Predicting Ability of Modified Short Physical Performance Battery

Kazuki Fukui, Noriaki Maeda, Junpei Sasada, Shogo Sakai, Tsubasa Tashiro, Toshiya Shima, Manabu Niitani, Yukio Urabe

**1. Introduction**

Frailty is a clinical state characterized by the increased vulnerability of an individual to stressors caused by a cumulative decline in multiple physiological systems. It is a multidimensional phenomenon that includes experiencing losses in physical, social and cognitive capabilities. However, in principle, it is a pre-disability syndrome and a reversible condition. It can be classified into three categories (non-frailty, pre-frailty and frailty) by several assessment tools. The most easily identifiable feature of frailty is considered to be low gait, along with a reduction in the ability to walk. Otherwise, there are remarkably few clinically reliable markers of pre-frailty, which of course is a significant impediment to developing an effective prevention strategy for reversing this condition.

In Japan, recent research has shown that the Kihon Checklist is an effective screening tool for estimating frailty. Various studies based on this instrument have demonstrated that individuals it determined to be frail had a 3-year risk of death that was 3.5 times higher than those considered non-frail by this metric. Other reports have indicated that frailty according to the Kihon Checklist could help predict the risk of adverse health outcomes, such as falls, disabilities, fractures, hospitalization, institutionalization and mortality in community-dwelling elderly individuals. However, while it is important to assess the status of elderly people for the prevention and early detection of frailty, the Kihon Checklist is the assessment tool composed of the only questionnaire, and unfortunately, it cannot assess the actual physical function capability like gait speed, muscle strength, and balance ability.

In contrast, the Short Physical Performance Battery (SPPB) developed by Guralnik et al., predicts adverse health events in older adults and the validity and reliability of the SPPB in the previous study demonstrated in older adults in several countries. In addition, the SPPB has been shown to detect early stages of frailty, even among high-functioning older adults. Compared with the Kihon Checklist, the SPPB is a much broader evaluation of physical function. It can be quickly executed using simple equipment, and it is sensitive to changes in functionality though time. However, it is not widely used in Japan, mainly because it is too easy for Japanese elderly to achieve a perfect score, thereby reducing its value as a sensitive indicator of physical performance capability. Therefore, it was necessary to modify this tool for high-functioning elderly adults, for example Japanese.

This project was accomplished by Makizako et al., who used a community-based scoring system to create a modified test, now called the SPPB-com. This assessment can readily measure physical functioning for Japanese elderly because it is mostly the same test as the SPPB, and mostly differs in how it is scored. Oddly, although the...
SPPB-com has the potential for predicting the risk of disability in Japanese community-dwelling elderly people, apparently, no research has yet been undertaken to define the criteria for 'frailty' regarding this specific population.

Therefore, this study aimed to investigate the capacity of the SPPB-com to distinguish across the categories of non-frailty, pre-frailty and frailty, classified by a total-Kihon Checklist score. Based on previous research, it was hypothesized that SPPB-com can accurately generate these three classifications and, therefore, be quite useful as a screening tool for determining which high-functioning elderly people might be most at risk of frailty.

2. Material and methods

2.1. Study design and subjects

The participants in this study were 132 elderly outpatients aged ≥ 65 years, who were registered at the orthopedic clinic in Hiroshima, Japan, from February 2018 to February 2019 and who had been visiting this clinic for > 2 months at least once a week. We excluded anyone who was not able to answer a detailed questionnaire or walk unaided (Figure 1). This study was reviewed and approved by the institutional review board (IRB) of the Niitani Clinic (NCL-18001).

2.2. Metrics

2.2.1. Sample characteristics

We recorded numerous aspects of our participants, including their age, sex, height, weight, Body Mass Index, family structure (living alone or not), history of falls in the past 12 months and their level of physical activity, as determined by the self-administered Physical Activity Questionnaire for Elderly Japanese (PAQ-EJ). This instrument is specifically designed to measure the typical activity patterns of daily life. We then converted the PAQ-EJ scores into metabolic equivalents of task (MET) hours per week (MET h/week). These assessments were measured in a self-report.

