



Reliability and validity of the Japanese version of the Barthel Index dyspnea among patients with respiratory diseases

Yamaguchi, Takumi

(Degree)

博士 (保健学)

(Date of Degree)

2021-09-25

(Date of Publication)

2022-09-01

(Resource Type)

doctoral thesis

(Report Number)

甲第8160号

(URL)

<https://hdl.handle.net/20.500.14094/D1008160>

※ 当コンテンツは神戸大学の学術成果です。無断複製・不正使用等を禁じます。著作権法で認められている範囲内で、適切にご利用ください。



博士論文

Reliability and validity of the Japanese version of the Barthel Index
dyspnea among patients with respiratory diseases
(呼吸器疾患患者におけるBarthel Index dyspnea日本語版の信頼性
、妥当性の検証)

令和3年7月6日

神戸大学大学院保健学研究科保健学専攻
Takumi Yamaguchi 山口 卓巳

Reliability and validity of the Japanese version of the Barthel Index dyspnea among patients with respiratory diseases

Takumi Yamaguchi ^{1, 2*}, Akio Yamamoto ^{2, 3}, Yutaro Oki ², Hideki Sakai ¹, Shogo Misu ^{1, 4}, Yusuke Iwata ², Masahiro Kaneko ⁵, Kaku Sawada ⁶, Yukari Oki ², Yuji Mitani ², Kumiko Ono ², Akira Ishikawa ²

1 Department of Rehabilitation, Kobe City Hospital Organization, Kobe City Medical Center West Hospital, 2-4, Ichibancho, Nagata-ku, Kobe, Hyogo, 653-0013, Japan;

2 Kobe University Graduate School of Health Sciences, 7-10-2, Tomogaoka, Suma-ku, Kobe, Hyogo, 654-0142, Japan;

3 Department of Faculty of Nursing, Osaka Medical College, 7-6 Hatchonishi-machi, Takatsuki, Osaka, 569-0095, Japan;

4 Department of Physical Therapy, Faculty of Nursing and Rehabilitation, Konan Women's University, 6-2-23, Morikita-machi, Higashinada-ku, Kobe, Hyogo, 658-0001, Japan;

5 Department of Respiratory Medicine, Kobe City Hospital Organization Kobe City Medical Center West Hospital, 2-4, Ichibancho, Nagata-ku, Kobe, Hyogo, 653-0013, Japan;

6 Department of Internal Medicine, Keiwakai Nishioka hospital, 1-52, Nishioka 4-jo 4-chome, Toyohira-ku, Sapporo, Hokkaido, 062-0034, Japan

Keywords: Activities of daily living, dyspnea, chronic obstructive pulmonary disease, interstitial lung disease, Japanese, reliability and validity

Abstract

Purpose: Japan has only a few respiratory disease-specific activity of daily living scales that are accepted outside of Japan, and they are not widely used. The Barthel Index dyspnea (BI-d), an improved version of the Barthel Index (BI), may be popular in Japan. The purpose of this study was to develop the Japanese version of BI-d (J-BI-d) and investigate its reliability and validity.

Patients and methods: The J-BI-d was developed using the basic guidelines for scale translation. The study included patients with chronic respiratory disease, receiving outpatient care at two centers between January 2019 and February 2020. Scores on the J-BI-d, modified Medical Research Council scale (mMRC scale), BI, respiratory function tests, and 6-minute walk distance (6MWD) test were measured. To verify the test-retest reliability, the J-BI-d was re-administered, and the intraclass correlation coefficient (ICC) was obtained. Internal consistency was verified by Cronbach's alpha reliability coefficient, and criterion-related validity was verified through a correlation analysis of the J-BI-d with mMRC scale and 6MWD test. Divergent validity was verified through correlation analysis between the J-BI-d and BI.

Results: Data for 57 participants (mean age 74.4 ± 8.3 years) were analyzed, and reliability testing was performed with 42 of them. The mean time to retest was 8.1 ± 3.0 days, and the ICC (2, 1) was 0.76 (95% CI: 0.62–0.85), indicating high reliability. Cronbach's alpha reliability coefficient was 0.81, indicating high internal consistency. Correlation coefficients of the J-BI-d with 6MWD test ($r = -0.46$, $p < 0.01$) and mMRC scale ($\rho = 0.76$, $p < 0.01$) indicated high criterion-related validity. The J-BI-d and BI had a weak negative correlation ($r = -0.29$, $p < 0.05$), indicating high divergent validity.

Conclusion: The results of this study demonstrate high reliability and appropriate validity of the J-BI-d in patients with chronic respiratory disease.

