Geriatrics & Gerontology International

DOI: 10.1111/ggi.14439

ORIGINAL ARTICLE

EPIDEMIOLOGY, CLINICAL PRACTICE AND HEALTH

Predictive validity of the modified Kihon Checklist for the incidence of functional disability among older people: A 3-year cohort study from the JAGES

Ryota Watanabe,[1,2] Taishi Tsuji,[2,3] Kazushige Ide,[2] Taiji Noguchi,[4] Mikako Yasuoka,[1] Koto Kamiji,[1,5] Shosuke Satake,[1,6] Katsunori Kondo[2,7] and Masayo Kojima[1]

[1]Department of Frailty Research, Research Institute, National Center for Geriatrics and Gerontology, Obu City, Aichi, Japan
[2]Department of Social Preventive Medical Sciences, Center for Preventive Medical Sciences, Chiba University, Chiba city, Chiba, Japan
[3]Faculty of Health and Sport Sciences, University of Tsukuba, Bunkyo city, Tokyo, Japan
[4]Department of Social Science, Research Institute, National Center for Geriatrics and Gerontology, Obu City, Aichi, Japan
[5]Center for Professional Development of Teachers, Shizuoka University, Shizuoka City, Shizuoka, Japan
[6]Department of Geriatric Medicine, Hospital, National Center for Geriatrics and Gerontology, Obu city, Aichi, Japan
[7]Department of Gerontological Evaluation, Research Institute, National Center for Geriatrics and Gerontology, Obu City, Aichi, Japan

Correspondence
Ryota Watanabe PhD, Department of Frailty Research, Research Institute, National Center for Geriatrics and Gerontology, 7-430 Morioka, Obu City, Aichi 474-8511 Japan.
Email: ryota.watanabe.1222@gmail.com

Received: 6 January 2022
Revised: 3 June 2022
Accepted: 18 June 2022

Aim: The original Kihon Checklist, validated to predict the incidence of functional disability, has been modified to capture both functional ability (can/cannot) and performing state (do/do not). However, the predictive validity of the modified Kihon Checklist remains unverified. Therefore, this study intends to verify the predictive validity of the modified Kihon Checklist and to clarify whether predictive discrimination differs between the classification method of functional ability and performing state.

Methods: The participants comprised 67 398 older people who responded to the Japan Gerontological Evaluation Study (2016). They were followed for 3.1 years on average. Cox's proportional hazards model with incidence of functional disability as the endpoint was used to calculate the hazard ratio, adjusted for sex and age. The independent variables were judged by two classification methods, functional ability and performing state, using nine indicators based on the modified Kihon Checklist. Additionally, we examined whether the two classification methods produced different C-index estimates.

Results: Incidence of functional disability occurred in 6232 participants (9.2%). The adjusted hazard ratio of those to whom the nine indicators applied was significantly higher than that of those to whom they did not. The range of the hazard ratio was 1.50–3.82 for both classification methods. The C-index was slightly higher when the classification was based on performing state than when it was based on on functional ability.

Conclusions: Although predictive discrimination was slightly higher for the performing state than for functional ability, the predictive validity of the modified Kihon Checklist was confirmed for both. **Geriatr Gerontol Int 2022; ••: ••–••.**

Keywords: functional ability, geriatric assessment, long-term care insurance, performing state, predictive discrimination.

Introduction

The world's population of older people is expected to increase by approximately 34%, from 1 billion in 2019 to 1.4 billion in 2030.[1] Consequently, it is necessary to effectively maintain and improve the health of older people.

The World Health Organization has indicated that healthy ageing should include functional ability as a health-related characteristic that enables a person to do something valuable and remain in a desirable state, and it aims to maximize older people's functional ability.[2] One of the four priority areas for achieving this aim is to "improve measurement, monitoring, and understanding."[2]

Japan has the world's oldest population,[3] and its government has been using the Kihon Checklist (KCL) as a tool for the early detection of older people who are at risk for functional disability.[4] The KCL is a 25-item self-administered questionnaire with two Yes/No options and is a comprehensive geriatric assessment that evaluates the functions of older people.[5] In this evaluation,

performing state (do/do not) is used as a criterion, and one point is added to the score if the content of each question is considered to limit daily function. A systematic KCL review has shown that a higher number of items predicts problematic health indicators, including the incidence of functional disability.[6] Furthermore, longitudinal studies have indicated that the KCL can predict the incidence of functional disability among older people.[7–9]

However, the Ministry of Health, Labour and Welfare (MHLW) has also developed a modified version of the KCL for the Public Survey of Long-Term Care Prevention and Needs in Spheres of Daily Life, which introduced three options for several of the responses in order to grasp both functional ability (can/-cannot) and performing state.[10] The option "I can, but don't do it" was added. This change was made to reflect the opinion of the municipal staff working on the prevention of functional decline during the development of the survey questionnaire. Municipalities use this modified KCL in the Public Survey of Long-Term Care Prevention and Needs in Spheres of Daily Life to identify issues and assess the risk of individuals developing long-term care needs in each community.[11] However, it is unclear whether the modified KCL can predict the incidence of functional disability among older people. Additionally, it is unclear whether predictive validity differs depending on whether the three-option response items are classified as focusing on functional ability or performing state.

