former smoked	679 (11.40)	383 (6.63)	14 (7.82)		
Never smoked	3462 (58.10)	4028 (69.72)	126 (70.39)		
Alcohol drinking frequency, n (%)				0.635	0.728
Drink more than once a month	1405 (23.59)	1763 (30.51)	54 (30.00)		
Drink but less than once a month	397 (6.67)	662 (11.46)	17 (9.44)		
Never drank	4154 (69.74)	3353 (58.03)	109 (60.56)		
Body mass index, median (IQR)	22.55 (4.92)	22.55 (4.91)	22.74 (5.02)	−0.493	0.622
Fallen last two years, n (%)	1118 (18.89)	1065 (18.55)	53 (30.11)	14.909	<0.001 *
Chronic disease, n (%)					
Hypertension	2108 (35.38)	2046 (35.41)	62 (34.44)	0.071	0.790
Dyslipidemia	778 (13.06)	753 (13.03)	25 (13.89)	0.113	0.737
Diabetes	513 (8.61)	487 (8.43)	26 (14.44)	8.029	0.005*
Chronic lung diseases	1122 (18.83)	1076 (18.62)	46 (25.56)	5.490	0.019*
Heart attack	1076 (18.06)	1037 (17.95)	39 (21.67)	1.632	0.201
Stroke	272 (4.57)	261 (4.52)	11 (6.11)	1.018	0.313
Kidney disease	505 (8.48)	479 (8.29)	26 (14.44)	8.523	0.004*
Memory-related disease	212 (3.56)	204 (3.53)	8 (4.44)	0.425	0.515
Arthritis or rheumatism	2759 (46.31)	2651 (45.88)	108 (60.00)	13.996	<0.001 *
Cognition, Median (IQR)	10.5 (6.50)	10.5 (6.50)	8.5 (6.62)	−3.448	0.001*
Walking speed					
Unable, n (%)	222 (3.73)	210 (3.97)	12 (7.19)	4.296	0.038*
Complete (m/s), Median (IQR)	0.82 (0.30)	0.68 (0.30)	0.63 (0.26)	−0.308	0.002*
Balance test score, Median (IQR)	4 (0)	4 (0)	4 (0)	−3.043	0.002*
Repeated chair stands					
Unable, n (%)	520 (8.73)	489 (8.47)	31 (17.22)	16.787	<0.001 *
Complete (s), Median (IQR)	13.54 (6.28)	10.69 (4.68)	12.54 (5.22)	−4.337	<0.001 *
Handgrip strength					
Unable, n (%)	93 (1.56)	90 (1.57)	3 (1.68)	0.012	0.914
Complete (N/m^2), Median (IQR)	36.50 (13.00)	10.69 (4.68)	12.54 (5.22)	−3.496	<0.001 *
SPPB, Median (IQR)	10 (3.00)	10 (3.00)	9 (4.00)	−5.264	<0.001 *

Note:
Memory-related disease included Alzheimer's disease, brain atrophy, Parkinson's disease.
Abbreviations: SPPB, short physical performance battery; s/m, second/meter.
* $P < 0.05$.

no physical performance data. The final sample size for this study was 5,958 (Fig. 1).

Among the 5,958 eligible participants who were followed from 2011 to 2016, 180 new hip fractures were reported. The overall incidence of hip fracture for people older than 60 years was 3.0 %. The baseline demographic characteristics and physical performance of participants are shown in Table 1. Of the 5,958 participants, 2,958 (49.65 %) were male, and the mean age was 67.49 ± 6.69 years. In the univariate analysis, participants with a hip fracture tended to be older and had lower levels of education and cognitive function and higher proportions of falls and chronic diseases (including diabetes, chronic lung diseases, kidney diseases, and arthritis or rheumatism). Regarding physical performance, participants with hip fractures performed poorer than those without hip fractures with respect to the walking speed, balance test score, repeated chair stands, and handgrip strength.

The results of the Cochran-Armitage trend test showed that there was a clear gradient in the risk of hip fracture associated with the SPPB performance (*P* for trend < 0.001), walking speed (*P* for trend = 0.001), balance test (*P* for trend < 0.001), repeated chair stands (*P* for trend < 0.001) and handgrip strength (*P* for trend = 0.003) (Fig. 2).