Introduction

Dyspnea has a serious impact on the prognosis and quality of life of patients suffering from chronic respiratory disease (CRD). Chronic and progressive dyspnea is the most characteristic symptom of chronic obstructive pulmonary disease (COPD),¹⁻² and is also the most important factor influencing the ability to perform activities of daily living (ADL).³ ADL restriction in COPD not only impairs health-related quality of life but is also a predictor of mortality.⁴⁻⁵ Moreover, dyspnea in interstitial lung disease (ILD) has been associated with ADL restriction.⁶ ADL restriction in ILD is associated with health-related quality of life and life prognosis.⁷⁻⁸ Therefore, estimating dyspnea during ADL in patients with CRD is considered essential.

Pulmonary rehabilitation based on quantitative evaluation for patients with CRD is effective for

improving dyspnea and the ability to perform ADL. A current guideline described that pulmonary rehabilitation for stable patients with COPD improves dyspnea, health status, and exercise tolerance (Evidence A).⁹ Pulmonary rehabilitation improves the physiologic response to and actual performance of ADL in patients with COPD.¹⁰ In addition, significant improvements of ADL in patients with ILD have been reported with pulmonary rehabilitation.¹¹ To clarify the degree of improvement in ADL after rehabilitation more clearly, a quantitative scale that reflects disease specificity is essential.¹²

Disease-specific assessment scales are needed for the evaluation of the ability to perform ADL in patients with CRD. The generic ADL scales, such as the Barthel Index (BI), mainly evaluate independence and do not include the element of dyspnea, which is an important symptom in patients with CRD.¹³ As a result, for example, the ability to perform ADL is sometimes overestimated when using these general ADL scales in patients with COPD.¹² Recently, CRD disease-specific ADL scales, which also include assessments about dyspnea, have been used globally.¹⁴⁻¹⁵ The Barthel Index dyspnea (BI-d), which is a modification of the BI developed in Europe, is one of the CRD disease-specific ADL scales including the element of dyspnea.¹⁶ The BI-d is a questionnaire with established guidelines for its measurement. It is reported to be associated with the 6-minute walk distance (6MWD, an index of exercise tolerance) and the modified Medical Research Council scale (mMRC scale, an index of dyspnea).

In Japan, it is necessary to promote the use of disease-specific ADL scales for CRD. Patients with COPD in Japan have several characteristics that differ from those of patients in other countries, for example, being aged and leaner. Hence, it is clinically and academically valuable to make international comparisons in ADL abilities, as assessed by standardized scales. However, and as far as the authors know, few Japanese versions of CRD disease-specific ADL scales, such as the Nagasaki University Respiratory ADL questionnaire and the pulmonary emphysema-ADL, are used globally.¹⁷ Although the modified version of the Pulmonary Functional Status and Dyspnea Questionnaire has been determined to be reliable and valid for the Japanese population, there are few reports of its use, and its use is not widespread in Japan.¹⁷ Moreover, as BI is one of the most widely used ADL scales in Japanese clinical settings, its modified version, BI-d, is also expected to be acceptable in Japan. However, BI-d has not been translated into Japanese yet, and its reliability and validity have not been examined. Therefore, the purpose of this study was to develop the Japanese version of the BI-d (J-BI-d), and investigate its reliability and validity.

Material and methods

Barthel Index dyspnea

BI-d is a scale whose reliability and validity have been established for patients with respiratory diseases.¹⁶ The BI-d was developed based on the BI items; it assesses dyspnea on the 10 ADL items of grooming, bathing, feeding, toilet use, using stairs, dressing, bowel movement, bladder discharge, mobility, wheelchair use, and transfers (bed to chair and back).¹⁶ The examiner judges each item on a five-point scale ranging from 0 to 4, where 0 indicates no sign of dyspnea during execution of ADL; 1 indicates slight dyspnea that does not prevent or slow down the execution of ADL; 2 indicates moderate dyspnea, which can slow down ADL; 3 indicates severe dyspnea, which can greatly slow down ADL; and 4 indicates an extremely severe level of dyspnea, which precludes or reduces ADL. The total BI-d score ranges from 0 (no dyspnea) to 100 (maximum level of dyspnea) according to the original BI grading score.

