Journal of Bone and Mineral Metabolism (2019) 37:935–942
https://doi.org/10.1007/s00774-019-00998-x

ORIGINAL ARTICLE

Predicting radiological vertebral fractures with a combined physical function and body composition scoring system

On Takeda[1,2] · Gentaro Kumagai[1] · Kanichiro Wada[1] · Hitoshi Kudo[1] · Toru Asari[1] · Seiya Ota[1] · Daisuke Chiba[1] · Shigeyuki Nakaji[2] · Yasuyuki Ishibashi[1]

Received: 16 December 2018 / Accepted: 4 March 2019 / Published online: 27 March 2019

Abstract

The objective of this study was to investigate the incidence of vertebral fractures (VFx) and the value of physical function (PF) and body composition (BC) for predicting VFx in a Japanese population. This study included 307 subjects (113 men, 194 women) at least 40 years of age who were assessed at community health check-ups in 2008 and 2016. PF was assessed by grip strength and by single-leg stance, timed up-and-go, and 30-s chair stand tests, each scored from 0 to 3 for a possible total of 12 points (higher scores reflect lower function). BC was scored on bioelectrical impedance measurements of trunk and appendage muscle volume, with 6 possible points. We diagnosed radiological VFx semiquantitatively on lateral views of the lumbar spine, and measured bone mineral status by quantitative ultrasound (QUS) of the calcaneus. We conducted logistic regression analysis with VFx as the dependent variable and age, sex, BMI, QUS, PF score, and BC score as independent variables. In 8 years, 36 participants (12%) sustained new VFx. After correcting for age, sex, BMI, and QUS, the odds of VFx increased with a PF score ≥ 8 (OR 5.6; 95% CI 1.21–25.90; $P = 0.028$) and increased further with a PF + BC score ≥ 9 (OR 8.1; 95% CI 1.80–36.00; $P < 0.01$). Both PF and BC are important for predicting fragility fractures. The scoring system used here may reflect small differences better than categorical (single cutoff) definitions of poor function.

Keywords Fragility fractures · Osteoporosis · Radiographic vertebral fractures · Risk prevention

Introduction

Fragility fractures increase mortality, morbidity, and economic burden in the elderly [1] and these burdens will increase as the global population ages [2]. The cost of treating all postmenopausal women to prevent fractures may be nearly as great as the cost of treating the fractures, emphasizing the need to develop effective screening and prevention methods to identify and treat individuals at higher risk [1]. Although current screening methods based on medical comorbidities and bone mineral density (BMD) are somewhat successful, Crandall et al. [3] reported that these methods may not predict the majority of osteoporotic fractures in women aged 50–64. Furthermore, caution is needed when adapting these methods for use outside their native context [4].

Another issue with screening methods is that an endpoint of major osteoporotic fractures and self-reported fractures may overlook subclinical vertebral fractures (VFx), which also increase morbidity and the risk of future fractures and decrease quality of life (QOL) [5–7].

There is growing evidence that relationships between muscle, bone, and fat are involved in fragility fractures, and that these factors should be considered when predicting such fractures [8, 9]. Cross-sectional studies reported that patients with fragility fractures had lower muscle volumes [10, 11]. However, most studies measure body composition (BC) by whole-body DXA [12], which is difficult to translate to screening methods and to reproduce in epidemiological studies. Studies indicate that physical function (PF) tests may also be useful for predicting falls and fractures [13, 14]. Although PF measurements are important targets in exercise interventions and programs to prevent fractures [15], there

✉ On Takeda
m03055ot@gmail.com; onkochishin@hirosaki-u.ac.jp

[1] Department of Orthopaedic Surgery, Hirosaki University Graduate School of Medicine, 5 Zaifu-cho, Hirosaki, Aomori 036-8562, Japan

[2] Department of Social Medicine, Hirosaki University Graduate School of Medicine, 5 Zaifu-cho, Hirosaki, Aomori 036-8562, Japan

is little information to indicate whether PF and BC measurements can be used to predict radiographic fragility fractures. Thus, this study aimed to determine the longitudinal relationships between PF and BC and the incidence of morphological VFx in a Japanese population, and to investigate whether BC and PF can predict radiographic VFx.

Materials and methods

Study population

Study data were derived from a community-based public health project, introduced by our institution in 2005, that aims to prolong a healthy lifespan in the general population. The project provides annual health check-ups to approximately 1000 volunteers who live in a city in northern Japan and are at least 20 years old, and collects basic anthropometric data and lifestyle information, nutritional status, medical history, biomechanical test results, and biochemical blood and urine test results, as well as questionnaire and test findings related to examinations by various specialists. In 2008, 886 people participated in this public health project. Previous studies of this population reported an association between lumbar spondylosis and mobility [16], and between neck pain and LDL cholesterol [17].