Table 2 separately presents the associations of hip fractures with the physical performance measures and SPPB scores. In model 1, there is a risk gradient for hip fracture associated with all physical performance measures (all *P* for trend <0.05). Compared with good performers (SPPB 10−12), fair (SPPB 7−9) and poor (SPPB 0−6) performers had a greater risk of hip fracture, with RRs (95 % CI) of 1.57 (1.09–2.26) and 2.25 (1.47–3.43), respectively. Compared with participants with walking speed ≥ 0.78 m/s, RRs (95 % CI) of those with walking speed of 0.44−0.60 m/s and those unable to complete the walking speed test were 1.62 (1.03–2.54) and 2.22 (1.11–4.43), respectively. Participants unable to complete the balance tests had a greater risk of hip fracture than those with scores of 4, with an RR (95 % CI) of 3.05 (1.81–5.14). Participants who were unable to complete repeated chair stands and those who took ≥ 16.7 s, or 13.7–16.6 s to complete had higher risks of hip fracture, compared with those who completed the test in less than 11.1 s, with RRs (95 % CI) of 3.02 (1.89–4.82), 2.46 (1.53–3.94), and 2.16 (1.37–3.39), respectively. However, no associations were observed between the handgrip strength and the risk of hip fracture.

After further adjustment for potential confounding factors in model

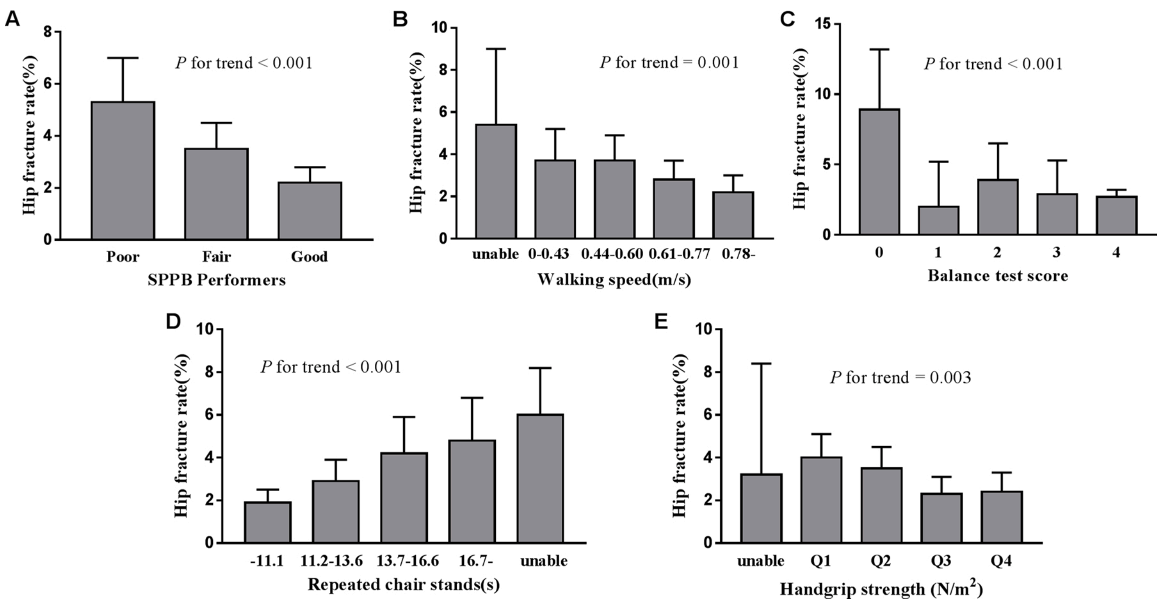

Fig. 2. Incidences of hip fracture at the levels of baseline physical performance.

2, a significant risk gradient for hip fracture was associated with the SPPB score (P for trend = 0.021) and repeated chair stands (P for trend < 0.001). Poor performers (SPPB 0−6) were approximately two times more likely to experience a hip fracture than good performers (SPPB 10−12) after multivariate adjustment (RR: 1.65; 95 % CI: 1.04–2.60). Participants unable to complete the balance tests had a higher risk of hip fracture than those with scores of 4, with an RR of 2.16 (1.14–4.11). The results of repeated chair stands and hip fractures were similar to those obtained using model 1. Compared with participants who completed repeated chair stands in less than 11.1 s, RRs (95 % CI) of those who were unable to complete and those who completed in ≥ 16.7 s or in 13.7–16.6 s were 2.45 (1.47–4.08), 2.12 (1.30–3.46) and 1.93 (1.22–3.06), respectively.

In the subset analysis, to explore whether cognition could modify the associations between physical performance and hip fracture, cognitive function was further included in model 3, and interactive effects of cognitive function and physical performance on hip fracture were examined. A total of 4560 participants with cognitive data were included, and the results are shown in Table 3. We found that cognition function had no effect on the relationship between physical performance and hip fracture. The results of the subset analysis were basically consistent with those of the whole sample analysis.

