THE GERONTOLOGICAL SOCIETY OF AMERICA®

Journals of Gerontology: Medical Sciences
cite as: *J Gerontol A Biol Sci Med Sci*, 2022, Vol. XX, No. XX, 1–9
https://doi.org/10.1093/gerona/glac059
Advance Access publication March 10, 2022

OXFORD

Research Article

Development and Validation of the Disability Index Among Older Adults

Xuxi Zhang, PhD,[1] Lizi Lin, PhD,[2] Xinying Sun, PhD,[3] Xiaoyan Lei, PhD,[1] Gordon G. Liu, PhD,[4] Hein Raat, MD, PhD,[5] and Yi Zeng, PhD[1,6,*]

[1]Center for Healthy Aging and Development Studies, National School of Development, Peking University, Beijing, China. [2]Guangdong Provincial Engineering Technology Research Center of Environmental Pollution and Health Risk Assessment, Department of Occupational and Environmental Health, School of Public Health, Sun Yat-sen University, Guangzhou, China. [3]Public Heath School, Health Science Center, Peking University, Beijing, China. [4]National School of Development, Peking University, Beijing, China. [5]Department of Public Health, Erasmus University Medical Center, Rotterdam, The Netherlands. [6]Center for the Study of Aging and Human Development, Duke University Medical School, Durham, North Carolina, USA.

*Address Correspondence to: Yi Zeng, PhD, Center for the Study of Aging and Human Development, Duke University Medical School, 40 Duke Medicine Cir, Durham, NC 27710, USA. E-mail: zengyi@nsd.pku.edu.cn

Received: October 24, 2021; Editorial Decision Date: February 21, 2022

Decision Editor: Lewis A. Lipsitz, MD, FGSA

Abstract

Background: With the rapid population aging, the challenge to provide care for disabled older adults is becoming bigger. This study aims to develop the Disability Index (DI) to assess disability to indicate care needs, and we evaluated the reliability and validity of the DI among older adults aged 65–105.

Methods: A total of 12 559 older adults (54.0% women; mean age = 84.3; *SD* = 11.2) from 22 provinces in China were investigated in 2017–2018. We developed the 21-item DI covering 4 subdomains, including Activities of Daily Living (ADL), Instrumental Activities of Daily Living (IADL), physical performance, and cognitive function. Cronbach's alpha was used to determine internal consistency. The convergent and divergent validity were assessed by Spearman rank order correlation coefficients and Pearson correlation coefficients. The known-group validity was assessed by Mann–Whitney *U* tests. The concurrent validity was assessed by the area under the receiver operating characteristic curve (AUC).

Results: In the total sample, the internal consistency of the full DI and its subdomains were satisfactory with Cronbach's alpha ≥0.70; the convergent and divergent validity of the 4 subdomains were supported by all the alternative measures; the known-group validity of the full DI and its subdomains were supported by clear discriminative ability; and the concurrent validity of the full DI was supported with all the AUCs ≥0.70. The reliability and validity of the full DI and its subdomains were additionally supported by age subgroups and sex subgroups.

Conclusions: The DI is a reliable and valid instrument to assess disability status among older adults.

Keywords: Activities of Daily Living, CLHLS, Cognitive function, Instrumental Activities of Daily Living, Physical performance

The prevalence rate of disability rises with age, and the increase is most prominent in the oldest (1). As the proportion of older people in the population continues to rise, disability is anticipated to become a major society issue (2). The challenge to provide long-term care services for disabled older adults is becoming bigger (2). Analyses and projections of care needs for disabled older adults is a crucial basis for governments and policymakers to develop optimum coping strategies to face the challenges of rapid population aging. It is widely accepted that care needs are closely related to the disability status (3). Therefore, the measurement of disability status among older adults plays an important role in public health policy studies.

From the perspective of care needs, there are 4 relevant dimensions regarding disability among older adults. First, as the direct indicator of the need for care, the requirement of assistance with basic Activities of Daily Living (ADL) is the most frequent approach used to define the disabled older adults (4). Second, Instrumental Activities of Daily Living (IADL) is additionally used to measure the level of functional impairments (5,6). Third, indicators of objectively-tested

physical performance, such as standing up from a chair, picking up a book from the floor, and turning around 360°, have additional predictive value beyond self-reported ADL/IADL in the evaluation of actual physical health changes among older adults (7). Fourth, cognitive functioning is a widely used measure of disability from the perspective of mental health (6). In many cases, the barrier of older adults to perform ADL and/or IADL independently is not their physical capability but their cognitive impairments (6,8).