For the present study, we looked for subjects who participated in our public health project check-ups in both 2008 and 2016, were at least 40 years old at the 2008 checkup, and had the following data available: quantitative ultrasound (QUS) of the calcaneus, bioelectrical impedance analysis (BIA), lumbar spine X-rays, and PF tests at both the 2008 and 2016 examinations. Of 374 subjects, 307 met all of the criteria and participated in the study. There were 113 men and 194 women (average age when enrolling in this study: men 56.6 ± 10.2 years; women 58.2 ± 9.4 years). This study was approved by the Ethics Committee of Hirosaki University Graduate School of Medicine, and all subjects provided written consent before participating in the study.

BC measurements

BC was measured with a BIA scanner (MC-190, Tanita, Japan), a method that has been validated in both young and older individuals and is correlated with DXA BC measurements [18, 19]. BIA has the advantage of being simple, fast, and relatively inexpensive, and it does not expose the subject to ionizing radiation. Its use and accuracy in epidemiological studies has been described [20]. BIA was used to measure the whole-body ratio of fat to body weight and to calculate the skeletal muscle index (SMI) for the trunk, legs, and appendages. The SMI, which is calculated as the muscle volume divided by height (meters) squared, is widely used as an index of corrected muscle mass [12]. SMI decrease (ΔSMI) was calculated as a percentage by dividing the difference between the SMI at baseline (2008) and follow-up (2016) by the baseline measurement. All measurements were taken in the morning after a night of fasting.

PF tests

PF tests were conducted on the same day as the health checkup. Participants were asked to wear lightweight clothing and flat-soled shoes appropriate for exercise, and physical tests were conducted in the afternoon after a light meal. The tests were supervised by experienced staff from our facility who were trained in the measurement criteria for each test. PF was assessed by the timed up-and-go (TUG) test, the single-leg stance test (SLST), the 30-s chair stand test (30SCST), and handgrip strength.

In the current study, subjects taking the TUG test were asked to stand up from a chair, walk briskly to a mark 3 meters away, walk back, and sit down. The chair had a seat height of 42 cm and did not have armrests. In accordance with the previous protocol [19], a cutoff of 12 s was used to differentiate between elderly individuals living in institutions and those living within the community [21].

The SLST is a reliable and reproducible test of balance [22] that predicts falls and self-reported fractures [23, 24]. For the SLST, we asked subjects to stand on one foot on a hard-surface floor. We then measured the time (seconds) from when the opposite foot left the floor to the time it touched the floor again. We used the best of three trials.

The 30SCST has been validated in a community-dwelling elderly population for measuring lower-extremity strength [25]. The test was designed to overcome the floor effect of the 5-repetition test, in which elderly participants may not be able to complete the task. Participants were asked to sit in a chair that had a 42-cm-high seat and no armrests, to cross their arms in front of their chest, and to stand and sit repeatedly for 30 s; we measured the number of full repetitions.

To measure handgrip strength, the subject used a handheld dynamometer while standing up. Two trials were performed for each hand, and the best value (in kilograms) was used.

PF and BC scoring system

We stratified results for the PF and BC tests using the original scoring system shown in Table 1. In this system, the TUG, 30SCST, and grip strength were scored (0–3 points each) according to cutoffs (mean minus 1, 2, or 3 SD) based on sex-specific *T* scores calculated from younger volunteers (age 20–40 years) from the same health check-ups. For feedback and educational purposes, the cutoff values were rounded to appropriate whole numbers. The

Table 1 Combined PF and BC scoring system for predicting fragility fractures

Points	PF tests				BC values	
	Grip strength[a] (kg)	TUG[a] (s)	30SCST[a]	SLST[b] (s)	SMI limbs	SMI trunk
Men						
0	≥40	≤6	≥20	≥30	≥8.2	≥9.6
1	<40	>6	<20	<30	<8.2	<9.6
2	<30	>7	<13	<20	<7.2	<8.8
3	<20	>8	<8	<10	<6.2	<8.1
Women						
0	≥25	≥6	≥20	≥30	≥6.4	≥8.2
1	<25	>6	<20	<30	<6.4	<8.2
2	<20	>7	<13	<20	<5.7	<7.6
3	<15	>8	<8	<10	<5.1	<7

In this scoring system, each test or measurement is assigned 0–3 points, with more points assigned for worse function. The sum of the PF tests scores is the PF score (0–12 points). The PF + BC score (0–18 points) is the sum of the PF score plus the scores for trunk and limb SMI

30SCST 30-s chair stand test, *BC* body condition, *PF* physical function, *SLST* single-leg stance test, *SMI* skeletal muscle index, *TUG* timed up-and-go test

[a]Cutoff values for grip strength, TUG, and 30SCT were calculated as the mean minus 1, 2, and 3 SD, calculated from gender-specific T scores for healthy volunteers (age 20–40 years) from the same community program. Cutoff values were rounded to the appropriate whole number

[b]Cutoff values of 20, 15, and 10 s were used for the SLST

SLST had a ceiling effect, since participants were stopped at 80 s even if they could remain standing longer, so the SLST was scored (0–3 points) using cutoff values of 10, 20, and 30 s as in previous studies [22, 23]. The total possible PF score was 12 points, with a higher score reflecting worse function. To explore the predictive value of adding BC values to this scoring system, limb and trunk SMI were scored (0–3 points each) according to cutoffs (young adult mean minus 1, 2, or 3 SD) calculated from younger volunteers given the same check-ups. The total possible score for PF + BC was 18 points.