This study aims to determine whether the modified KCL has predictive validity for the incidence of functional disability. In addition, it aims to examine the extent to which predictive validity differs depending on whether the three options are classified as functional ability or performing state.

Methods

This study utilized the longitudinal cohort data from the Japan Gerontological Evaluation Study (JAGES). The JAGES is an ongoing cohort study investigating social and behavioral factors related to health decline, including the incidence of functional decline or cognitive impairment among individuals aged ≥65 years.[12,13] The baseline survey that was undertaken was self-reported and distributed by mail to 129 311 people aged ≥65 years between October and November 2016. The participants were older people who did not receive public long-term care insurance (LTCI) certification and were selected from 18 municipalities in 10 prefectures of Japan. Furthermore, random sampling methods were used in 10 large municipalities, and complete enumeration survey methods were used in eight smaller municipalities. Among 92 234 respondents (response rate 71.3%), sex, age, and area of residence were validly collected from 84 376 respondents. Out of the valid respondents, 83 079 (98.5%) participants were successfully linked to the incident records of LTCI certification. Those participants who answered the 25 questions of the KCL completely were included in the study, and 15 681 participants who did not complete the questionnaire were excluded.

The analytical sample for this study comprised 67 398 participants (32 314 men and 35 084 women).

Ethical approval for the study was obtained from the National Center for Geriatrics and Gerontology (application number: 1274–2, approval date: December 18, 2020), Chiba University (application number: 3442, approval date: December 22, 2020), and the Japan Agency for Gerontological Evaluation Study (application number: 2019–01, approval date: November 26, 2020).

Dependent variable

The dependent variable was the incidence of functional disability.[7–9,14–16] and information was collected from those eligible for LTCI benefits who were identified as having functional disability. Eligibility determination for LTCI was based on nationally standardized procedures, including physician evaluation of physical and cognitive functions.[17] In addition, verification of certification status during an average of 3.1 years of follow-up (from 2016 to 2019) was performed by linking cohort participants to records in the national LTCI database.

Independent variables

Table 2 lists the nine independent variables used in the study. The KCL comprises 25 questions in seven domains, namely the Instrumental Activities of Daily Living (IADL) domain, physical domain, nutrition domain, eating domain, houseboundness domain, memory domain, and mood domain. Each question statement is listed in Table 1. Based on the scores of the original KCL,[5] a total of nine indicators, which have been validated to predict the incidence of functional disability,[7,9,14–16] were evaluated using the modified KCL in accordance with the frailty,[9,18] risk assessment scale,[7] and seven domains from the MHLW.[4,5,15]

Frailty, as assessed by the KCL, was related to the Cardiovascular Health Study criteria of Fried *et al.* (2001), a common measure of frailty.[19] From a total KCL score of 25 points, 0–3 points was defined as robust; 4–7 points as pre-frail; and 8 points or more as frailty.[9,18] The sensitivity and specificity for frailty was reported to be 89.5% and 80.7%, respectively and for pre-frailty to be 70.3% and 78.3%.[18]

A risk assessment scale was calculated by extracting 10 important items from the KCL according to the criteria of Tsuji *et al.* and adding points for each question, sex, and age (range: 0–48 points). Additionally, a cutoff of 16/17 points was used.[7]

The seven domains, based on the original KCL presented by the MHLW consist of the following questions: (i) multi-domain were nos. 1–20, (ii) physical domain were nos. 6–10, (iii) nutrition domain were nos. 11 and 12, (iv) eating domain were nos. 13–15, (v) houseboundness domain were nos. 16–17, (vi) memory domain were nos. 18–20, and (vii) mood domain were nos. 21–25. The at-risk categories are scored based on number of difficulties in each domain as follows: (i) ≥10 (multi-domain), (ii) ≥3 (physical domain), (iii) both of two (nutrition domain), (iv) ≥2 (eating domain), (v) if no. 16 applies (houseboundness domain), (vi) ≥1 (memory domain), and (vii) ≥2 (mood domain).

The modified KCL included three options within questions 1–3 and 6–8. The three answer options were as follows: (i) I can and do it, (ii) I can, but don't do it, and (iii) I cannot do it.[11] The "I can, but don't do it" item were conducted both when classified as applicable and non-applicable to functional decline. When "I can, but don't do it" is classified as non-applicable, the focus is on functional ability (can/cannot), and when it is classified as applicable, the focus is on performing state (do/do not).