Most of the widely used disability measurements, such as Barthel Index (BI), Groningen Activity Restriction Scale (GARS), and Lawton and Brody Instrumental Activities of Daily Living Scale (LB-IADL), are based on information from self-reported questionnaires focusing on ADL and/or IADL only (9). Self-reported ADL and IADL are based on respondents' perceptions of their own functioning (9), which is subject to the subjective, personal interpretation of the older person. Performance-based measurements that use objectively-tested physical performance and cognitive function generally show less measurement errors (10). In general, performance-based measurements have better face validity, reproducibility and have a higher sensitivity to change (9,11). A combination of both self-reported and performance-based measurements can be utilized in better characterizing the degree of disability in daily life (11,12). However, to the best of our knowledge, there is no measurement yet to assess the disability status of older adults from the perspective of care needs covering the four relevant dimensions, and including both self-reported ADL and IADL as well as objectively-tested physical performance and cognitive function.

Therefore, in the current study, first, we developed the Disability Index (DI) to assess disability including ADL, IADL, physical performance, and cognitive function to indicate care needs. Second, we evaluated the reliability and validity of the DI among older adults aged 65 to 105, including the internal consistency, convergent and divergent validity, known-group validity, and concurrent validity. Furthermore, we explored the reliability and validity for distinct age and sex subgroups separately.

Materials and Methods

Disability Index

Based on experiences of the Frailty Index (FI) and the Successful Aging Index (SAI) (13,14), we developed the DI with 21 items following the principle of FI and focused on 4 domains including the ADL, IADL, physical performance, and cognitive function as mentioned in the Introduction. The principle of the FI is to count deficits in health, functioning, and vulnerability; the assumption is that the more deficits a person has, the more likely that person is to be frail (14). The SAI (13) adopted the principle of FI but narrowed down the items to focus on successful aging, which was shown to be relevant for identifying components of successful aging (15).

Table 1 presents the details on the corresponding items of the 4 subdomains of the DI. The ADL was assessed by a self-reported scale on performing 6 basic daily activities (16,17), including (a) bathing, (b) dressing, (c) toilet, (d) indoor transfer, (e) continence, and (f) eating. The IADL was assessed by a self-reported scale on performing 8 activities that are important for independent living (16,17), including (a) visiting neighbors, (b) going shopping, (c) preparing and cooking meal, (d) doing laundry, (e) walking 2 km continuously, (f) carrying 5 kg, (g) squatting and standing up 3 times, and (h) going out by public transportation. The domain of physical performance was measured by 6 objective examinations (7,16), including (a) picking up a book from the floor, (b) standing up from a chair, (c) turning around 360°, (d) putting hand behind neck, (e) putting hand behind lower back, and (f) raising arms upright. The domain of cognitive function was measured by the total score of the Chinese version of the Mini-Mental State Examination (MMSE) that includes 24 items regarding orientation, naming foods, registration, attention and calculation, copy a figure, recall, and language, with a total score ranging from 0 to 30 (18). Supplementary Table 1 presents the details on items and coding methods of the MMSE. According to previously widely accepted cutoff points of the MMSE total score (13,19), we classified cognitive functions into 4 categories

Table 1. The Disability Index (DI), Number of Items, Description of Items, Coding Methods of Responses, and Score Calculation

Subdomains of the Disability Index (DI)	Number of Items	Description of Items	Score Range	Score Description
Activities of daily living (ADL)	6	Six items including bathing, dressing, toilet, indoor transfer, continence, and eating	0–100	A higher score represents a higher level of ADL disability
Instrumental activities of daily living (IADL)	8	Eight items including visiting neighbors independently, going shopping independently, preparing and cooking meal independently, doing laundry independently, walking 2 km continuously, carrying 5 kg, squatting and standing up 3 times, and going out by public transportation independently	0–100	A higher score represents a higher level of IADL disability
Physical performance	6	Six items including picking up a book from the floor, standing up from a chair, turning around 360°, putting hand behind neck, putting hand behind lower back, and raising arms upright	0–100	A higher score represents a higher level of physical performance disability
Cognitive function	1	The item derived from the total score of the Chinese version of the Mini-Mental State Examination (MMSE), which is a 24-item scale assessing different domains of cognitive functioning including calculation, language, orientation, and recall	0–100	A higher score represents a higher level of cognitive disability
Full DI	21	Supplementary Table 2 presents the detailed coding methods and score calculation of the DI as well as its subdomains	0–100	A higher score represents a higher level of general disability status

including severe impairment (0~17), moderate impairment (18~21), mild impairment (22~25), and no cognitive impairment (26~30).