Translation process

The translation procedure of the J-BI-d followed the basic guidelines for scale translation.¹⁸ First, we contacted the original author to obtain permission to translate the scale into Japanese. Second, two bilingual respiratory physicians and a physiotherapist translated the scale from English to Japanese in turn. Third, a group of experts in respiratory diseases was recruited to examine and integrate more appropriate wording in comparison to the forward-translated scale. Next, a bilingual respiratory physician and a physiotherapist performed the back-translation, which was different from the forward translation. The back-translated scale was reviewed by the original author to confirm its equivalence to the original version. After obtaining consent for using the scale's back translation from the original author, cognitive debriefing was performed on 10 patients with CRD. Cognitive debriefing is a survey

method used to check whether a patient understands concepts and items. The average time required for measurement was 196.7 seconds (SD = 88.0 seconds). Since no incomprehensible expressions or deficiencies were pointed out by the patients, we decided to finalize this tentative Japanese version of the scale.

Study design and participants

A cross-sectional study was performed to investigate the reliability and validity of the J-BI-d. The study participants were outpatients with CRD (confirmed diagnosis of COPD or ILD by a respiratory physician), who were recruited from outpatient care in two hospitals between January 2019 and February 2020. The inclusion criteria were: a stable condition and no infection or acute exacerbation within the past three months, no unstable or severe cardiac disease, and no disease that would interfere with activities of daily living (e.g., stroke or orthopedic disease). The exclusion criteria were a diagnosis of asthma only, acute exacerbation during the study period, difficulty in understanding the questionnaire, and not providing consent for the study. Participants' personal information used in this study was anonymized with a research-specific number and a correspondence table, and was stored on electronic media. The files were managed with a password so that only the researcher could view them. This study was approved by the ethics committee of Kobe University (N611-1); Kobe City Hospital Organization, Kobe City Medical Center, West Hospital (N17-003); and Nishioka Hospital (N1806). All participants provided written informed consent in accordance with the ethical standards set in the Declaration of Helsinki.

Measurements

Basic patient data were obtained from medical records. The J-BI-d was measured twice to verify intrarater reliability. The participants for the second measurement were those who had an appointment with a respiratory physician and whose interval from the first measurement was approximately one to four weeks. The anchor index, a self-administered questionnaire assessing health status on a 3-point scale, was used to analyze whether the condition of the participants in the second measurement had changed from the initial measurement.¹⁹ Participants whose status change was confirmed by the anchor index were excluded from the retest. The lung function test was performed according to the American Thoracic Society's guidelines²⁰ and measured FVC, FEV₁. The mMRC scale was used as a measure of dyspnea. The 6-minute walk test was performed to assess functional exercise capacity according to standard procedures.²¹ The mMRC scale has been validated as a measure of dyspnea, and the 6MWD as a test of exercise tolerance.⁹ For divergent validity, we tested whether the BI-d measures different concepts than the BI. Specifically, we confirmed if there was low association between J-BI-D and BI.

Statistical analysis

Data were presented as counts and percentages, means and standard deviations, or median and interquartile range, as appropriate. The Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) taxonomy was applied for statistical analysis of validity and reliability.²² The floor effect and ceiling effect were visually confirmed by histograms. The statistical analyses were performed with R (ver. 3. 6. 1) on an EZR platform (Saitama Medical Center, Jichi Medical University, Saitama, Japan), and with R studio (ver. 1. 2. 5042). P values less than 0.05 were considered statistically significant.

Construct Validity

The criterion-related validity of the J-BI-d in comparison to the mMRC scale and 6MWD test was evaluated using Spearman's rank correlation coefficient and Pearson's correlation coefficient, respectively. The strength of the association was defined as very weak for < 0.2, weak for 0.2-0.35, moderate for 0.35-0.5, and strong for >0.5.²³ Comparisons between the J-BI-d and grades of the mMRC scale was performed using the Kruskal-Wallis test to confirm the significant main effect. Multiple comparison tests between the J-BI-d and mMRC scale were performed using Mann-Whitney U test and adjusted for significance levels by the Bonferroni test. To verify the divergent validity of the BI-d, Pearson's correlation analysis between the BI-d and BI was performed.