Statistical analysis

The Cox proportional hazards model was used to estimate hazard ratios (HRs) and construct 95% confidence intervals (CIs) when nine indicators were applicable, in comparison with those that were not applicable. In addition to the crude model, the analysis further adjusted for age and sex.

We calculated Harrell's C statistics (C-index) to clarify whether predictive discrimination differs by using two different

Table 1 Distribution of participants' responses at baseline and the cumulative incidence of functional disability during the follow-up period

No.	$n = 67\,398$	Proportion (%)	Cumulative incidence	HR[†]	95% CI	*P*-value
Sex						
Male	32 314	47.9	9.0	1.00		
Female	35 084	52.1	9.5	1.08	1.02–1.13	0.003
Age (years)						
65–69	21 519	31.9	1.7	1.00		
70–74	18 976	28.2	5.0	3.02	2.68–3.41	<0.001
75–79	15 170	22.5	10.8	6.83	6.09–7.65	<0.001
80–84	8083	12.0	22.7	15.65	13.98–17.52	<0.001
85+	3650	5.4	39.9	32.48	28.94–36.46	<0.001
1 Do you go out by bus or train by yourself?[‡]						
I cannot	1832	2.7	43.5	2.94	2.71–3.19	<0.001
I do not	7643	11.3	23.2	2.14	2.02–2.27	<0.001
2 Do you go shopping to buy daily necessities by yourself?						
I cannot	1069	1.6	45.7	3.36	3.06–3.70	<0.001
I do not	9717	14.4	16.8	1.76	1.66–1.86	<0.001
3 Do you manage your own deposits and savings at the bank?[‡]						
I cannot	1694	2.5	27.4	2.31	2.10–2.55	<0.001
I do not	10 653	15.8	14.6	1.61	1.52–1.71	<0.001
4 Do you sometimes visit your friends?						
No	25 094	37.2	12.1	1.59	1.51–1.68	<0.001
5 Do your family or friends turn to you for advice?						
No	12 793	19.0	14.8	1.72	1.62–1.81	<0.001
6 Do you normally climb stairs without using a handrail or wall for support?[‡]						
I cannot	6660	9.9	29.1	2.64	2.49–2.80	<0.001
I do not	24 968	37.0	15.5	1.95	1.85–2.05	<0.001
7 Do you normally stand up from a chair without any aids?[‡]TF1						
I cannot	3711	5.5	31.9	2.63	2.46–2.81	<0.001
I do not	13 005	19.3	18.4	1.91	1.81–2.01	<0.001
8 Do you normally walk continuously for 15 min?[‡]						
I cannot	1812	2.7	35.5	2.72	2.51–2.96	<0.001
I do not	12 282	18.2	15.8	1.88	1.78–1.99	<0.001
9 Have you experienced a fall in the past year?						
Yes	15 314	22.7	14.5	1.54	1.46–1.62	<0.001
10 Do you have a fear of falling while walking?						
Yes	25 756	38.2	14.9	1.89	1.79–1.99	<0.001
11 Have you lost 2 kg or more in the past 6 months?						
Yes	8463	12.6	14.7	1.67	1.57–1.78	<0.001
12 Height: cm, Weight: kg, BMI: kg/m^2 If BMI is less than 18.5, this item is scored						
<18.5 kg/m^2	4654	6.9	15.9	1.59	1.47–1.71	<0.001
13 Do you have any difficulties eating tough foods compared to 6 months ago?						
Yes	18 000	26.7	13.1	1.34	1.27–1.41	<0.001
14 Have you choked on your tea or soup recently?						
Yes	11 796	17.5	12.8	1.28	1.21–1.36	<0.001
15 Do you often experience having a dry mouth?						
Yes	13 433	19.9	13.9	1.47	1.39–1.55	<0.001
16 Do you go out at least once a week?						
<1 day/week	2157	3.2	27.8	2.26	2.08–2.47	<0.001
17 Do you go out less frequently compared to last year?						
Yes	9664	14.3	20.0	2.07	1.96–2.19	<0.001
18 Do your family or your friends point out your memory loss? e.g., "You ask the same question over and over again."						
Yes	8753	13.0	15.5	1.63	1.53–1.73	<0.001
19 Do you make a call by looking up phone numbers?						
No	3394	5.0	20.6	2.15	1.99–2.33	<0.001
20 Do you find yourself not knowing today's date?						
Yes	15 402	22.9	13.6	1.47	1.40–1.55	<0.001
21 In the last 2 weeks have you felt a lack of fulfillment in your daily life?						
Yes	8112	12.0	16.3	1.82	1.71–1.93	<0.001

(*Continues*)