4. Discussion

This study evaluates the association between physical performance measures and hip fracture risk over approximately 4 years of follow-up in a large population of community-dwelling older people in China. The SPPB score has been used to predict falls, death, disability, hospitalization and nursing home admission [16–23]. Most studies have focused on the effect of a single physical performance score on the risk of hip fracture, but the relationship between the SPPB and hip fracture is unknown. To our knowledge, there is little in the literature on the association between hip fracture and the SPPB score.

However, in some previous epidemiological studies [24,42], poor physical performance on a single exam at baseline was related to an increased future incident risk of hip fracture. The relative risk factors from those studies were also not completely consistent. In the Osteoporotic Fractures (SOF) in Men study, older men who had poor performance with respect to walking speed and repeated chair stands and inability to complete the handgrip strength test at baseline had high hip fracture risk after multivariate adjustment [24]. In the EPIDOS prospective study of older women, lower walking speed and difficulty performing balance tests at baseline were statistically significantly associated with an increased risk of subsequent hip fracture after multivariate adjustment, while handgrip strength at baseline was not significantly related to hip fracture risk after multivariate adjustment [42].

In this study, inability to complete an exam or a lower score tended to be associated with an increased risk of hip fracture. The results obtained after multivariate adjustment (model 2 in Table 2) showed that those with poor performance on the SPPB (SPPB score 0−6) were approximately two times more likely to experience a hip fracture than good performers (SPPB score 10−12). The SPPB can provide information on several motor domains, such as the static and dynamic balance, coordination, and strength of lower limbs, which are often used to assess the lower extremity strength in older adults. Lower limb neuromuscular function is associated with hip fracture, and thus, the SPPB score was found to provide a good indication of global frailty in the elderly and appears to be a simple and reliable method for identifying people with hip fracture risk.

Among the different measures, repeated chair stands appeared pertinent because they significantly distinguished a gradient of risk for hip fracture between the low and high end of the functional spectrum after multivariate adjustment. The repeated chair stands measurement may be an efficient tool for screening elderly individuals with a higher risk of hip fracture and may easily identify large high-risk groups in the community. None of the other single physical performance measurements used in this study (walking speed, balance tests and handgrip strength), after multivariate adjustment, was able to distinguish a gradient of risk for hip fracture over the whole functional spectrum. For balance tests, those unable to complete the tests had approximately two times the risk of hip fracture than those who completed tests with a score of 4 after multivariate adjustment.

These findings are consistent with those of previous studies in which balance tests [24,42] and repeated chair stands [24,43] were associated with the risk of hip fracture in the elderly after multivariate adjustment. There were no relationships between walking speed and handgrip strength and hip fracture after multivariate adjustment, and the walking speed result was inconsistent with a previous study [24,40,42,44]. The handgrip strength result was consistent with the SOF study [24] but not consistent with the EPIDOS study [42]. This difference in the results is

Table 2
Hip fracture rate and RR (95 % CI) of hip fracture for SPPB, each SPPB variable considered individually and handgrip strength in all sample (N = 5958).

	Event/Total	RR (95 %CI)	
		Model 1	Model 2
SPPB Performers			
Per SD		0.71 (0.62, 0.82)*	0.79 (0.67, 0.92)*
Good (10−12)	68/3073	1.00	1.00
Fair (7−9)	56/1580	1.57 (1.09, 2.26) *	1.40 (0.97, 2.03)
Poor (0−6)	42/799	2.25 (1.47, 3.43)*	1.65 (1.04, 2.60)*
P for trend		<0.001*	0.021*
Walking speed(m/s)			
Per SD		0.31 (0.08, 1.13)	0.48 (0.12, 1.86)
≥0.78	34/1569	1.00	1.00
0.61−0.77	44/1595	1.24 (0.79, 1.96)	1.11 (0.70, 1.76)
0.44−0.60	47/1269	1.62 (1.03, 2.54)*	1.38 (0.87, 2.20)
≤0.43	30/805	1.57 (0.94, 2.62)	1.35 (0.80, 2.28)
Unable	12/222	2.22 (1.11, 4.43)*	1.17 (0.47, 2.88)
P for trend		0.010 *	0.212
Balance test score			
4	136/4960	1.00	1.00
3	9/308	1.01 (0.51, 2.00)	0.92 (0.46, 1.84)
2	12/308	1.32 (0.72, 2.43)	1.15 (0.62, 2.13)
1	3/151	0.64 (0.20, 2.04)	0.49 (0.15, 1.60)
0	19/214	3.05 (1.81, 5.14)*	2.16 (1.14, 4.11)*
P for trend		0.001*	0.210
Repeated chair stands(s)			
Per SD		1.19 (1.07, 1.32)*	1.16 (1.03, 1.30)*
≤11.1	58/2988	1.00	1.00
11.2–13.6	33/1156	1.47 (0.95, 2.27)	1.41 (0.91, 2.19)
13.7–16.6	30/713	2.16 (1.37, 3.39)*	1.93 (1.22, 3.06)*
≥16.7	28/578	2.46 (1.53, 3.94)*	2.12 (1.30, 3.46)*
Unable	31/520	3.02 (1.89, 4.82)*	2.45 (1.47, 4.08)*
P for trend		<0.001*	<0.001*
Handgrip strength (N/m^2)			
Per SD		0.80 (0.68, 0.94)*	0.87 (0.70, 1.07)
Quartile 4	34/1431	1.00	1.00
Quartile 3	33/1462	0.90 (0.55, 1.47)	0.81 (0.49, 1.34)
Quartile 2	48/1381	1.39 (0.89, 2.18)	1.01 (0.59, 1.72)
Quartile 1	61/1531	1.51 (0.97, 2.34)	1.14 (0.65, 2.01)
Unable	3/93	1.15 (0.34, 3.86)	0.42 (0.06, 3.25)
P for trend		0.023*	0.592