Diagnosis of vertebral fractures and osteoporosis

VFx were evaluated on lateral-view X-rays of the lumbar spine using the semiquantitative method described by Genant et al. [26]. Two examiners, one with 13 years of experience as a spine surgeon (GK) and one with 7 years of experience as an orthopedic surgeon (OT), counted the grade 1 or greater fractures in the T12–L5 vertebrae. The examiners were blinded to the demographic data, such as age and sex, of the subjects. Cohen's κ for inter-rater reliability was 0.897 (95% CI, 0.76–1.04; $P < 0.001$), demonstrating very good agreement between the raters. We compared results between subjects who had sustained a new VFx (VFx group) and those with no new fractures (non-VFx group) at follow-up.

Bone QUS

Bone quality was assessed at baseline by QUS (AOS-100 scanner, Hitachi, Japan) of the left calcaneus (unless that region had been injured). At follow-up, we used the osteosono index (OSI), calculated as (speed of sound)^2 × transmission index.

Statistical analysis

Baseline characteristics were compared using Fisher's exact test for categorical variables and the Mann–Whitney *U* test for non-categorical values. Comparison of age, PS and BC with single vs. multiple VFx was conducted using the Mann–Whitney *U* test. The association of age, PS and BC with the semiquantitative grade of VFx was conducted using the Kruskal–Wallis test with post hoc Steel–Dwass analysis.

The relationship between VFx, PF, and BC was analyzed by logistic regression as in previous studies [27]. We used ROC curves to determine the optimal cutoff values for our proposed PF and BC scoring system. We used a stepwise method with VFx as the dependent variable and age, sex, BMI, OSI, and PF score as independent variables. The BC score was added as an independent variable to analyze the effect of body composition on VFx. A P value < 0.05 was considered statistically significant.

Results

VFx incidence and severity

At follow-up, 36 participants (12%; 12 men and 24 women) had sustained new VFx; of these, 10 had new VFx at multiple levels. The peak incidence of VFx was at T12, with most fractures occurring in the transitional zone from the thoracic to the lumbar vertebrae. By maximum fracture grade, 23 participants were classified as grade 1, nine as grade 2, and four as grade 3 (Fig. 1). The VFx group was significantly older than the non-VFx group (65.8 ± 8.8 vs. 56.5 ± 9.3; $P < 0.001$), but the other demographic characteristics and the prevalence of medical comorbidities were statistically similar (Table 2). The self-reported bodily pain category of the SF-36 did not differ between groups.

Age, PF and BC scores among those with single and multiple VFxs did not differ significantly. Comparison of age, PF, PF + BC scores among the semiquantitative VFx grades showed no difference among groups (Table 3).

Longitudinal associations among VFx, PF, and BC

Baseline PF was worse in the VFx group than the non-VFx group in the SLST (52.0 ± 27.9 vs. 61.8 ± 22.0, $P = 0.02$) and 30SCST (17.5 ± 5.6 vs. 19.5 ± 4.9, $P = 0.03$). Although the baseline BC was similar in the two groups (Table 2), the rate of SMI decline (ΔSMI) was significantly higher in the VFx than the non-VFx group (ΔSMI trunk 3.4 ± 5.9 vs. 1.5 ± 4.5, $P = 0.02$; ΔSMI appendicular 4.6 ± 5.0 vs. 2.5 ± 5.1, $P = 0.01$).