Table 1 Continued

No.	$n = 67\,398$	Proportion (%)	Cumulative incidence	HR[†]	95% CI	*P*-value
22 In the last 2 weeks have you felt a lack of joy when doing the things you used to enjoy?						
Yes	6010	8.9	18.7	1.94	1.82–2.07	<0.001
23 In the last 2 weeks have you felt difficulty in doing what you could do easily before?						
Yes	13 127	19.5	17.9	1.91	1.82–2.02	<0.001
24 In the last 2 weeks have you felt helpless?						
Yes	11 405	16.9	16.9	1.71	1.62–1.81	<0.001
25 In the last 2 weeks have you felt tired without a reason?						
Yes	13 801	20.5	15.9	1.75	1.66–1.84	<0.001

Frailty: of the total score of nos. 1–25: 0–3, robust; 4–7, pre-frailty; 8 or more, frailty. Risk assessment scale: risk applicable for 17 points or more of the 12 items for sex, age, nos. 1–3, nos. 6–10, no. 12, and no. 17. Multi-domain: risk applicable for 10 or more of the 20 items from no. 1 to no. 20. Physical domain: risk applicable for 3 or more of the 5 items from no. 6 to no. 10. Nutrition domain: risk applicable for the 2 items from no. 11 and no. 12. Eating domain: risk applicable for 2 or more of the 3 items from no. 13 to no. 15. Houseboundness domain: risk applicable for no. 16 of the 2 items from no. 16 and no. 17. Memory domain: risk applicable for 1 or more of the 3 items from no. 18 to no. 20. Mood domain: risk applicable for 2 or more of the 5 items from no. 21 to no. 25.

[†]Adjusted sex and age.

[‡]Because there are three options, two methods (can/cannot, do/do not) are shown.

95% CI, 95% confidence interval; BMI, body mass index; HR, hazard ratio.

classification methods: functional ability and performing state. The C-index estimates the probability of agreement between predicted and observed responses. A value of 0.5 indicates that predictive discrimination is not possible, while a value of 1.0 indicates that participants with different outcomes are completely separated.[20] The C-index is similar to the area under the curve (AUC), but it is a measure of discriminability that considers the censoring nature of the data.[20] In addition, statistical tests[21] were used to determine if the estimates of the C-index are different when the three options are classified as functional ability or performing state.

Although all participants were ineligible for LTCI certification at baseline, some of those who responded to the self-administered questionnaire who needed assistance with daily living and those who did not respond. To confirm the robustness of the results, we performed a sensitivity analysis excluding those who reported needing daily living assistance or those who did not respond. In addition, we conducted follow-up after periods of up to 1 and 2 years to compare with previous studies. Furthermore, local governments often judge unanswered questions as applicable.[22] Therefore, as a sensitivity analysis, the participants who did not answer the KCL completely were included in the analysis, and the unanswered questions were considered as applicable.

All statistical analyses were conducted using Stata/MP 17.0 (Stata Corp, College Station, TX, USA), with $P < 0.05$ indicating statistical significance.

Results

The mean age (standard deviation: SD) was 73.6 (6.0) years. During a mean follow-up period of 3.1 years, the incidence of functional disability in 6232 participants was 9.2%. Table 1 lists the prevalence of each applicable question, the impairment rate, and the incidence of functional disability for the 25 modified KCL items. The incidence of functional disability was significantly higher for all questions.

Table 2 shows the results of examining the applicability and incidence of functional disability for each of the nine indicators. The adjusted HRs for sex and age were calculated for each of the nine indicators, and the HR ranged from 1.50 to 3.82, indicating that all indicators were significantly associated with the incidence of functional disability.

Table 3 shows the C-index of the relationship between the incidence of functional disability and the total score of the nine indicators. The C-index was 0.54–0.81. Consequently, the C-index was significantly higher in three out of four indicatiors, except for the risk assessment scale indicators, when classified by performing state compared with when classified by functional ability.

Table 4 shows the sensitivity, specificity, and positive predictive values of each indicator for the incidence of functional ability. Sensitivity ranged from 3.9% in the nutritional domain to 79.7% in the risk assessment scale. Specificity ranged from 69.0% in the memory domain to 98.8% in the nutrition domain. Positive predictive values ranged from 13.1% in the memory domain to 47.4% in the multi-domain.

The sensitivity analysis excludes those who responded that they required assistance at baseline and those who did not respond. Compared with the main results, the C-index tended to decrease, but the trend of the results remained the same (Table S1). The results of the C-index with a maximum follow-up period of 1 and 2 years were also given (the larger the C-index, the shorter the follow-up period; Table S2). Furthermore, a sensitivity analysis was performed, which included non-respondents, on the items of the KCL, and non-respondents' items were treated as applicable. The results show that the C-index improved for all indicators except the risk assessment scale and the nutrition domain; sensitivity improved for all indicators; and specificity worsened for all indicators; however, the overall trend was the same.

Discussion

We examined the predictive validity of the original KCL's criteria for the incidence of functional disability by dividing it into functional ability and performing state in the modified KCL. The results suggested that the two classifications using the modified KCL had the predictive validity of incidence of functional disability. Regarding the handling of the three options, the predictive discrimination was higher when the evaluation was focused on performing state rather than on functional ability.