Note:
Model 1: Adjusted for age.
Model 2: Included covariate in model 1, and further adjusted for gender, body mass index, education level, falls and chronic diseases (include: diabetes, chronic lung diseases, kidney disease, arthritis or rheumatism).
Handgrip strength (N/m^2) in quartiles.
Abbreviations: SD, standard deviation; SPPB, short physical performance battery; Q, quartile; s/m, second/meter. s,second.
* $P<0.05$.

likely due to the different populations studied. The previous studies included only men or women, while this study included both men and women.

The study has the following strengths. We studied a large population of Chinese community-dwelling elderly people and followed them for almost 4 years. The participants were well characterized with a comprehensive set of measurements at baseline and follow-up. In addition, the physical performance tests were objective measurements representing basic aspects of physical performance. However, some limitations of this study should be noted. First, the diagnosis of hip fracture was based only on the participants' self-reports, which may have resulted in recall bias. Several measures were taken to reduce recall bias during the period of design and data collection, including careful research questions, an appropriate data collection method, and well-trained investigators. In particular, the interviewers introduced the location of the hip bone in detail to let the respondents understand the

Table 3
Hip fracture rate and RR (95 % CI) of hip fracture for SPPB, each SPPB variable considered individually and handgrip strength in subset sample (N = 4560).

	Event/ Total	RR (95 %CI)		
		Model 1	Model 2	Model 3
SPPB Performers				
Per SD		0.66 (0.56, 0.78)*	0.72 (0.60, 0.86)*	0.72 (0.60, 0.87)*
Good (10−12)	46/2475	1.00	1.00	1.00
Fair (7−9)	38/1165	1.75 (1.12, 2.71)*	1.55 (0.99, 2.42)	1.53 (0.98, 2.41)
Poor (0−6)	31/517	3.21 (1.96, 5.28)*	2.46 (1.45, 4.19)*	2.41 (1.41, 4.13) *
P for trend		<0.001*	0.001*	0.001*
Walking speed(m/s)				
Per SD		0.20 (0.04, 0.93)*	0.29 (0.06, 1.44)	0.30 (0.06, 1.52)
≥0.78	24/1272	1.00	1.00	1.00
0.61−0.77	25/1226	1.06 (0.60, 1.87)	0.99 (0.56, 1.75)	0.99 (0.56, 1.75)
0.44−0.60	34/934	1.87 (1.09, 3.19)*	1.63 (0.95, 2.82)	1.61 (0.94, 2.78)
≤0.43	23/581	1.99 (1.10, 3.60)*	1.71 (0.93, 3.13)	1.68 (0.91, 3.08)
Unable	9/147	3.09 (1.39, 6.87)*	1.73 (0.63, 4.71)	1.70 (0.62, 4.64)
P for trend		0.001*	0.023*	0.028*
Balance test score				
4	95/3881	1.00	1.00	1.00
3	8/225	1.41 (0.67, 2.95)	1.28 (0.61, 2.70)	1.27 (0.60, 2.67)
2	9/210	1.66 (0.82, 3.39)	1.43 (0.70, 2.93)	1.38 (0.67, 2.83)
1	2/101	0.73 (0.18, 3.05)	0.54 (0.13, 2.29)	0.53 (0.12, 2.21)
0	12/136	3.56 (1.87, 6.79)*	2.57 (1.18, 5.63)*	2.47 (1.13, 5.41)*
P for trend		0.001*	0.088	0.118
Repeated chair stands(s)				
Per SD		1.16 (1.02, 1.32)*	1.13 (0.99, 1.30)	1.13 (0.98, 1.30)
≤11.1	47/2420	1.00	1.00	1.00
11.2–13.6	21/896	1.20 (0.72, 2.03)	1.18 (0.69, 2.00)	1.17 (0.69, 1.98)
13.7–16.6	20/515	2.00 (1.17, 3.43)*	1.74 (1.00, 3.00)*	1.71 (0.99, 2.97)
≥16.7	18/381	2.44 (1.38, 4.30)*	2.16 (1.20, 3.87)*	2.11 (1.18, 3.80)*
Unable	20/346	3.00 (1.72, 5.23)*	2.39 (1.31, 4.37)*	2.31 (1.26, 4.24)*
P for trend		<0.001*	0.001*	0.001*
Handgrip strength (N/m^2)				
Per SD		0.82 (0.68, 0.99)*	0.90 (0.70, 1.14)	0.92 (0.72, 1.17)
Quartile 4	30/1223	1.00	1.00	1.00
Quartile 3	24/1183	0.79 (0.46, 1.36)	0.70 (0.40, 1.24)	0.70 (0.40, 1.23)
Quartile 2	30/1008	1.16 (0.69, 1.95)	0.81 (0.43, 1.52)	0.79 (0.42, 1.49)
Quartile 1	39/1045	1.40 (0.84, 2.31)	1.04 (0.55, 1.99)	1.00 (0.52, 1.91)
Unable	2/57	1.27 (0.29, 5.51)	0.62 (0.08, 4.82)	0.59 (0.08, 4.62)
P for trend		0.089	0.774	0.888