Considering the limitations of a binary framework, the DI was developed as a continuum-based index with higher resolution and greater granularity (13). Ordinal items were given values, and the scores from each item were combined to form a continuous value (13). Fraction values ranged 0–100 were assigned according to the number of potential answers. For example, in a 3-point scale, the lowest level of functioning would receive 100, the middle level 50, and the highest level 0. In the DI, items of ADL, IADL, and physical performance domains were measured on a 3-point scale ranging from "able to perform the activity without difficulty" to "unable to perform the activity independently," with an exception of item 4–6 in the physical performance domain which ranged from "both hands can perform the activity" to "neither hand can perform the activity." Each item from the ADL, IADL, and physical performance domains was converted into its own index, averaged, then placed into the full DI as a single item. The 4 categories of the MMSE total score were used to assigned fractional values with 100 for severe impairment (score range 0~17), 67 for moderate impairment (18~21), 33 for mild impairment (22~25), and 0 for no cognitive impairment (26~30) to obtain the item score of cognitive function domain. A total score of the full DI (21 items) between 0 and 100 was calculated by averaging the items, with higher values indicating a higher level of disability. Supplementary Table 2 presents the detailed coding and calculation methods of the DI.

Data Resource and Study Population

For evaluating the reliability and validity of the DI, we used the Chinese Longitudinal Healthy Longevity Survey (CLHLS) that has been conducted between 1998 and 2021. The CLHLS is a nationwide cohort study conducted in community settings of randomly selected half of the counties and cities in 23 of 31 provinces covering 85% of the total population in China (7). Data were collected with the CLHLS questionnaire of both self-reported and performance-based measurements at baseline and the follow-up surveys. For objective questions such as physical performance and cognitive function, no proxy was adopted. Trained interviewers performed the survey at the participants' homes. Ethical committee procedures have been followed in all study sites and approval has been provided by the Biomedical Ethics Committee, Peking University (IRB00001052-13074) (7,20). Written informed consent was obtained from all participants or their legal representatives for participation at baseline and the follow-up surveys. More information regarding the CLHLS has been described in detail elsewhere (7,21).

In the current study, we adopted a cross-sectional design and used data of 15 498 participants aged 65–105 years old from the eighth wave of the CLHLS conducted in 2017–2018. Participants with missing data on one or more items of the DI (N = 2 939) were excluded. Thus, the final analyses included 12 559 participants.

In addition, we divided the total sample into 4 age subgroups, including 65–79 years old (n = 4 704), 80–89 years old (n = 3 318), 90–99 years old (n = 2 743), and 100–105 years old (n = 1 791).

Measures Used for the Evaluation of the DI

Self-rated health was assessed by 1 item with a 5-point Likert scale (22) with 1 representing "very good" and 5 representing "very bad"; the score of self-rated health ranges from 1 to 5 with a higher score representing worse self-rated health.

Cognitive score was assessed by 7 items that were not included in the MMSE. Among which, 4 items were from Community Screening Instrument for Dementia (CIS-D) (23) including questions describing the use of a hammer, naming an elbow, locating the nearest store and pointing to the window, and 3 items were from the Telephone Interview for Cognitive Status (TICS-M) (24) including naming the tool of cutting paper, naming the plant, and knowledge of the President of China. The cognitive score ranges from 0 to 7 with a higher score representing a higher level of cognitive function.

Informal care was assessed by the care hours per day provided by family members who take care of participants' daily life, ranging from 0 to 24 (25). A cutoff point of 8 hours/day was adopted according to a previous study on informal care hours (26). Participants receiving informal care greater than 8 hours/day were categorized as "heavy informal care burden."

Activity limitation was assessed by 1-item Global Activity Limitation Index (GALI) (27). Participants who indicated their function to be limited or severely limited were classified as having a "limited function."

Health during the past 2 weeks was assessed by one question of "Have you felt not well regarding your health during the past 2 weeks?" Participants who indicated not feeling well were classified as "Feel not well regarding health."

Sociodemographic Factors

Age (in years), sex (male/female), educational level (without schooling/1–6 years of education/>6 years of education), current marital status (married/divorced or separated/widowed/never married), residence (urban/rural), and living situation (living alone/living with family members) were assessed.

Statistical Analyses

Descriptive statistics were used to describe the scale scores (28). Reliability was measured by the internal consistency with the Cronbach's alpha; a Cronbach's alpha value of 0.7 or above was considered satisfactory reliability of internal consistency (29).