Table 2 Incidence of functional disability in nine indicators (Cox proportional hazard ratio)

	Total number of people $n = 67\,398$	Proportion (%)	Cumulative incidence	Crude			Model 1		
				HR	95% CI	*P*-value	HR	95% CI	*P*-value
Evaluating "functional ability" (including indicators that do not have a three-option question)									
Frailty									
Robust	40 171	59.6	4.7	1.00			1.00		
Pre-frailty	19 108	28.4	11.2	2.49	2.34–2.65	<0.001	1.89	1.77–2.01	<0.001
Frailty	8119	12.0	26.9	6.83	6.42–7.27	<0.001	3.82	3.58–4.08	<0.001
Risk assessment scale									
≤16	49 424	73.3	3.6	1.00			1.00		
≥17	17 974	26.7	24.9	8.15	7.71–8.61	<0.001	3.65	3.33–4.00	<0.001
Multi-domain									
Not applicable	65 930	97.8	8.4	1.00			1.00		
Applicable	1468	2.2	47.4	8.32	7.68–9.00	<0.001	3.64	3.35–3.96	<0.001
Physical domain									
Not applicable	62 948	93.4	7.6	1.00			1.00		
Applicable	4450	6.6	32.9	5.34	5.03–5.66	<0.001	2.82	2.65–3.00	<0.001
Nutrition domain									
Not applicable	66 401	98.5	9.0	1.00			1.00		
Applicable	997	1.5	24.7	3.20	2.81–3.63	<0.001	2.29	2.01–2.60	<0.001
Eating domain									
Not applicable	56 836	84.3	8.1	1.00			1.00		
Applicable	10 562	15.7	15.5	2.04	1.93–2.16	<0.001	1.50	1.42–1.59	<0.001
Houseboundness domain									
Not applicable	65 241	96.8	8.6	1.00			1.00		
Applicable	2157	3.2	27.8	3.85	3.54–4.19	<0.001	2.26	2.08–2.47	<0.001
Memory domain									
Not applicable	45 605	67.7	7.4	1.00			1.00		
Applicable	21 793	32.3	13.1	1.85	1.76–1.94	<0.001	1.56	1.48–1.64	<0.001
Mood domain									
Not applicable	53 666	79.6	7.1	1.00			1.00		
Applicable	13 732	20.4	17.6	2.69	2.55–2.83	<0.001	1.97	1.87–2.07	<0.001
Evaluating "performing state"									
Frailty									
Robust	32 147	47.7	4.1	1.00			1.00		
Pre-frailty	22 946	34.0	9.0	2.29	2.14–2.45	<0.001	1.79	1.67–1.92	<0.001
Frailty	12 305	18.3	23.2	6.62	6.20–7.06	<0.001	3.77	3.53–4.04	<0.001
Risk assessment scale									
≤16	43 586	64.7	2.9	1.00			1.00		
≥17	23 812	35.3	20.9	8.16	7.67–8.68	<0.001	3.27	2.98–3.59	<0.001
Multi-domain									
Not applicable	64 347	95.5	7.9	1.00			1.00		
Applicable	3051	4.5	37.9	6.34	5.94–6.75	<0.001	3.18	2.97–3.40	<0.001
Physical domain									
Not applicable	53 962	80.1	6.3	1.00			1.00		
Applicable	13 436	19.9	20.9	3.65	3.47–3.84	<0.001	2.33	2.21–2.46	<0.001

Model 1, adjusted for sex and age.

95% CI, 95% confidence interval; HR, hazard ratio.

Table 3 Comparison between functional ability and performing state of the C-index for each indicator ($n = 67\,398$)

	Evaluating "functional ability" (including indicators that do not have three-option questions)		Evaluating "performing state"		*P*-value*
	C-index	95% CI	C-index	95% CI	
Frailty	0.717	0.710–0.724	0.722	0.715–0.728	<0.001
Risk assessment scale	0.813	0.808–0.819	0.813	0.808–0.818	0.590
Multi-domain	0.712	0.706–0.719	0.718	0.711–0.724	<0.001
Physical domain	0.686	0.679–0.692	0.694	0.688–0.701	<0.001
Nutrition domain	0.557	0.551–0.563			
Eating domain	0.586	0.579–0.593			
Houseboundness domain	0.535	0.532–0.539			
Memory domain	0.586	0.579–0.593			
Mood domain	0.633	0.627–0.640			

*Statistical analysis was performed to compare the C-index of functional ability and performing state.

95% CI, 95% confidence interval; C-index, Harrell's C statistics.