Note:
Model 1: Adjusted for age.
Model 2: Included covariate in model 1, and further adjusted for gender, body mass index, education level, falls and chronic diseases (include: diabetes, chronic lung diseases, kidney disease, arthritis or rheumatism).
Model 3: Included covariates in model 2, and further adjusted for cognitive.
Handgrip strength (N/m^2) in quartiles.

Abbreviations: SD, standard deviation; SPPB, short physical performance battery; Q, quartile; s/m, second/meter; s,second.
* $P < 0.05$.

specific definition of hip fracture and ensure the accuracy of the respondent's answer in the survey. In addition, chronic diseases were self-reported and thus the information may not have been completely accurate. Moreover, some elderly patients would not remember their trivial falls in the last two years, which may have resulted in recall bias as it has been suggested that individuals are more likely to recall a fall that resulted in injury [45]. Thus, the difference in falls between non-fracture and hip fracture participants may have been overestimated. Although we could not explore this bias with our data, several measures were taken to reduce recall bias during the period of design and data collection, including careful research questions, an appropriate data collection method, and well-trained investigators. Finally, many participants were excluded from the study due to a lack of physical performance data, which may have led to bias.

In conclusion, a large prospective observational study of Chinese community-dwelling elderly people found that the SPPB score and performance on balance tests and repeated chair stands were significantly associated with hip fracture. The study confirmed that poor physical performance was associated with an increased risk of hip fracture in older persons and suggested that physical performance is an important tool to predict hip fracture.

Contributors

Bi-Xia Zhong contributed to the study design, data cleaning, analysis, and interpretation, as well as manuscript drafting and revision.

Hai-Li Zhong contributed to the study design, data interpretation, as well as manuscript drafting and revision.

Guan-Qun Zhou contributed to the data interpretation, and manuscript drafting.

Wen-Qi Xu contributed to the data cleaning and interpretation.

Ying Lu contributed to the manuscript drafting.

Qian Zhao contributed to the study design, study conduct, data interpretation, manuscript revision, and takes responsibility for the integrity of the data analysis.

All authors read and approved the final version of the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

Funding

This work was supported by Guangdong Science and Technology Department of China [grant numbers 2014A020212689].

Ethics

Each participant included in this study signed a written informed consent form before taking the survey. Ethics approval for the data collection in CHARLS was obtained from the Biomedical Ethics Review Committee of Peking University (IRB00001052-11015).

Data sharing and collaboration

The data underlying the results presented in the study is from the China Health and Retirement Longitudinal Study, an open database on geriatrics (http://charls.pku.edu.cn/pages/data/111/zh-cn.html).

Provenance and peer review

This article was not commissioned and was externally peer reviewed.