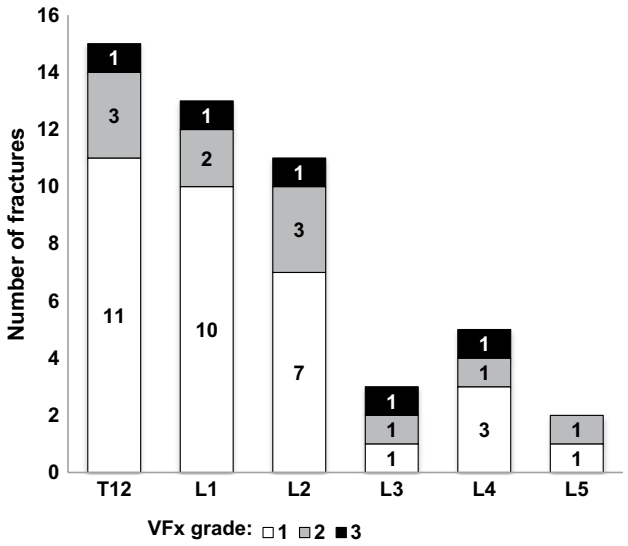

Fig. 1 Incidence and grade of fractures per vertebral level

Table 2 Background data for VFx and non-VFx subjects

	Non-VFx ($n = 271$)	VFx ($n = 36$)	P
Age, years	56.5 ± 9.3	65.8 ± 8.8	< 0.001*
Women, no. (%)	170 (63)	24 (67)	0.72
Height, cm	158.6 ± 8.5	156.8 ± 8.6	0.29
Weight, kg	58.6 ± 10.2	57.8 ± 10.1	0.64
BMI, kg/m^2	23.2 ± 3.1	23.4 ± 2.7	0.79
OSI	2.64 ± 0.37	2.60 ± 0.57	0.19
Existing VFx, no. (%)	12 (4.4)	2 (5.6)	0.67
Osteoporosis treatment (%)	6 (2.2)	3 (8.3)	0.08
Trunk SMI, kg/m^2	9.0 ± 0.9	9.0 ± 0.8	0.53
Leg SMI, kg/m^2	5.5 ± 0.8	5.3 ± 0.7	0.13
Appendicular SMI, kg/m^2	7.2 ± 1.1	6.9 ± 0.9	0.24
Body fat, %	25.6 ± 8.0	27.3 ± 6.5	0.16
ΔSMI trunk	− 1.5 ± 4.5	− 3.4 ± 5.9	0.02*
ΔSMI legs	− 2.1 ± 5.6	− 4.0 ± 5.6	0.06
ΔSMI appendicular	− 2.5 ± 5.1	− 4.8 ± 5.0	0.01*
Grip strength, kg	35.9 ± 11.2	32.5 ± 10.7	0.08
TUG, s	7.0 ± 1.6	7.2 ± 1.7	0.56
SLST, s	61.8 ± 22.0	52.0 ± 27.9	0.02*
30CST	19.5 ± 4.9	17.5 ± 5.6	0.03*
PF score	2.5 ± 1.9	3.5 ± 2.8	0.14
PF + BC score	4.7 ± 2.3	6.1 ± 3.4	0.10

Non-categorical variables are shown as mean ± SD

30CST 30-s chair stand test, *BC* body composition, *BMI* body mass index, *PF* physical function, *OSI* osteo-sono index, *SLST* single-leg stance test, *SMI* skeletal muscle index, *TUG* timed up-and-go test, *VFx* vertebral fracture

*$P < 0.05$. Non-categorical variables were tested by Mann–Whitney *U* test and categorical variables by Fisher's exact test

As shown in Table 4, stepwise logistic regression showed that age and a PF score of 8 or more significantly increased the odds of VFx by 5.6 (95% CI 1.21–25.90, $P = 0.028$. For the combined PF + BC score, we determined a cutoff of 9 to be optimal, and the odds ratio rose to 8.1 (95% CI 1.80–36.00, $P < 0.01$).

Discussion

In this study, we examined the relationship between BC, PF, and radiographic VFx. Over an 8-year period, 36 participants (12%) sustained a new VFx. Baseline SLST and 30SCST values were worse in the VFx group than in the non-VFx group, but the baseline SMI and body fat ratio were comparable in the two groups. We proposed a scoring system for PF and BC to help identify individuals with a high risk of VFx, as shown in Table 1. Using this scoring system, we found that a PF score of 8 or more increased

Table 3 Difference of age, PF, and BC between single or multiple VFX and VFx grade

	Single VFx ($n=26$)	Multiple VFx ($n=10$)	P
Age, years	65.5 ± 9.3	68.5 ± 7.8	0.62
PF score	2.5	3.0	0.53
PF + BC score	3.0 ± 3.1	4.0 ± 3.4	0.52
VFx grade	1	2	3	P
Age, years	64.7 ± 9.0	66.2 ± 9.8	71.5 ± 3.5	0.22
PF score	3.2 ± 2.7	3.4 ± 2.6	5.3 ± 3.9	0.41
PF + BC score	4.3 ± 3.1	4.1 ± 2.9	6.5 ± 3.7	0.40

Mann–Whitney *U* test, Kruskal–Wallis test with post hoc Steel–Dwass analysis

Table 4 Relationship between PF, BC, and VFx

	OR	95% CI	P
Model 1[a]			
Intercept	4.7×10^{-4}	0.00–0.01	< 0.01
Age	1.1	1.1–1.2	< 0.01
PF score ≥ 8	5.6	1.2–25.9	0.028
Model 2[b]			
Intercept	4.7×10^{-4}	2.9×10^{-5}–7.7×10^{-2}	< 0.01
Age	1.1	1.1–1.1	< 0.01
PF + BC score ≥ 9	8.1	1.8–36.0	< 0.01

Results of stepwise logistic regression with VFx as a dependent variable and age, sex, BMI, OSI, and PF score (model 1) or PF + BC score (model 2) as independent variables

BC body composition, *PF* physical function, *VFx* vertebral fractures

[a]Model 1: higher age and a PF score of 8 or more significantly increased the risk of vertebral fractures

[b]Model 2: a PF + BC score of 9 or more significantly increased the risk of vertebral fractures

the odds of VFx by 5.6, and a combined PF + BC score of 9 or more increased the odds ratio to 8.1.