Table 4 Sensitivity, specificity, and positive predictive value for the incidence of functional disability in each indicator ($n = 67\,398$)

	Evaluating "functional ability" (including indicators that do not have three-option questions)			Evaluating "performing state"		
	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Sensitivity (%)	Specificity (%)	Positive predictive value (%)
Frailty	35.1	90.3	26.9	45.8	84.6	23.2
Risk assessment scale	71.8	77.9	24.9	79.7	69.2	20.9
Multi-domain	11.2	98.7	47.4	18.6	96.9	37.9
Physical domain	23.5	95.1	32.9	45.0	82.6	20.9
Nutrition domain	3.9	98.8	24.7			
Eating domain	26.3	85.4	15.5			
Houseboundness domain	9.6	97.5	27.8			
Memory domain	45.9	69.0	13.1			
Mood domain	38.7	81.5	17.6			

The results of this study support those of previous studies that predicted incidence of functional disability based on the original KCL.[7,9,14–16] The sensitivity analysis results, which excluded those who required care or who did not respond at baseline for robustness, also indicated the same trend. Therefore, we concluded that the predictive validity of incidence of functional disability was confirmed in the modified KCL.

The AUC, which indicates the predictive discrimination of incidence of functional disability assessed by the indicator for frailty, was 0.76 (95% CI: 0.74–0.79)[9] in a previous study. The C-index of 0.72 (0.71–0.72), which indicates predictive discrimination, was lower in the present study compared with a previous study. Similarly, the AUC was reported for seven domains of the MHLW ranged from 0.62 (95% CI: 0.59–0.65) to 0.83 (0.81–0.85) and sensitivity from 13.5% to 71.4%.[15] While the present study indicated the C-index ranged from 0.54 (0.53–0.54) to 0.71 (0.71–0.72) and sensitivity from 3.9% to 45.9%, all values are lower in the present study. However, specificity and positive predictive value were higher in this study than in previous studies. Therefore, false negatives are more likely to appear and false positives are less likely in the current study than in previous studies. There are two reasons for this difference in predictive discrimination, sensitivity, specificity, and positive predictive value. First, the follow-up period of this study was longer, averaging 3.1 years, compared with 1 year in previous studies.[15] Tables S2 and S3 show that values of the C-index and sensitivity increased when the follow-up period was shortened. In contrast, specificity and positive predictive value decreased when the follow-up period was shortened, approaching the values of a previous study.[15] Second, the difference in the target population of the analysis influences the level of predictive discrimination. In the current study, only those who responded to all 25 questions of the modified KCL were included in the analysis; however, previous studies included those who did not respond to all questions in the analysis and treated the unanswered questions as corresponding to functional decline.[15] Other studies have included all respondents to the original KCL but used telephone or face-to-face contact to obtain answers for previously unanswered questions.[9] Furthermore, previous studies have indicated that the risk of death is higher for those who do not answer the questions.[23] In sensitivity analyses in which unanswered questions were considered applicable, the C-index and sensitivity scores improved and the specificity scores decreased, approaching those of previous studies.

In this study, we conducted two forms of assessments, one for functional ability and one for performing state. The C-index was higher when the focus was on performing state than when it was on functional ability, except for the risk assessment scale. Functional ability (can/cannot) may differ from performance status (do/do not) of activities of daily living, while performing state asks about actual performance status, which may be directly reflected in

physical activity. Specifically, no. 8, "Do you normally walk continuously for 15 min?" asks for physical activity. Furthermore, nos. 1–3 are IADL items, and IADL-independent persons are known to have high levels of physical activity.[24] Studies examining the relationship between physical activity and disability[25] and all-cause mortality[26] have reported a dose–response relationship, with higher levels of physical activity associated with lower levels of disability and all-cause mortality. One study compared functional ability and performing state in some of the questions of the KCL among community-dwelling older adults.[27] It was reported that performing state scores were lower than scores for functional ability; the results of this study support those of previous studies. The International Classification of Functioning, Disability, and Health proposed by the WHO[28] explains that functional ability is evaluated under a standardized evaluation setting, while performing state is evaluated in various different environments, including with respect to the individual's socioeconomic status. Predictive discrimination may have been higher in the performing state, reflecting differences in individual environment and physical activity.

The fact that only the risk assessment scale failed to present a difference in the C-index by the classification method is partly explained by its inclusion of sex and age. Actually, when the C-index of the risk assessment scale excluding sex and age was calculated, it was also higher in the performing state; that is, 0.72 for performing state and 0.71 for functional ability ($P = 0.016$).

However, the predictive validity of all nine indicators for incidence of functional disability was confirmed for the functional ability and performing state classification methods. Therefore, it is recommended that the classification method should be based on purpose, such as classification by functional ability when identifying high-risk older people, or classification by performing state when extracting a wide range of high-risk older people.