Reliability

Internal consistency, test-retest reliability, and smallest detectable change (SDC) were analyzed. The scale's internal consistency was evaluated using Cronbach's alpha. Test-retest reliability was for analyzing the J-BI-d's reproducibility in the test (day 1) and the retest (day 2), we used the intraclass correlation coefficient (ICC (2,1)). The ICC and its 95% confidence intervals were calculated based on the absolute-agreement, two-way mixed-effects model. The ICC value is considered acceptable if it is >0.70 .²⁴ SDC is based on standard error of measurement (SEM), which is the variability in measurements (SD) of the same individual, with a confidence of 95%, and is expressed in the unit of the measurement. It was estimated by computing the square root of the within-subject variance of the patients ($SEM_{\text{agreement}} = \sqrt{\sigma_{\text{between measurement}} + \sigma_{\text{residual}}}$). Variance components were obtained from a multilevel mixed effects model (restricted maximum-likelihood estimates). Since SDC is the smallest amount of change in individuals that can be detected beyond measurement error with a confidence of 95%, it is calculated as $1.96 * \sqrt{2} * SEM$.²⁵

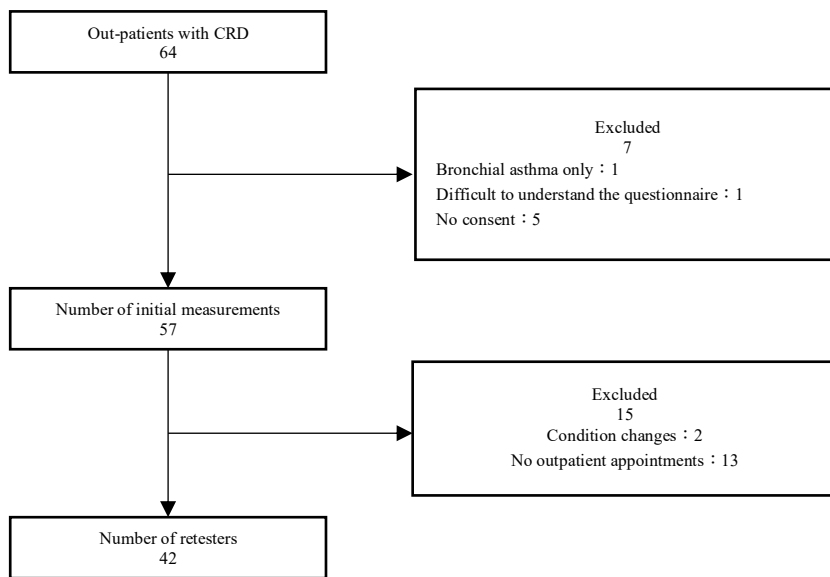


Figure 1. Flow diagram of patients with chronic respiratory diseases enrolled in this study.

Results

Sixty-four patients with outpatient CRD were convened, out of whom 57 were included in the analysis of this study (Figure 1). Fifty (87.7%) participants of the patient group in this study had COPD (involving an asthma-COPD overlap) and seven (12.3%) had ILD (Table 1). With regard to the severity level of the patients with COPD, eight (16%) were in GOLD Stage I, 26 (52%) were in Stage II, 11 (22%) were in Stage III, and five (10%) were in Stage IV. For the mMRC scale, 10 patients were in grade 0 (18%), 19 were in grade 1 (33%), 19 were in grade 2 (33%), nine were in grade 3 (16%), and none were in grade 4 (0%). The mean value of J-BI-d was 14 ± 12.6 , and the ADL of the present target population was relatively mild and varied. The histogram of the total J-BI-d scores showed a high percentage, biased toward lower scores (Figure 2). No ceiling effect was found; however, a floor effect was suspected.

Construct Validity

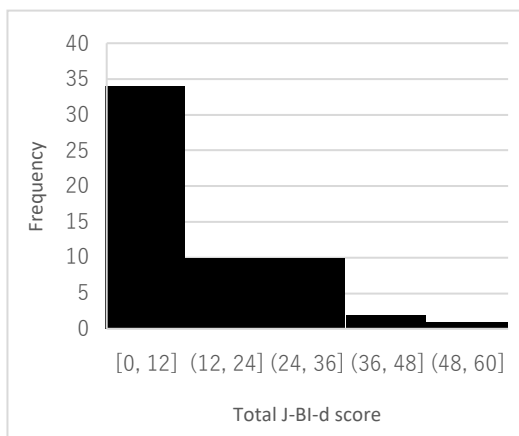
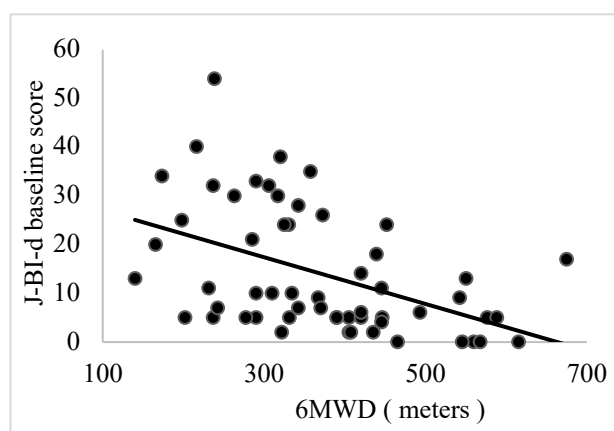
There was a moderate negative correlation between the BI-d and 6MWD test ($r = -0.46$, $p < 0.01$) (Figure 3). There was a strong positive correlation between the BI-d and mMRC scale ($\rho = 0.76$, $p < 0.01$). Significant main effects of the mMRC scale grades among J-BI-d scores were confirmed by the Kruskal-Wallis test. Multiple comparisons revealed that there were significant differences in the J-BI-d scores among all four grades of the mMRC scale (Figure 4). This result indicates that as the mMRC scale grade increases, the J-BI-d score increases. The J-BI-d and BI had a weak negative correlation ($r = -0.29$, $p < 0.05$), indicating high divergent validity.