VFx is a hallmark of osteoporosis, and the presence of morphological VFx increases morbidity [5] and lowers QOL [6]. The cumulative incidence of radiological VFx in a Japanese rural population by decade of life (40 s, 50 s, 60 s, and 70 s) was reported to be 2.1%, 8.3%, 10.0%, and 12.2% for men and 2.1%, 6.1%, 18.0%, and 22.4% for women, respectively [28]. The incidence of VFx in our study population was 12%; when stratified by age, the incidence was 4.1%, 6.3%, 15.7%, and 33.3% for the entire population in their 40 s, 50 s, 60 s, and 70 s or above, respectively. The incidence of fractures in women by age group was 2.4%, 4.2%, 20% and 33.3% in their 40 s, 50 s, 60 s and 70 or above, respectively. This was comparable to previous studies conducted in Japan. Our study did not have an adequate number of male VFx to compare to previous studies of radiographic VFx [29–33].

Efforts to predict fragility fractures have focused on medical comorbidities and BMD, and have had some success. However, these screening methods must be calibrated for different populations [3, 34]. Moreover, after testing the performance of three predictive strategies (FRAX, fracture risk assessment tool; SCORE, simple calculated osteoporosis risk estimate; and OST, osteoporosis self-assessment tool), Crandall et al. [3] reported that these methods may not be able to predict the majority of osteoporotic fractures in women aged 50–64. After following women aged 70 or older over a 10-year period, Ensrud et al. [35] reported that the predictive strength of FRAX was similar to that of simpler tools. Rubin et al. [36] reported that FRAX and even simpler tools were not superior to age alone in predicting fractures in women aged 45–90 over a 3-year period. These results suggest that risk factors for osteoporosis are of limited value for predicting fractures, and that the focus should be broadened to include risk factors for falls in both middle-aged and elderly women.

BC is associated with bone strength and fractures [8]. Hida et al. [11] used whole-body DXA to measure muscle volume in 216 women with acute osteoporotic VFx diagnosed by MRI, and found that the appendicular SMI was lower and the prevalence of sarcopenia was higher in women with acute VFx compared to healthy control subjects. Walsh et al. [37] reported that a positive association between SMI and BMD disappeared when corrected for physical activity. The VFx and non-VFx groups in our study had a similar baseline SMI and prevalence of sarcopenia, but the rate of muscle decline was greater in the VFx group. This result is consistent with studies showing lower muscle volume in patients diagnosed with fragility fractures [11, 38, 39]. Because the baseline muscle volume was similar in the VFx and non-VFx groups, the loss of muscle volume may not be not a cause, but rather the result of a common confounding factor for bone strength and muscle volume, such as aging, a sedentary lifestyle, or comorbidities. Although a causal relationship between BC and fractures remains to be clarified, our study showed that muscle mass is unlikely to be an

independent predictor, and thus a cause, of VFx. Sarcopenia, as defined by the Asia Working Group for Sarcopenia (AWGS), was less prevalent in our study population (12.5%) than that reported elsewhere (20%) [12], possibly because our population included younger individuals or because the study was conducted in a region where the main industry is agriculture, and most workers are self-employed and continue to work into their later years.

In this study, we used a scoring system (PF score) to assess physical function based on the TUG test, SLST, 30SCST, and grip strength (Table 1). A PF score of 8 or more increased the odds of VFx by 5.6, which was significant after correcting for age, sex, BMI, and OSI. The relationship between poor PF and susceptibility to falls and subsequent fractures has been studied. Poor PF is hypothesized to increase the incidence of falls, which in turn increases the number of fractures. A cross-sectional study by Prato et al. [13] showed that grip strength in the first quartile increased the odds ratio of self-reported falls by 2.3. Lee et al. [40] found that the incidence of VFx was higher in women with a lower jump strength, but was not associated with SMI. These reports are consistent with our current results, which showed that taken separately, poor PF but not BC increased the risk of VFx. Longitudinal studies have demonstrated associations between PF measurements and self-reported fractures in elderly individuals living in the general community [14, 24, 27, 41]. Although the method of diagnosing fractures differed, these studies concluded that a combination of PF tests, rather than a single measurement, is best for predicting osteoporotic fractures. These results agree with our findings that a combination of PF tests, rather than a single measurement, is required for predicting VFx. Different tests measure different abilities, and a decline in several areas may be necessary to increase the risk of fracture.