The strength of this study derives from its confirmation of the predictive validity of incidence of functional disability using the modified KCL used in epidemiological investigation among 67 398 older people living in 18 municipalities. This indicator is used by municipalities to assess the risk of developing LTCI and to identify local issues.[11] It enables municipalities to understand the risk of LTCI with existing data. However, the current study also includes the following limitations. First, it is limited to participants who answered all 25 questions of the KCL, and thus there is a selection bias towards a relatively healthy population. However, a similar trend was confirmed in the sensitivity analysis of the questions in the KCL left unanswered in the analysis. Second, the incidence of functional disability in this study was defined by LTCI, which is a voluntary reporting system, and not all people with reduced ADL apply for LTCI.

In conclusion, the predictive validity of the modified KCL for the incidence of functional disability was confirmed. However, some indicators might show false negatives with high specificity and low sensitivity. Although predictive discrimination was slightly higher for the performing state than for functional ability, the predictive validity of the modified KCL was confirmed for both. Therefore, it is essential to choose the classification appropriately based on the purpose for which the modified KCL is being used.

Acknowledgements

This study used data from JAGES (the Japan Gerontological Evaluation Study). This study was supported by grants from the JSPS (Japan Society for the Promotion of Science; KAKENHI grant JP15H01972, JP20H03954), Health Labour Sciences (research grant H28-Choju-Ippan-002), the Japan Agency for Medical Research and Development (AMED) (JP18dk0110027, JP18ls0110002, JP18le0110009, JP20dk0110034, JP21lk0310073, JP21dk0110037), and the Open Innovation Platform with Enterprises, Research Institute and Academia (OPERA, JPMJOP1831) from the Japan Science and Technology (JST). It also received grants from the Innovative Research Program on Suicide Countermeasures (1–4), the Sasakawa Sports Foundation, the Japan Health Promotion & Fitness Foundation, the Chiba Foundation for Health Promotion & Disease Prevention, the 8020 Research Grant for fiscal 2019 from the 8020 Promotion Foundation (adopted number: 19-2-06), and grants from the Meiji Yasuda Life Foundation of Health and Welfare and the Research Funding for Longevity Sciences from National Center for Geriatrics and Gerontology (29-42, 30-22, 20-19, 21-20, 21-19).

Disclosure statement

The authors declare no conflict of interest.

Data availability statement

Data are from the JAGES study. All enquiries are to be addressed at the data management committee via e-mail: dataadmin.ml@jages.net. All JAGES datasets have ethical or legal restrictions for public deposition due to inclusion of sensitive information from the human participants.

References

1 World Health Organization. Decade of healthy ageing: Plan of Action 2020 [Cited 28 Dec 2021]. Available from URL: https://cdn.who.int/media/docs/default-source/decade-of-healthy-ageing/final-decade-proposal/decade-proposal-final-apr2020-en.pdf?sfvrsn=b4b75ebc_25&download=true
2 World Health Organization. World Report on Ageing and Health, 2015 [Cited 28 Dec 2021]. Available from URL: https://www.who.int/publications/i/item/9789241565042
3 Arai H, Ouchi Y, Toba K *et al*. Japan as the front-runner of super-aged societies: perspectives from medicine and medical care in Japan. *Geriatr Gerontol Int* 2015; **15**: 673–687.
4 Ministry of Health, Labour and Welfare. Preventing Long-Term Care Manual, Revised edition; 2012 (In Japanese). [Cited 28 Dec 2021]. Available from URL: https://www.mhlw.go.jp/topics/2009/05/dl/tp0501-1_1.pdf
5 Ministry of Health, Labour and Welfare. Utilize of the Kihon Checklist; 2005 (In Japanese). [Cited 28 Dec 2021]. Available from URL: https://www.wam.go.jp/wamappl/bb05Kaig.nsf/0/842faf619ff7dbee492570dc0023a07f/$FILE/siryou2.pdf
6 Sewo Sampaio PY, Sampaio RA, Yamada M, Arai H. Systematic review of the Kihon checklist: is it a reliable assessment of frailty? *Geriatr Gerontol Int* 2016; **16**: 893–902.
7 Tsuji T, Kondo K, Kondo N, Aida J, Takagi D. Development of a risk assessment scale predicting incident functional disability among older people: Japan Gerontological evaluation study. *Geriatr Gerontol Int* 2018; **18**: 1433–1438.
8 Tsuji T, Takagi D, Kondo N, Kondo K. Development of risk assessment scales for needed support/long-term care certification: a longitudinal study using the Kihon checklist and medical assessment data. *Nihon Koshu Eisei Zasshi* 2017; **64**: 246–257 (In Japanese).
9 Satake S, Shimokata H, Senda K, Kondo I, Toba K. Validity of Total Kihon checklist score for predicting the incidence of 3-year dependency and mortality in a community-dwelling older population. *J Am Med Dir Assoc* 2017; **18**: e551–552.e556.
10 Ministry of Health, Labour and Welfare. Briefing on Various Surveys, for the Formulation of Long-Term Care Insurance Business Plans; 2016 (In Japanese). [Cited 28 Dec 2021]. Available from URL: https://www.mhlw.go.jp/file/05-Shingikai-12301000-Roukenkyoku-Soumuka/0000142077.pdf
11 Ministry of Health, Labour and Welfare. *Instructions of Conducting the Public Survey of Long-Term Care Prevention and Needs in Spheres of Daily Life (the Version of 23 October, 2019)*; 2019 (In Japanese). [Cited 28 Dec 2021].