Table 1. Baseline characteristics of the participants

Variable	Results (n=57)
Male ^a , n (%)	45 (78.9)
Age ^b , year	74.4 ± 8.3
BMI ^b , kg/m ²	22.2 ± 3.9
COPD ^a , n (%)	50 (87.7)
asthma-COPD overlap ^a , n (%)	22 (38.6)
ILD ^a , n (%)	7 (12.3)
FEV ₁ /FVC ratio ^b , %	58.0 ± 16.5
FEV ₁ ^b , %predicted	61 ± 20.6
GOLD stage (with COPD)	
Stage I, II, III, IV	8, 26, 11, 5
mMRC ^c	1 [1]
CAT ^b	15.8 ± 10.6
6MWD ^b , m	370.4 ± 122.4
BI ^b	99.4 ± 2.5
J-BI-d ^b	14 ± 12.6
LTOT ^a , n (%)	16 (28)

Notes : ^an (%), ^bmean & SD, ^cmedian[IQR]

BMI = body mass index, ILD = interstitial lung disease, FEV₁ = forced expiratory volume in one second, FVC = forced vital capacity, GOLD = Global Initiative for Chronic Obstructive Lung Disease, mMRC = modified Medical Research Council, CAT = COPD Assessment Test, 6MWD = 6-min walk distance, BI = Barthel Index, J-BI-d = Japanese version of Barthel Index dyspnea, LTOT = long-term oxygen therapy

**Figure 2.** Histogram of J-BI-d total score**Figure 3.** Correlation between meters in the 6MWD test and J-BI-d scale.

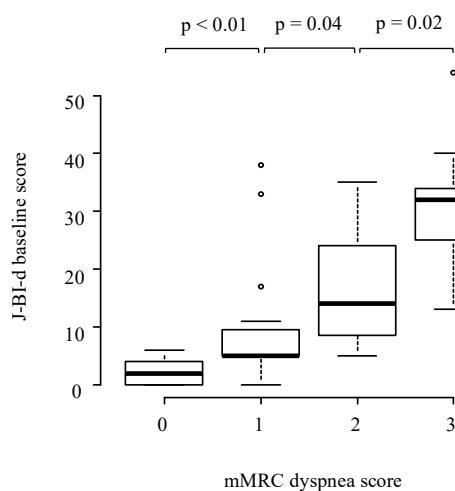


Figure 4. Box plot of the distribution of J-BI-d scores according to mMRC scale dyspnea score grading measured at baseline (mMRC grade 4 : n = 0)

Table 2. The Cronbach's alpha coefficient by deleting item by item

Delete item	Cronbach's alpha
Bathing	0.78
Bladder	0.80
Bowels	0.81
Dressing	0.77
Feeding	0.79
Grooming	0.79
Mobility	0.84
Stairs	0.77
Toilet use	0.80
Transfers	0.78

Reliability

The Cronbach's alpha coefficient for the J-BI-d was 0.81, indicating high internal consistency. Cronbach's alpha coefficient from deleting item by item is shown in table 2. Test-retest reliability was analyzed in 42 patients, excluding 2 patients due to change in condition and 13 patients who did not have a medical appointment in the stipulated period (Figure 1). The average interval between the initial test and the retest was 8.1 days (SD = 3.0 days). The ICC (2,1) for the J-BI-d was 0.76 (95% CI: 0.62–0.85), indicating high intrarater reliability. SEM was 7.19, SDC was 19.92.

Discussion

The present study was the first to examine the reproducibility, internal consistency, and validity of the translated J-BI-d in patients with CRD. The results of the study showed that the J-BI-d had a high ICC value and Cronbach's alpha coefficient, as well as a strong correlation with both the mMRC scale and 6MWD test. These results indicate high reproducibility, internal consistency, and validity of the J-BI-d.