2.2.2. Measuring risk exposures

Frailty was assessed using the Kihon Checklist, which is a 25-item self-administered questionnaire developed by the Japanese Ministry of Health, Labour and Welfare to identify frail older individuals who are at risk of requiring new certification for long-term care insurance. The Kihon Checklist is a comprehensive evaluation method that focuses on social and psychological aspects in addition to the physical dimensions of frailty. Given its ability to assess infirmity across multiple domains, the Kihon Checklist is widely regarded as an effective screening tool, and its validity and reliability have been formally confirmed previously. The questions on the Kihon Checklist require a simple ‘yes’ or ‘no’ answer and are scored as 1 or 0 points, respectively. A score of ≥ 8 points is considered to indicate frailty, whereas a score of 4–7 points indicates pre-frailty, and < 3 points indicates non-frailty. In addition, the Kihon Checklist includes an assessment of cognitive function, which is largely based on three specific questions, and a yes answer to any of them is considered to indicate cognitive function decline.

2.2.3. Outcome measures - muscle strength

Handgrip strength and maximum isometric knee extension strength were measured as indicators of muscle strength by 3 physical therapy. Handgrip strength was tested twice with the dominant hand using a digital dynamometer (T.K.K.5401, Takei Kiki Kogyo, Japan), and the best of these two test results was used in the current analysis.

The maximum isometric knee extensor strength was assessed using a hand-held dynamometer (Mobie, Sakai Med Co., Tokyo, Japan) with the participants seated on a treatment table with their knees and hips at 90° flexion. The average of the highest of two recordings (two from each side) for each participant was used.

The maximum walking speed was measured over a 10-m distance. The participants were instructed to walk as fast as possible 1 m before the starting line and to continue at this pace until 1 m after the finishing line. The timing commenced when the participant’s swing leg crossed the starting line and ended when the swing leg passed over the finishing line. Two trials were performed, and the shortest time was used for calculations.

2.2.4. Physical performance

Both the SPPB (0–12 points) and the SPPB-com, which is based on a community-based score (0–10 points) were used to measure physical performance. The SPPB consists of 3 assessments: 1) a balance test, 2) a 4-m walk test, and 3) a repeated chair stand test. Each SPPB component (balance, walk and chair stand) is scored from 0 to 4, with a score of 0 representing an inability to carry out the test, and 4 reflecting the best performance. For the balance test, the participants were asked to keep their feet in a side-by-side, then in a semi-tandem, and finally in a fully tandem position, each for 10 s. For the 4-m walk test, participants walked with their usual speed. For the repeated chair stand test, subjects were asked to stand up and sit down five times as quickly as possible. This task was only conducted one time.

The SPPB-com test is the same as the SPPB, but it is scored differently, as shown in Table 1. For the balance test, participants who cannot keep a tandem position receive 0 points, those who can maintain this position for 1 to 10 s get 1 point and holding it for longer earns 2 points. Other tests (4-m walk test and chair stands test) can earn between 0 and 4 points, and the highest possible overall score is 10 points.

2.3. Statistical analysis

Prior to the analysis, we classified participants into three groups classified by their Kihon Checklist score: non-frailty, pre-frailty and frailty. Continuous variables are presented as means ± standard deviations and all data were normally distributed. For the 3-group analysis of variance, the Bonferroni post hoc test was used. Differences in the prevalence of cognitive decline, living alone and experience with falls over the past 12 months among the three groups

---

**Figure 1.** Participant recruitment flow diagram.
were evaluated with the Chi-square test. We also performed a discriminant analysis using the stepwise forward procedure to identify the three frailty groups. A receiver operating characteristics (ROC) of the factors selected by discriminant analysis was employed to calculate the area under the curve (AUC) and the cut-off point, to better distinguish the non-frailty from the frailty respondents. Statistical analyses were carried out using SPSS Statistics for Windows, version 25.0 (IBM SPSS Inc., Tokyo, Japan), with a significance threshold of 0.05.

3. Results

Demographic data for participants stratified by frailty group is shown in Table 2. There were 28 participants (21.2%) in the non-frailty group, 41 (31.1%) in the pre-frailty and 63 (47.7%) in the frailty group. Our analysis of variance showed that across these three categories, there were significant differences in age, handgrip strength, maximum walking speed, knee extensor strength and in the scores on the PAQ-EJ, SPPB and the SPPB-com instruments. In addition, the SPPB-com score was able to reveal a significant difference between the pre-frailty and frailty groups. In the Chi-square test, there were material differences in the prevalence of cognitive decline, living alone and the number of falls (Table 2).

The results of the discriminant analyses are shown in Table 3. This procedure was conducted to determine which evaluation instrument can best discriminate among the three frailty groups, and it became clear that the SPPB-com score was found to be the most effective in this regard. Indeed, it was able to classify frailty groups into the three categories with an accuracy of 62.5% following validation (Table 3).