One advantage of our study is that our risk stratification includes both PF and BC. Previous studies have reported the risk of these factors independently, but few have considered both factors for predicting radiographic vertebral fractures. Our results showed that the addition of BC to PF was cumulative in predicting fractures. The PF scoring system used in this study may reflect small differences better than a categorical definition of poor function (i.e., one cutoff value), and this may also be an advantage when tracking the value of interventions such as physical training and exercise. Our population was younger and more physically fit than populations in previous studies, for example, none of our subjects met the criteria for poor PF described by Chun et al. [24]. We believe that populations that are growing older but have a relatively low degree of functional decline, like our study population, can benefit the most from fracture prevention strategies. It is important to screen for and treat high-risk individuals before PF is severely impaired, when there is greater potential to conserve function and thereby conserve medical resources. Additional studies are warranted to confirm these results and determine whether targeted improvement of the physical measures used in our PF scoring system can prevent fractures.

Our study had several limitations. We excluded volunteers with incomplete data, which may have excluded volunteers whose condition was too poor to complete the PF tests. Thus, our data may be representative of a relatively fit population. Second, the study was conducted in a rural community with a relatively high percentage of elderly people compared to urban populations in Japan, and our results may not be representative of Japan as a whole.

Third, DXA measurements were not available in this study. DXA is widely used for the diagnosis and screening of osteoporosis and is a valuable factor for predicting fractures. However, there are screening tools that do not utilize DXA that have shown some success [3]. Because our physical function tests do not require any special equipment, we believe our scoring system may be useful for screening purposes where DXA is unavailable.

The fourth limitation was the unavailability of thoracic X-rays. Incident vertebral fractures most frequently occur in the thoracolumbar region, but also occur in the upper thorax [29, 42]. Future studies should include the thoracic spine to validate our results.

Finally, our study did not consider symptoms such as knee or back pain. Knee pain is associated with falls in women [43], and knee pain and osteoarthritis are associated with osteoporotic fractures [44]. Pain is a possible feature and confounding factor in BC, PF, and VFx, and future studies should encompass all of these factors for a more complete understanding of the factors that increase the risk of VFx.

Our study also had several advantages. We used radiographic VFx as the outcome, in contrast to other studies based on self-reported fractures. Although the presence of subclinical radiographic VFx may seem insignificant, these fractures increase morbidity and decrease QOL [5, 6]. Second, our study included younger participants, who are ideal targets for preventative intervention.

In conclusion, the incidence of VFx over an 8-year period was 12% in our study population. Baseline SMI was not a factor in the risk of future VFx, but the rate of decline in muscle volume was greater in those who sustained VFx. In our proposed PF scoring system, which assigns a point value based on results from the TUG test, 30SCST, SLST, and grip strength, a cutoff of 8 points identified an increased risk of future VFx. The OR of this point system was increased by adding points for lower SMI in the trunk and limbs.

Acknowledgements This work was supported by a Grant-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology

of Japan (No. 18200044), by the Japanese Society for the Promotion of Science (No. 21500676), by a Health and Labour Sciences Research Grant, and by JOA-Subsidized Science Project Research from the Japanese Orthopaedic Association (No. 2015-02).

Compliance with ethical standards

Conflict of interest The authors have no conflicts of interests to declare.

Ethical approval This study was approved by the Ethics Committee of Hirosaki University Graduate School of Medicine and was performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. Each individual participant gave informed consent before participating.