The J-BI-d was developed by following the basic guidelines for scale translation.¹⁸ Both the forward and back translation of the BI-d involved independent bilinguals, and the original authors approved the questions of the J-BI-d as accurately reflecting the intent of the questions in the original version. When cognitive debriefing for the J-BI-d was conducted with 10 patients with CRD, none of the patients reported encountering incomprehensible words or expressions, and there were no missing values in the responses. The average time taken for cognitive debriefing in this study was about 3 minutes, which is less than the 9 minutes taken for administering the previous Japanese version of the Pulmonary Functional Status and Dyspnea Questionnaire—a disease-specific ADL scale for COPD.²⁶ For busy clinicians, a quick and easily understandable assessment scale is strongly required. Therefore, the J-BI-d is clinically useful as a questionnaire for patients with CRD.

The result of this study proved the J-BI-d's high reliability in patients with CRD. Since this study excluded patients with symptomatic changes, as detected in the anchor index at the time of retesting, the effect on the J-BI-d scores is expected to be minimal. The ICC obtained from the measured data was 0.76, indicating excellent intrarater reliability.²⁷ The average interval period between the initial test and retest was 8.1 ± 3.0 days in the current study. Generally, it is suggested that the test-retest interval period be from one to two weeks. The high reliability of this study was demonstrated from the

data obtained within this recommended appropriate interval. The test-retest ICC of the original English version of the scale was 0.99, which is higher than the ICC value obtained in the present study. The result may have been influenced by the difference in the test-retest interval period, since the interval period in the original English version of the scale was shorter by 48 hours than that in the current study. In terms of internal consistency, the Cronbach's alpha coefficient of the J-BI-d was 0.81, which is as high as that of the original English version. Previous studies have suggested that high internal consistency is demonstrated by a Cronbach's alpha coefficient of over 0.7.²⁸ Therefore, the J-BI-d can be used for repeated ADL assessments of patients with CRD, for instance, in the case of assessing patients before and after rehabilitation.

The current study demonstrated the J-BI-d's sufficient criterion-related validity from multiple perspectives of analyses. First, the J-BI-d had a significant positive correlation with the mMRC scale and 6MWD test. This suggests that higher J-BI-d scores imply greater dyspnea and more severely impaired exercise tolerance. In addition, the J-BI-d total score was significantly different among the grades of the mMRC scale. These results suggest that ADL limitations caused by dyspnea may be stratified by the J-BI-d. A weak correlation between the J-BI-d and BI suggests that the J-BI-d evaluates a different aspect of ADL ability as compared to the BI. The correlation between BI-d and BI in the original English version was low ($r = -0.38$) as well, similar to the results of this study. Based on these results, it can be assumed that the J-BI-d provides an accurate measure of dyspnea during ADL in Japanese patients with CRD.

The J-BI-d developed in this study is expected to facilitate the study of limitations in performing ADL due to dyspnea in Japanese patients with CRD, which is an important step forward, as it has been difficult to compare the differences in performing ADL due to dyspnea internationally. Respiratory disease-specific ADL scales that have been developed and used in Japan, such as the Nagasaki University Respiratory ADL questionnaire and pulmonary emphysema-ADL, are hardly used outside of Japan. Thus, comparing the ability to perform ADL among patients with CRD in Japan to the ability of patients in other countries has been difficult. In addition, the reliability and validity of the London Chest Activity of Daily Living Scale, which is used internationally, have not been verified in the Japanese population because this scale does not capture the lifestyle of Japanese people and has not been translated into Japanese. In contrast, the Pulmonary Functional Status and Dyspnea Questionnaire is an instrument whose reliability and validity have been proven among the Japanese population, but there have been few reports of its use in Japan. The BI is one of the most widely used ADL scales in Japan; hence, the J-BI-d, which has been created and translated based on the BI, is expected to be accepted in actual clinical practice in Japan. We hope that the J-BI-d will be widely used both clinically and in research, and that the number of reports on the ADL of CRD patients in Japan will increase.