Table 3 presents the ROC analysis aimed at predicting frailty and pre-frailty. In the frailty, the SPPB score with a cut-off-point of 11 gave the best trade-off between sensitivity and specificity, with an AUC = 0.77. In contrast, the SPPB-com score with a cut-off point of 5 gave the best trade-off between sensitivity and specificity, with an

<table>
<thead>
<tr>
<th>Variable</th>
<th>Normalized discriminant function</th>
<th>SPPB-com (score)</th>
<th>0.625*</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPPB-com: short physical performance battery-community based.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>^* Step length parameter with higher discriminant value between groups.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4 Sensitivity and specificity of Short Physical Performance Battery and Short Physical Performance Battery-community based identifying frailty and pre-frailty in participants.

<table>
<thead>
<tr>
<th>Cut-off point</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>AUC (CI 95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frailty SPPB 10/11 79 73 0.77 (0.68–0.85)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPPB-com 4/5 67 73 0.80 (0.73–0.87)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-frailty SPPB 11/12 82 65 0.71 (0.58–0.84)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPPB-com 5/6 71 69 0.76 (0.64–0.87)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2 Demographic differences according to Kihon Checklist scores.

<table>
<thead>
<tr>
<th>Total (n = 132)</th>
<th>Non-frailty (n = 28)</th>
<th>Pre-frailty (n = 41)</th>
<th>Frailty (n = 63)</th>
<th>p for trend</th>
<th>Post-hoc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td>79.4±7.0</td>
<td>76.0±7.2</td>
<td>78.5±5.8</td>
<td>81.4±7.0</td>
<td>0.002</td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td>20.9±7.2</td>
<td>23.2±7.4</td>
<td>22.2±6.8</td>
<td>19.1±7.1</td>
<td>0.691</td>
</tr>
<tr>
<td>Grip strength (kg)</td>
<td>21.6±7.1</td>
<td>23.4±7.3</td>
<td>23.1±6.7</td>
<td>19.6±7.0</td>
<td>0.016</td>
</tr>
<tr>
<td>Knee extensor strength (N/kg)</td>
<td>3.88±1.25</td>
<td>4.73±1.12</td>
<td>4.15±1.19</td>
<td>3.31±1.07</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Maximum walking speed (m/s)</td>
<td>1.23±0.38</td>
<td>1.55±0.33</td>
<td>1.29±0.32</td>
<td>1.05±0.34</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>PAQ-EJ (MET hours/week)</td>
<td>49.9±47.5</td>
<td>74.8±41.2</td>
<td>57.6±60.7</td>
<td>33.8±32.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>SPPB (score)</td>
<td>9.87±2.43</td>
<td>11.5±1.3</td>
<td>10.7±1.7</td>
<td>8.65±2.62</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>SPPB-com (score)</td>
<td>4.61±1.61</td>
<td>6.29±1.41</td>
<td>4.88±1.42</td>
<td>3.70±1.10</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sex (male)</td>
<td>36 (27.3%)</td>
<td>7 (25.0%)</td>
<td>11 (26.8%)</td>
<td>18 (28.6%)</td>
<td>0.937</td>
</tr>
<tr>
<td>Cognitive decline (n)</td>
<td>69 (52.9%)</td>
<td>6 (19.4%)</td>
<td>15 (31.9%)</td>
<td>48 (76.2%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Living alone (n)</td>
<td>37 (28.0%)</td>
<td>1 (4.0%)</td>
<td>14 (34.1%)</td>
<td>22 (34.9%)</td>
<td>0.005</td>
</tr>
<tr>
<td>Experience of fall (n)</td>
<td>52 (39.4%)</td>
<td>6 (21.4%)</td>
<td>13 (31.7%)</td>
<td>33 (52.4%)</td>
<td>0.010</td>
</tr>
</tbody>
</table>


a. Significant difference between frailty and non-frailty (p < 0.01).
b. Significant difference between frailty and pre-frailty (p < 0.01).
c. Significant difference between non-frailty and pre-frailty (p < 0.01).
Efficacy of Modified SPPB for Frailty Detection 215

AUC = 0.80. On the other hand, in the pre-frailty, the SPPB score with a cut-off point of 12 gave the best trade-off between sensitivity and specificity, with an AUC = 0.71, while an SPPB-com score with a cut-off point of 6 gave the best compromise between sensitivity and specificity, with an AUC = 0.76. The AUC results of the SPPB-com score were better than the SPPB scores for both frailty and pre-frailty (Figure 2).