References

[1] B. Gullberg, O. Johnell, J.A. Kanis, World-wide projections for hip fracture, Osteoporos. Int. 7 (5) (1997) 407–413, https://doi.org/10.1007/pl00004148.
[2] OASS Beijing, China Health Promotion Foundation, White Paper, China, 2008, 2008.
[3] P. Haentjens, J. Magaziner, C.S. Colon-Emeric, D. Vanderschueren, K. Milisen, B. Velkeniers, S. Boonen, Meta-analysis: excess mortality after hip fracture among older women and men, Ann. Intern. Med. 152 (6) (2010) 380–390, https://doi.org/10.7326/0003-4819-152-6-201003160-00008.
[4] O. Johnell, J.A. Kanis, An estimate of the worldwide prevalence, mortality and disability associated with hip fracture, Osteoporos. Int. 15 (11) (2004) 897–902, https://doi.org/10.1007/s00198-004-1627-0.
[5] L. Si, T.M. Winzenberg, Q. Jiang, M. Chen, A.J. Palmer, Projection of osteoporosis-related fractures and costs in China: 2010–2050, Osteoporos. Int. 26 (7) (2015) 1929–1937, https://doi.org/10.1007/s00198-015-3093-2.
[6] L. Nyberg, Y. Gustafson, D. Berggren, B. Brännström, G. Bucht, Falls leading to femoral neck fractures in Lucid Older People, J. Am. Geriatr. Soc. 44 (2) (1996) 156–160, https://doi.org/10.1111/j.1532-5415.1996.tb02432.x.
[7] J.D. Moreland, J.A. Richardson, C.H. Goldsmith, C.M. Clase, Muscle weakness and falls in older adults: a systematic review and meta-analysis, J. Am. Geriatr. Soc. 52 (7) (2004) 1121–1129, https://doi.org/10.1111/j.1532-5415.2004.52310.x.
[8] V.S. Stel, J.H. Smit, S.M.F. Pluijm, P. Lips, Balance and mobility performance as treatable risk factors for recurrent falling in older persons, J. Clin. Epidemiol. 56 (7) (2003) 659–668, https://doi.org/10.1016/s0895-4356(03)00082-9.
[9] S. Park, Tools for assessing fall risk in the elderly: a systematic review and meta-analysis, Aging Clin. Exp. Res. 30 (1) (2018) 1–16, https://doi.org/10.1007/s40520-017-0749-0.
[10] P.D. Neuls, T.L. Clark, N.C. Van Heuklon, J.E. Proctor, B.J. Kilker, M.E. Bieber, A. V. Donlan, S.A. Carr-Jules, W.H. Neidel, R.A. Newton, Usefulness of the berg balance scale to predict falls in the elderly, J. Geriatr. Phys. Ther. 34 (1) (2011) 3–10, https://doi.org/10.1097/JPT.0b013e3181ff2b0e.
[11] L.K. Boulgarides, S.M. McGinty, J.A. Willett, C.W. Barnes, Use of clinical and impairment-based tests to predict falls by community-dwelling older adults, Phys. Ther. 83 (4) (2003) 328–339.
[12] A.Y.Y. Chiu, S.S.Y. Au-Yeung, S.K. Lo, A comparison of four functional tests in discriminating fallers from non-fallers in older people, Disabil. Rehabil. 25 (1) (2003) 45–50.
[13] S.R. Lord, H.B. Menz, A. Tiedemann, A physiological profile approach to falls risk assessment and prevention, Phys. Ther. 83 (3) (2003) 237–252.
[14] U. Granacher, T. Muehlbauer, A. Gollhofer, R.W. Kressig, L. Zahner, An intergenerational approach in the promotion of balance and strength for fall prevention - a mini-review, Gerontology 57 (4) (2011) 304–315, https://doi.org/10.1159/000320250.