Available from: https://www.mhlw.go.jp/content/12301000/000560423.pdf
12 Kondo K. Progress in aging epidemiology in Japan: the JAGES project. *J Epidemiol* 2016; **26**: 331–336.
13 Kondo K, Rosenberg M. World Health Organization. *Advancing universal health coverage through knowledge translation for healthy ageing: lessons learnt from the Japan gerontological evaluation study*. Geneva: World Health *Organization,* 2018.
14 Kamegaya T, Yamaguchi H, Hayashi K. Evaluation by the basic checklist and the risk of 3 years incident long-term care insurance certification. *J Gen Fam Med* 2017; **18**: 230–236.
15 Tomata Y, Hozowa A, Ohomori-Matsuda K *et al*. Validation of the Kihon checklist for predicting the risk of 1-year incident long-term care insurance certification: the Ohsaki cohort 2006 study. *Nihon Koshu Eisei Zasshi* 2011; **58**: 3–13 (In Japanese).
16 Hamazaki Y, Morikawa Y, Nakamura K, Morimoto S, Nakagawa H. Relationship between participation in a secondary screening test for the frail elderly and new onset of the need for long-term care among the elderly in a rural town. *Nihon Koshu Eisei Zasshi* 2012; **59**: 801–809 (In Japanese).
17 Tsutsui T, Muramatsu N. Japan's universal long-term care system reform of 2005: containing costs and realizing a vision. *J Am Geriatr Soc* 2007; **55**: 1458–1463.
18 Satake S, Senda K, Hong YJ *et al*. Validity of the Kihon checklist for assessing frailty status. *Geriatr Gerontol Int* 2016; **16**: 709–715.
19 Fried LP, Tangen CM, Walston J *et al*. Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci* 2001; **56**: M146–M156.
20 Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996; **15**: 361–387.
21 Newson RB. Comparing the predictive powers of survival models using Harrell's C or Somers' D. *The Stata Journal* 2010; **10**: 339–358.
22 Ministry of Health, Labour and Welfare. *Instructions of Community General Support Center*; 2007 (In Japanese). [Cited 25 Mar 2021]. Available from: https://www.mhlw.go.jp/topics/2007/03/tp0313-1.html
23 Tolonen H, Laatikainen T, Helakorpi S, Talala K, Martelin T, Prättälä R. Marital status, educational level and household income explain part of the excess mortality of survey non-respondents. *Eur J Epidemiol* 2010; **25**: 69–76.
24 Crevenna R, Dorner TE. Association between fulfilling the recommendations for health-enhancing physical activity with (instrumental) activities of daily living in older Austrians. *Wien Klin Wochenschr* 2019; **131**: 265–272.
25 Chen T, Honda T, Chen S, Narazaki K, Kumagai S. Dose-response association between accelerometer-assessed physical activity and incidence of functional disability in older Japanese adults: a 6-year prospective study. *J Gerontol A Biol Sci Med Sci* 2020; **75**: 1763–1770.
26 Ekelund U, Tarp J, Steene-Johannessen J *et al*. Dose-response associations between accelerometry measured physical activity and sedentary time and all cause mortality: systematic review and harmonised meta-analysis. *BMJ* 2019; **366**: l4570.
27 Suzuki N, Makigami K, Goto A, Yokokawa H, Yasumura S. Comparison of ability-based and performance-based IADL evaluation of community-dwelling elderly using the Kihon checklist and TMIG index of competence. *Nihon Ronen Igakkai Zasshi* 2007; **44**: 619–626.
28 World Health Organization. *The International Classification of Functioning, Disability and Health (ICF)*. 2001 [Cited 25 Mar 2021]. Available from: https://www.who.int/classifications/international-classification-of-functioning-disability-and-health

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's website:

Table S1. Comparison between functional ability and performing state of the C-index for each indicator, and sensitivity analysis excluding those with potential for care at baseline.
Table S2. Comparison between functional ability and performing state of the C-index for each indicator and sensitivity analysis for different follow-up periods.
Table S3. Sensitivity specificity, positive predictive value for the incidence of functional disability in each indicator and sensitivity analysis for different follow-up periods.

How to cite this article: Watanabe R, Tsuji T, Ide K, et al. Predictive validity of the modified Kihon Checklist for the incidence of functional disability among older people: A 3-year cohort study from the JAGES. Geriatr. Gerontol. Int. 2022;1–8. https://doi.org/10.1111/ggi.14439