However, several limitations prevent the generalization of the results of this study. First, all of the patients in the study were outpatients with settled symptoms. Most of them were in mMRC scale grade 1–2 and there was no patient in mMRC scale grade 4. Thus, the results of this study may not apply to patients with the most severe dyspnea, such as those in mMRC scale grade 4. This observation is similar to that in the original English version of the scale, wherein patients with the most severe dyspnea were not included in the results.¹⁶ Thus, the applicability of the J-BI-d to cases of severe respiratory distress, including hospitalized patients, should be investigated in future studies. Second, the patients in this study were a mix of those having COPD and ILD. The pathogenesis of COPD and ILD is different, and most of the similar studies to date have confined their coverage to COPD.¹⁴⁻¹⁵ However, COPD and ILD have in common the occurrence of chronic, progressive increases in dyspnea and reduced exercise tolerance.^{1-2,6} Since patients with COPD and ILD were also included in the original English version of the scale, it can be assumed that the J-BI-d can be used for both sets of patients. Third, the sensitivity of changes before and after pulmonary rehabilitation could not be investigated using the J-BI-d. Further investigation of this issue is required. Fourth, the sample size was inadequate. COSMIN requires at least 80 participants²²; and it is undeniable that this led to a lack of testing power. Additionally, most of the subjects have COPD; therefore, it is not possible to generalize the results to all patients with CRD. In future, we would like to conduct a similar analysis by narrowing the target population to ILD patients.

Conclusion

This study demonstrated that the J-BI-d had a high reliability and sufficient validity as a tool to assess the ability to perform ADL in Japanese patients with CRD. The J-BI-d can be adopted as a new tool to adequately measure ADL in patients with CRD in Japan. Further studies should use the J-BI-d to evaluate the effects of the intervention for and prognosis of CRD.

Acknowledgments

The authors would like to thank Michele Vitacca MD FERS ICS S.Maugeri IRCCS Pavia (Italy) for their constructive advice on the manuscript. The authors would also like to thank Susumu Sato from the Department of Respiratory Medicine, Graduate School of Medicine, Kyoto University and Kenji Kuno from the Japan International Cooperation Agency for their help with the translation. Finally, the authors would like to thank the members of the Department of Rehabilitation, Kobe City Medical Center, West Hospital, including Chisato Nagatani, Kiyohiko Kado, Takashi Kadooka, Ryoya Fujisaki, and Kohei Otake for their cooperation in the acquisition of data, and the members of the Department of Community Health Sciences, Kobe University Graduate School of Health Sciences, including Kaoru Hanaie, Kentaro Iwata, Yoji Yamada, Kanji Yamada, Ken Umehara, Keita Ohashi, Kazumo Miura, Saki Kakuta, and Rina Omote for their constructive comments on this paper. We would like to thank Editage (www.editage.com) for English language editing.

Authors' contributions

Takumi Yamaguchi made substantial contributions to the literature search, data collection, study design, and analysis of data, and was involved in manuscript preparation. Akio Yamamoto made substantial contributions to the literature search, data collection, study design, and analysis of data, and was involved in the review of the manuscript. Yutaro Oki made substantial contributions to the literature search, study design, and analysis of data, and was involved in the review of the manuscript. Hideki Sakai made substantial contributions to data collection and was involved in the review of the manuscript. Shogo Misu made substantial contributions to the literature search, study design, and analysis of data, and was involved in the review of the manuscript. Yusuke Iwata made substantial contributions to the data collection, study design, and analysis of data, and was involved in the review of the manuscript. Masahiro Kaneko made substantial contributions to the data collection, study design, and analysis of data, and was involved in the review of the manuscript. Kaku Sawada made substantial contributions to the data collection, study design, and analysis of data, and was involved in the review of the manuscript. Yukari Oki made substantial contributions to the literature search, study design, and analysis of data, and was involved in the review of the manuscript. Yuji Mitani made substantial contributions to the literature search, study design, and analysis of data, and was involved in the review of the manuscript. Kumiko Ono made substantial contributions to the literature search, study design, and analysis of data, and was involved in the review of the manuscript. Akira Ishikawa made substantial contributions to the study design and analysis of data and was involved in the review of the manuscript.

All authors who contributed to the data analysis, drafting or revising of the paper, and agreed on the journal to which the article will be submitted, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

References

1. O'Donnell DE, Milne KM, James MD, de Torres JP, Neder JA. Dyspnea in COPD: New Mechanistic Insights and Management Implications. *Adv Ther.* 2020;37(1):41–60.
2. O'Donnell DE, Neder JA. The enigma of dyspnoea in COPD: A physiological perspective. *Respirology.* 2020;25(2):134–136.
3. Ries AL, Bauldoff GS, Carlin BW, et al. Pulmonary rehabilitation: Joint ACCP/AACVPR evidence-based clinical practice guidelines. *Chest.* 2007;131(5):4S–42S.