[15] J.M. Guralnik, L. Ferrucci, E.M. Simonsick, M.E. Salive, R.B. Wallace, Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability, N. Engl. J. Med. 332 (9) (1995) 556–562, https://doi.org/10.1056/NEJM199503023320902.
[16] N. Veronese, F. Bolzetta, E.D. Toffanello, S. Zambon, M. De Rui, E. Perissinotto, A. Coin, M. Corti, G. Baggio, G. Crepaldi, G. Sergi, E. Manzato, Association between short physical performance battery and falls in older people: the progetto veneto anziani study, Rejuv. Res. 17 (3) (2014) 276–284, https://doi.org/10.1089/rej.2013.1491.
[17] S. Vasunilashorn, A.K. Coppin, K.V. Patel, F. Lauretani, L. Ferrucci, S. Bandinelli, J. M. Guralnik, Use of the short physical performance battery score to predict loss of ability to walk 400 meters: analysis from the InCHIANTI study, J. Gerontol. Series A 64A (2) (2009) 223–229, https://doi.org/10.1093/gerona/gln022.
[18] Y. Rolland, V. Lauwers-Cances, M. Cesari, B. Vellas, M. Pahor, H. Grandjean, Physical performance measures as predictors of mortality in a cohort of community-dwelling older French women, Eur. J. Epidemiol. 21 (2) (2006) 113–122, https://doi.org/10.1007/s10654-005-5458-x.
[19] J.M. Guralnik, L. Ferrucci, C.F. Pieper, S.G. Leveille, K.S. Markides, G.V. Ostir, S. Studenski, L.F. Berkman, R.B. Wallace, Lower extremity function and subsequent disability: consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery, J. Gerontol. A Biol. Sci. Med. Sci. 55 (4) (2000) M221–M231, https://doi.org/10.1093/gerona/55.4.m221.
[20] B.W. Penninx, L. Ferrucci, S.G. Leveille, T. Rantanen, M. Pahor, J.M. Guralnik, Lower extremity performance in nondisabled older persons as a predictor of subsequent hospitalization, J. Gerontol. A Biol. Sci. Med. Sci. 55 (11) (2000) M691–M697, https://doi.org/10.1093/gerona/55.11.m691.
[21] S. Studenski, S. Perera, D. Wallace, J.M. Chandler, P.W. Duncan, E. Rooney, M. Fox, J.M. Guralnik, Physical performance measures in the clinical setting, J. Am. Geriatr. Soc. 51 (3) (2003) 314–322, https://doi.org/10.1046/j.1532-5415.2003.51104.x.
[22] S. Volpato, M. Cavalieri, F. Sioulis, G. Guerra, C. Maraldi, G. Zuliani, R. Fellin, J. M. Guralnik, Predictive value of the short physical performance battery following hospitalization in older patients, J. Gerontol. Series A Biol. Sci. Med. Sci. 66A (1) (2010) 89–96, https://doi.org/10.1093/gerona/glq167.
[23] G.V. Ostir, K.S. Markides, S.A. Black, J.S. Goodwin, Lower body functioning as a predictor of subsequent disability among older Mexican Americans, J. Gerontol. A Biol. Sci. Med. Sci. 53 (6) (1998) M491–M495, https://doi.org/10.1093/gerona/53a.6.m491.
[24] P.M. Cawthon, R.L. Fullman, L. Marshall, D.C. Mackey, H.A. Fink, J.A. Cauley, S. R. Cummings, E.S. Orwoll, K.E. Ensrud, Physical performance and risk of hip