4. Discussion

This study investigated whether the SPPB-com was a tool that can effectively determine the difference between non-frailty, pre-frailty and frailty in our group of community-dwelling elderly outpatients. The data we generated showed that the SPPB-com was the more valid discriminant tool compared with SPPB for the purpose of identifying pre-frailty in elderly people at an orthopedic clinic. More generally, after performing the discriminant analysis between the three groups, we found that the SPPB-com score was able to effectively identify the differences across all our three categories of frailty.

In a previous study of healthier and relatively high-functioning older adults with normal mobility, the SPPB was also able to detect early signs of frailty even before a decrease in normal walking speed occurred.20 Our frailty group had reduced muscle strength, walking speed and physical activity compared with the non-frailty group, and similar outcomes have been previously reported although there were normative difference value on sex because on the same rate of males and females in each group.21 However, the results from our present study showed that the walking speed of the pre-frailty subjects had already decreased, but the SPPB score could only distinguish between the frailty and the non-frailty, because it could not reflect changes in walking ability for the pre-frailty. Thus, for our sample, the SPPB could find frailty but could not find pre-frailty, and it would therefore be impossible to identify the need for early intervention.

In contrast, the SPPB-com score could generate a significant difference between our three frailty groups. The SPPB-com has been adjusted for Japanese community-dwelling elderly people by Makizako et al.,14 and even for this population, achieving a perfect score on each of the component tests is now very difficult. In particular, it was previously too easy to get a perfect score of 4 on the 4-m walk test on the SPPB because the average walking speed for Japanese older adults is 1.2–1.3 m/s and this constitutes a greater than perfect score on this component of the SPPB.21,22 Also, our results showed that the pre-frailty individuals demonstrated decreased walking speed compared with non-frailty individuals, and the SPPB-com score showed the same results, suggesting that it can determine the difference between non-frailty and pre-frailty individuals.

Even better, Verghese and Xue reported that with a cut-off of 8 points, the SPPB had 52% sensitivity and 70% specificity in frailty identification in elderly people with good functionality.12 In a longitudinal study, Vasunilashorn et al. found that a cut-off of 10 points was an important predictor of mobility loss (inability to walk 400 m), with 69% sensitivity and 84% specificity, for a population aged ≥65 years.24 However, our present investigation showed that a SPPB score of 11 could discriminate frailty from non-frailty, and a value of 12 could distinguish between non-frailty and pre-frailty.

Some researchers have reported that physical performance, as measured by the SPPB, is a credible predictor of disability among older adults.25 On the other hand, the participants in those studies were not high function elderly people. In contrast, a 2017 study revealed a mean SPPB score for Japanese community-dwelling older people of 11.6, and 80% could get a perfect score.14 The participants in the present study were elderly Japanese people at an orthopedic clinic, and most of them were frailty or pre-frailty according to the total Kihon Checklist score. They had a mean SPPB score of 9.8, i.e., less than 11. These results suggest that the SPPB could discriminate between the extremes of non-frailty and frailty but could not assess the possible effects of an intervention for pre-frailty.

On the other hand, a value of 5 on the SPPB-com could discriminate frailty from non-frailty, and a value of 6 could distinguish between non-frailty and pre-frailty. These results showed that the SPPB-com is a better tool for discriminating among the non-frailty, pre-frailty and frailty compared with the SPPB. We found that a one point increase in the SPPB-com score reduced the risk of needing long-term care by a remarkable 23%.15 So, a difference of only one point in the cut-off score can distinguish frailty or pre-frailty compared with being non-frailty, in the context of the general decline of physical functionality of elderly people. Because the SPPB has a ceiling effect, due to high-functioning elderly people often achieving a perfect score, it offers little assessment capability for this population. In contrast, with the SPPB-com, only 10% can readily achieve a perfect score, even in populations of Japanese community-dwelling older adults.14 Therefore, because it eliminates the ceiling effect, the SPPB-com is a superior instrument for assessing the impact of intervention not only for Japanese community-dwelling elderly people but also for high-functioning older adults throughout the world.

4.1. Strengths and limitations

This study is the first to describe the discrimination capability of the new SPPB-com criteria for frailty, classified by the Kihon Checklist score. In addition, the muscle strength and physical performance of
the pre-frailty in orthopedic clinic patients were revealed from both outcome measures and the score on the SPPB-com.

A limitation of this study is that our SPPB scores were higher than in previous studies because our subjects were only Japanese elderly people at an orthopedic clinic. A second weakness is that we cannot calculate the power and utilized a relatively small sample of subjects. and they were all from an orthopedic clinic, with over half rated as frailty.

**Conflict of interest**

The authors declare that they have no conflict of interest.

**References**