References

1. Cummings SR, Melton LJ (2002) Epidemiology and outcomes of osteoporotic fractures. Lancet 359:1761–1767
2. Hernlund E, Svedbom A, Ivergård M, Compston J, Cooper C, Stenmark J, McCloskey EV, Jönsson B, Kanis JA (2013) Osteoporosis in the European Union: medical management, epidemiology and economic burde. A report prepared in collaboration with the International Osteoporosis Foundation (IOF) and the European Federation of Pharmaceutical Industry Associations (EFPIA). Arch Osteoporos 8:136
3. Crandall CJ, Larson JC, Watts NB, Gourlay ML, Donaldson MG, LaCroix A, Cauley JA, Wactawski-Wende J, Gass ML, Robbins JA, Ensrud KE (2014) Comparison of fracture risk prediction by the US Preventive Services Task Force strategy and two alternative strategies in women 50–64 years old in the Women's Health Initiative. J Clin Endocrinol Metab 99:4514–4522
4. Sandhu SK, Nguyen ND, Center JR, Pocock NA, Eisman JA, Nguyen TV (2010) Prognosis of fracture: evaluation of predictive accuracy of the FRAX algorithm and Garvan nomogram. Osteoporos Int 21:863–871
5. Kanis JA, Johnell O, Oden A, Borgstrom F, Zethraeus N, De Laet C, Jonsson B (2004) The risk and burden of vertebral fractures in Sweden. Osteoporos Int 15:20–26
6. Cockerill W, Lunt M, Silman AJ, Cooper C, Lips P, Bhalla AK, Cannata JB, Eastell R, Felsenberg D, Gennari C, Johnell O (2004) Health-related quality of life and radiographic vertebral fracture. Osteoporos Int 15:113–119
7. Lindsay R, Silverman SL, Cooper C, Hanley DA, Barton I, Broy SB, Licata A, Benhamou L, Geusens P, Flowers K, Stracke H (2001) Risk of new vertebral fracture in the year following a fracture. JAMA 285:320–323
8. Rolland Y, Van Kan GA, Benetos A, Blain H, Bonnefoy M, Chassagne P, Jeandel C, Laroche M, Nourhashemi F, Orcel P, Piette F (2008) Frailty, osteoporosis and hip fracture: causes, consequences and therapeutic perspectives. J Nutr Health Aging 12:335–346
9. Scott D, Seibel M, Cumming R, Naganathan V, Blyth F, Le Couteur DG, Handelsman DJ, Waite LM, Hirani V (2017) Sarcopenic obesity and its temporal associations with changes in bone mineral density, incident falls, and fractures in older men: the concord health and ageing in men project. J Bone Miner Res. 32:575–583
10. Mokhtarzadeh H, Anderson DE (2016) The role of trunk musculature in osteoporotic vertebral fractures: implications for prediction, prevention, and management. Curr Osteoporos Rep 14:67–76
11. Hida T, Shimokata H, Sakai Y, Ito S, Matsui Y, Takemura M, Kasai T, Ishiguro N, Harada A (2016) Sarcopenia and sarcopenic leg as potential risk factors for acute osteoporotic vertebral fracture among older women. Eur Spine J 25:3424–3431
12. Chen LK, Liu LK, Woo J, Assantachai P, Auyeung TW, Bahyah KS, Chou MY, Chen LY, Hsu PS, Krairit O, Lee JS (2014) Sarcopenia in Asia: consensus report of the Asian Working Group for Sarcopenia. J Am Med Dir Assoc 15:95–101
13. Prato SCF, Andrade SM, Cabrera MAS, Dip RM, Santos HGD, Dellaroza MSG, Mesas AE (2017) Frequency and factors associated with falls in adults aged 55 years or more. Rev Saude Publica 51:37
14. Stel VS, Pluijm SM, Deeg DJ, Smit JH, Bouter LM, Lips P (2004) Functional limitations and poor physical performance as independent risk factors for self-reported fractures in older persons. Osteoporos Int 15:742–750
15. Watson SL, Weeks BK, Weis LJ, Harding AT, Horan SA, Beck BR (2017) High-intensity resistance and impact training improves bone mineral density and physical function in postmenopausal women with osteopenia and osteoporosis: the LIFTMOR randomized controlled trial. J Bone Miner Res 33:211–220
16. Chiba D, Tsuda E, Wada K, Kumagai G, Sasaki E, Nawata A, Nakagomi S, Takahashi I, Nakaji S, Ishibashi Y (2016) Lumbar spondylosis, lumbar spinal stenosis, knee pain, back muscle strength are associated with the locomotive syndrome: Rural population study in Japan. J Orthop Sci 21:366–372
17. Kumagai G, Wada K, Tanaka T, Kudo H, Asari T, Chiba D, Ota S, Nakaji S, Ishibashi Y (2018) Associations between neck symptoms and LDL cholesterol in a cross-sectional population-based study. J Orthop Sci 23:277–281
18. Shafer KJ, Siders WA, Johnson LK, Lukaski HC (2009) Validity of segmental multiple-frequency bioelectrical impedance analysis to estimate body composition of adults across a range of body mass indexes. Nutrition 25:25–32
19. Bosy-Westphal A, Later W, Hitze B, Sato T, Kossel E, Gluer CC, Heller M, Müller MJ (2008) Accuracy of bioelectrical impedance consumer devices for measurement of body composition in comparison to whole body magnetic resonance imaging and dual X-ray absorptiometry. Obes Facts 1:319–324
20. Dehghan M, Merchant AT (2008) Is bioelectrical impedance accurate for use in large epidemiological studies? Nutr J 7:26
21. Bischoff HA, Stähelin HB, Monsch AU, Iversen MD, Weyh A, Von Dechend M, Akos R, Conzelmann M, Dick W, Theiler R (2003) Identifying a cut-off point for normal mobility: a comparison of the timed 'up and go' test in community-dwelling and institutionalised elderly women. Age Ageing 32:315–320
22. Giorgetti MM, Harris BA, Jette A (1998) Reliability of clinical balance outcome measures in the elderly. Physiother Res Int 3:274–283
23. Chomiak T, Pereira FV, Hu B (2015) The single-leg-stance test in Parkinson's disease. J Clin Med Res 7:182–185
24. Chun SH, Cho B, Yang HK, Ahn E, Han MK, Oh B, Shin DW, Son KY (2017) Performance on physical function tests and the risk of fractures and admissions: findings from a national health screening of 557,648 community-dwelling older adults. Arch Gerontol Geriatr 68:174–180
25. Jones CJ, Rikli RE, Beam WC (1999) A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. Res Q Exerc Sport 70:113–119
26. Genant HK, Wu CY, van Kuijk C, Nevitt MC (1993) Vertebral fracture assessment using a semiquantitative technique. J Bone Miner Res 8:1137–1148
27. Cawthon PM, Blackwell TL, Marshall LM, Fink HA, Kado DM, Ensrud KE, Cauley JA, Black D, Orwoll ES, Cummings SR, Schousboe JT (2014) Physical performance and radiographic