fractures in older men, J. Bone Miner. Res. 23 (7) (2008) 1037–1044, https://doi.org/10.1359/jbmr.080227.
[25] Y. Zhao, Y. Hu, J.P. Smith, J. Strauss, G. Yang, Cohort profile: the china health and retirement longitudinal study (CHARLS), Int. J. Epidemiol. 43 (1) (2014) 61–68, https://doi.org/10.1093/ije/dys203.
[26] Y. Ren, J. Hu, B. Lu, W. Zhou, B. Tan, Prevalence and risk factors of hip fracture in a middle-aged and older Chinese population, Bone 122 (2019) 143–149, https://doi.org/10.1016/j.bone.2019.02.020.
[27] E. Freiberger, P. de Vreede, D. Schoene, E. Rydwik, V. Mueller, K. Frändin, M. Hopman-Rock, Performance-based physical function in older community-dwelling persons: a systematic review of instruments, Age Ageing 41 (6) (2012) 712–721, https://doi.org/10.1093/ageing/afs099.
[28] L. Ferrucci, B.W. Penninx, S.G. Leveille, M.C. Corti, M. Pahor, R. Wallace, T. B. Harris, R.J. Havlik, J.M. Guralnik, Characteristics of nondisabled older persons who perform poorly in objective tests of lower extremity function, J. Am. Geriatr. Soc. 48 (9) (2000) 1102–1110, https://doi.org/10.1111/j.1532-5415.2000.tb04787.x.
[29] J.M. Guralnik, E.M. Simonsick, L. Ferrucci, R.J. Glynn, L.F. Berkman, D.G. Blazer, P.A. Scherr, R.B. Wallace, A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission, J. Gerontol. 49 (2) (1994) M85–M94, https://doi.org/10.1093/geronj/49.2.M85.
[30] W. Huang, Y. Zhou, Effects of education on cognition at older ages: evidence from China's great famine, Soc. Sci. Med. 98 (2013) 54–62, https://doi.org/10.1016/j.socscimed.2013.08.021.
[31] J. Li, P.Z. Cacchione, N. Hodgson, B. Riegel, B.T. Keenan, M.T. Scharf, K. C. Richards, N.S. Gooneratne, Afternoon napping and cognition in chinese older adults: findings from the China health and retirement longitudinal study baseline assessment, J. Am. Geriatr. Soc. 65 (2) (2017) 373–380, https://doi.org/10.1111/jgs.14368.
[32] V.C. Pangman, J. Sloan, L. Guse, An examination of psychometric properties of the mini-mental state examination and the standardized mini-mental state examination: implications for clinical practice, Appl. Nurs. Res. ANR 13 (4) (2000) 209–213, https://doi.org/10.1053/apnr.2000.9231.
[33] M.F. Folstein, S.E. Folstein, P.R. McHugh, "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician, J. Psychiatr. Res. 12 (3) (1975) 189–198, https://doi.org/10.1016/0022-3956(75)90026-6.
[34] X. Lei, J.P. Smith, X. Sun, Y. Zhao, Gender differences in cognition in China and reasons for change over time: evidence from CHARLS, J. Econ. Ageing 4 (2014) 46–55, https://doi.org/10.1016/j.jeoa.2013.11.001.
[35] J.J. McArdle, G.G. Fisher, K.M. Kadlec, Latent variable analyses of age trends of cognition in the health and retirement study, 1992-2004, Psychol. Aging 22 (3) (2007) 525–545, https://doi.org/10.1037/0882-7974.22.3.525.
[36] P.S. Mathuranath, P.J. Nestor, G.E. Berrios, W. Rakowicz, J.R. Hodges, A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia, Neurology 55 (11) (2000) 1613–1620, https://doi.org/10.1212/01.wnl.0000434309.85312.19.
[37] W.G. Cochran, Some methods for strengthening the common χ2Tests, Biometrics 10 (4) (1954) 417–451, https://doi.org/10.2307/3001616.
[38] P. Armitage, Tests for linear trends in proportions and frequencies, Biometrics 11 (3) (1955) 375–386, https://doi.org/10.2307/3001775.
[39] A.V. Schwartz, D.E. Sellmeyer, K.E. Ensrud, J.A. Cauley, H.K. Tabor, P.J. Schreiner, S.A. Jamal, D.M. Black, S.R. Cummings, Older women with diabetes have an increased risk of fracture: a prospective study, J. Clin. Endocrinol. Metab. 86 (1) (2001) 32–38, https://doi.org/10.1210/jcem.86.1.7139.
[40] J.A. Cauley, P.M. Cawthon, K.E. Peters, S.R. Cummings, K.E. Ensrud, D.C. Bauer, B. C. Taylor, J.M. Shikany, A.R. Hoffman, N.E. Lane, D.M. Kado, M.L. Stefanick, E. S. Orwoll, Risk Factors for Hip Fracture in Older Men: The Osteoporotic Fractures in Men Study (MrOS), J. Bone Miner. Res. 31 (10) (2016) 1810–1819, https://doi.org/10.1002/jbmr.2836.
[41] O. Zaslavsky, S. Zelber-Sagi, S.L. Gray, A.Z. LaCroix, R.L. Brunner, R.B. Wallace, M. J. O'Sullivan, B. Cochrane, N.F. Woods, Comparison of frailty phenotypes for prediction of mortality, incident falls, and hip fracture in older women, J. Am. Geriatr. Soc. 64 (9) (2016) 1858–1862, https://doi.org/10.1111/jgs.14233.
[42] P. Dargent-Molina, F. Favier, H. Grandjean, C. Baudoin, A.M. Schott, E. Hausherr, P.J. Meunier, G. Breart, Fall-related factors and risk of hip fracture: the EPIDOS prospective study, Lancet 348 (9021) (1996) 145–149, https://doi.org/10.1016/s0140-6736(96)01440-7.
[43] S.R. Cummings, M.C. Nevitt, W.S. Browner, K. Stone, K.M. Fox, K.E. Ensrud, J. Cauley, D. Black, T.M. Vogt, Risk factors for hip fracture in White Women. Study of osteoporotic fractures research group, N. Engl. J. Med. 332 (12) (1995) 767–773, https://doi.org/10.1056/NEJM199503233321202.
[44] B.C. Taylor, P.J. Schreiner, K.L. Stone, H.A. Fink, S.R. Cummings, M.C. Nevitt, P. J. Bowman, K.E. Ensrud, Long-term prediction of incident hip fracture risk in elderly white women: study of osteoporotic fractures, J. Am. Geriatr. Soc. 52 (9) (2004) 1479–1486, https://doi.org/10.1111/j.1532-5415.2004.52410.x.
[45] S.R. Cummings, M.C. Nevitt, S. Kidd, Forgetting Falls. The Limited Accuracy of Recall of Falls in the Elderly, J. Am. Geriatr. Soc. 36 (7) (1988) 613–616, https://doi.org/10.1111/j.1532-5415.1988.tb06155.x.