4. Vitacca M, Paneroni M. Rehabilitation of patients with coexisting COPD and heart failure. *COPD J Chronic Obstr Pulm Dis*. 2018;15(3):231–237.
5. Yohannes AM, Baldwin RC, Connolly M. Mortality predictors in disabling chronic obstructive pulmonary disease in old age. *Age Ageing*. 2002;31(2):137–140.
6. Raghu G, Collard HR, Egan JJ, et al. An official ATS/ERS/JRS/ALAT statement: Idiopathic pulmonary fibrosis: Evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med*. 2011;183(6):788–824.
7. Matsuda T, Taniguchi H, Ando M, et al. COPD assessment test for measurement of health status in patients with idiopathic pulmonary fibrosis: A cross-sectional study. *Respirology*. 2017;22(4):721–727.
8. Nishiyama O, Taniguchi H, Kondoh Y, et al. A simple assessment of dyspnoea as a prognostic indicator in idiopathic pulmonary fibrosis. *Eur Respir J*. 2010;36(5):1067–1072.
9. Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease (2020 Report); 2020. Available from: https://goldcopd.org/wp-content/uploads/2019/12/GOLD-2020-FINAL-ver1.2-03Dec19_WMV.pdf. Accessed October 5th, 2020.
10. Vaes AW, Delbressine JML, Mesquita R, et al. Impact of pulmonary rehabilitation on activities of daily living in patients with chronic obstructive pulmonary disease. *J Appl Physiol*. 2018;126(3):607–615.
11. Kozu R, Senjyu H, Jenkins SC, Mukae H, Sakamoto N, Kohno S. Differences in response to pulmonary rehabilitation in idiopathic pulmonary fibrosis and chronic obstructive pulmonary disease. *Respiration*. 2011;81(3):196–205.
12. Yohannes AM, Roomi J, Waters K, Connolly MJ. A comparison of the Barthel index and Nottingham extended activities of daily living scale in the assessment of disability in chronic airflow limitation in old age. *Age Ageing*. 1998;27(3):369–374.
13. Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. *J Clin Epidemiol*. 1989;42(8):703–709.
14. Garrod R, Bestall JC, Paul EA, Wedzicha JA, Jones PW. Development and validation of a standardized measure of activity of daily living in patients with severe COPD: The London chest activity of daily living scale (LCADL). *Respir Med*. 2000;94(6):589–596.
15. Lareau SC, Meek PM, Roos PJ, Calif LL, Ariz T. Development and testing of the modified version of the Pulmonary Functional Status and Dyspnea Questionnaire (PFSDQ-M). *Heart Lung*. 1998;27(3):159–168.
16. Vitacca M, Paneroni M, Baiardi P, et al. Development of a Barthel Index based on dyspnea for patients with respiratory diseases. *Int J COPD*. 2016;11(1):1199–1206.
17. Janaudis-Ferreira T, Beauchamp MK, Robles PG, Goldstein RS, Brooks D. Measurement of activities of daily living in patients with COPD: A systematic review. *Chest*. 2014;145(2):253–271.
18. Wild D, Grove A, Martin M, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health*. 2005;8(2):94–104.

19. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in medicine: A practical guide*. Cambridge, England: Cambridge University Press; 2011.
20. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *Eur Respir J*. 2005;26(2):319–338.
21. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: Guidelines for the six-minute walk test. *Am J Respir Crit Care Med*. 2002;166(1):111–117.
22. COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs) user manual. Version 1.0 dated February 2018. Available from: https://www.cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018-1.pdf. Accessed October 5th, 2020. Accessed May 5th, 2021
23. Guyatt GH, King DR, Feeny DH, et al. Generic and specific measurement of health-related quality of life in a clinical trial of respiratory rehabilitation. *J Clin Epidemiol*. 1999;52(3):187-192
24. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60(1):34-42
25. de Vet HC, Terwee CB, Knol DL, et al. When to use agreement versus reliability measures. *J Clin Epidemiol*. 2006;59(10):1033-9
26. Kojima S, Ando M, Okazawa H, Sakakibara S, Hashimoto S, Nakajima S. [Development of a Japanese version of the modified questionnaire on the pulmonary functional status and dyspnea]. *Nihon Kokyuki Gakkai Zasshi [The Journal of the Japanese Respiratory Society]*; 2004;42(6):486–490. [in Japanese].
27. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol Assess*. 1994;6(4):284–290.
28. Bland JM, Altman DG. Statistics notes: Cronbach's alpha. *BMJ*. 1997;314(7080):572.