and clinical vertebral fractures in older men. J Bone Miner Res 29:2101–2108
28. Yoshimura N, Kinoshita H, Takijiri T, Oka H, Muraki S, Mabuchi A, Kawaguchi H, Nakamura K, Nakamura T (2008) Association between height loss and bone loss, cumulative incidence of vertebral fractures and future quality of life: the Miyama study. Osteoporos Int 19:21–28
29. Ross PD, Fujiwara S, Huang C, Davis JW, Epstein RS, Wasnich RD, Kodama K, Melton LJ III (1995) Vertebral fracture prevalence in women in Hiroshima compared to Caucasians or Japanese in the US. Int J Epidemiol 24:1171–1177
30. Fujiwara S, Kasagi F, Masunari N, Naito K, Suzuki G, Fukunaga M (2003) Fracture prediction from bone mineral density in Japanese men and women. J Bone Miner Res 18:1547–1553
31. Lopes JB, Danilevicius CF, Takayama L, Caparbo VF, Menezes PR, Scazufca M, Kuroishi ME, Pereira RMR (2011) Prevalence and risk factors of radiographic vertebral fracture in Brazilian community-dwelling elderly. Osteoporos Int 22:711–719
32. Baddoura R, Arabi A, Haddad-Zebouni S, Khoury N, Salamoun M, Ayoub G, Okais J, Awada H, Fuleihan GEH (2007) Vertebral fracture risk and impact of database selection on identifying elderly Lebanese with osteoporosis. Bone 40:1066–1072
33. Ballane G, Cauley JA, Luckey MM, El-Hajj Fuleihan G (2017) Worldwide prevalence and incidence of osteoporotic vertebral fractures. Osteoporos Int 28:1531–1542
34. Fraser LA, Langsetmo L, Berger C, Ioannidis G, Goltzman D, Adachi JD, Papaioannou A, Josse R, Kovacs CS, Olszynski WP, Towheed T (2011) Fracture prediction and calibration of a Canadian FRAX® tool: a population-based report from CaMos. Osteoporos Int 22:829–837
35. Ensrud KE, Lui LY, Taylor BC, Schousboe JT, Donaldson MG, Fink HA, Cauley JA, Hillier TA, Browner WS, Cummings SR (2009) A comparison of prediction models for fractures in older women: is more better? Arch Intern Med 169:2087–2094
36. Rubin KH, Abrahamsen B, Friis-Holmberg T, Hjelmborg JV, Bech M, Hermann AP, Barkmann R, Glüer CC, Brixen K (2013) Comparison of different screening tools (FRAX®, OST, ORAI, OSIRIS, SCORE and age alone) to identify women with increased risk of fracture: a population-based prospective study. Bone 56:16–22
37. Walsh MC, Hunter GR, Livingstone MB (2006) Sarcopenia in premenopausal and postmenopausal women with osteopenia, osteoporosis and normal bone mineral density. Osteoporos Int 17:61–67
38. Blain H, Vuillemin A, Teissier A, Hanesse B, Guillemin F, Jeandel C (2001) Influence of muscle strength and body weight and composition on regional bone mineral density in healthy women aged 60 years and over. Gerontology 47:207–212
39. Segal NA, Torner JC, Yang M, Curtis JR, Felson DT, Nevitt MC (2008) Muscle mass is more strongly related to hip bone mineral density than is quadriceps strength or lower activity level in adults over age 50 year. J Clin Densitom 11:503–510
40. Lee EY, Lee SJ, Kim KM, Seo DH, Lee SW, Choi HS, Kim HC, Youm Y, Kim CO, Rhee Y (2017) Lower jump power rather than muscle mass itself is associated with vertebral fracture in community-dwelling elderly Korean women. Calcif Tissue Int 100:585–594
41. Wihlborg A, Englund M, Åkesson K, Gerdhem P (2015) Fracture predictive ability of physical performance tests and history of falls in elderly women: a 10-year prospective study. Osteoporos Int 26:2101–2109
42. van der Klift M, Pols HA, Hak AE, Witteman JC, Hofman A, de Laet CE (2002) Bone mineral density and the risk of peripheral arterial disease: the Rotterdam Study. Calcif Tissue Int 70:443–449
43. Muraki S, Akune T, Ishimoto Y, Nagata K, Yoshida M, Tanaka S, Oka H, Kawaguchi H, Nakamura K, Yoshimura N (2013) Risk factors for falls in a longitudinal population-based cohort study of Japanese men and women: the ROAD Study. Bone 52:516–523
44. Arden NK, Crozier S, Smith H, Anderson F, Edwards C, Raphael H, Cooper C (2006) Knee pain, knee osteoarthritis, and the risk of fracture. Arthritis Rheum 55:610–615