To examine the convergent and divergent validity, we hypothesized that self-rated health is stronger related to the IADL and physical performance domains, and less to the other 2 domains. We hypothesized that the informal care from family members is stronger related to the ADL and IADL domains, but less so to the other 2. We also hypothesized that the cognitive score is strongly related to the cognitive function domain, but less so to the other 3. A satisfactory convergent validity was defined as a statistically significant correlation between a domain score and the score of a relevant alternative measure of the same domain, with a greater correlation indicating stronger validity (30). If each alternative measure exhibited a greater correlation with the related DI subdomain(s) but a lower correlation with the other DI subdomains, divergent validity was inferred (30). The convergent and divergent validity were assessed using Spearman rank order correlation coefficients (self-rated health) (31) and Pearson correlation coefficients (informal care hours from family members and cognitive score) (30).

To assess the known-group validity, the ability of the DI to discriminate between groups differing with regard to receiving informal care, presence of activity limitation and feeling not well regarding health, Mann–Whitney U tests were applied (31,32). We hypothesized that relatively higher scores of the full DI and each subdomain will be present in the subgroups characterized by (a) *Heavy informal*

care burden, (b) *Limited function*, and (c) *Feel not well regarding health*.

We utilized the following alternative measures as the criterion for examining concurrent validity of the full DI: (a) receiving informal care and (b) the presence of activity limitation. The receiver operating characteristic (ROC) curve analysis was used to assess the concurrent validity (30). Accuracy was measured by the area under the ROC curve (AUC). AUCs of 0.7–0.8 were considered acceptable, 0.8–0.9 excellent, and greater than 0.9 outstanding (33).

All analyses were conducted among the total sample as well as by age subgroups. In addition, we conducted all the analyses stratified by sex subgroups and found similar results (Supplementary Tables 5–8). All analyses were performed with SPSS version 24.0 (IBM SPSS Statistics for Windows, Armonk, NY: IBM Corp). The level of significance was p value < .05.

Results

Participant Characteristics

Table 2 presents the demographics of the total sample as well as the age subgroups' subsamples. The mean age of the total sample was 84.3 (*SD* = 11.2) years and 54.0% were women.

Score Distributions

Table 3 presents the score distributions of the DI and its subdomains. A floor effect (>25% of the respondents had the lowest possible score (31)) was observed in the full DI (total sample and 65–79 years old), ADL (total sample and all age subgroups), IADL (total sample, 65–79 years old and 80–89 years old), physical performance (total sample and all age subgroups except 100–105 years old), and cognitive function (total sample and all age subgroups except 100–105 years old) domains. A ceiling effect (>25% had the highest possible score (31)) was observed in the IADL (100–105 years old) and cognitive function (90–99 years old and 100–105 years old) domains.

Internal Consistency

Table 3 presents the internal consistency of the DI and its domains. The Cronbach's alpha of the full DI and the ADL, IADL, and physical performance domains were 0.94, 0.87, 0.95, and 0.72, respectively in the total sample. The Cronbach's alpha of the full DI, as well as the ADL and IADL domains was ≥0.70 in each age subgroup. The Cronbach's alpha of the physical performance domain varied between 0.66 and 0.68 in all age subgroups. Supplementary Table 3 presents that the Cronbach's alpha of the MMSE was ≥0.70 in the total sample and each age subgroup.

Table 2. Characteristics of Participants (n = 12 559)

Characteristics	Total (n = 12 559)	Age Subgroups				p Value
		65–79 years (n = 4 704)	80-89 years (n = 3 318)	90-99 years (n = 2 743)	100-105 years (n = 1 791)	
Age (range 65–105 years)	84.3 ± 11.2	72.5 ± 4.3	84.1 ± 2.8	93.5 ± 2.8	101.5 ± 1.5	p < .01*
Female	6 788 (54.0%)	2 320 (49.3%)	1 701 (51.3%)	1 455 (53.0%)	1 312 (73.3%)	p < .01†
Educational level						p < .01†
Without schooling	5 007 (46.4%)	1 014 (24.8%)	1 299 (49.3%)	1 431 (59.8%)	1 263 (75.4%)	
1–6 years of education	3 630 (33.6%)	1 748 (42.8%)	903 (34.3%)	673 (28.1%)	306 (18.3%)	
>6 years of education	2 151 (19.9%)	1 322 (32.4%)	434 (16.5%)	289 (12.1%)	106 (6.3%)	
Current marital status						p < .01†
Married	5 253 (42.2%)	3 328 (71.4%)	1 345 (40.9%)	512 (18.9%)	68 (3.8%)	
Divorced or separated	271 (2.2%)	140 (3.0%)	78 (2.4%)	40 (1.5%)	13 (0.7%)	
Widowed	6 800 (54.7%)	1 138 (24.4%)	1 837 (55.8%)	2 144 (79.1%)	1 681 (94.9%)	
Never married	115 (0.9%)	57 (1.2%)	32 (1.0%)	16 (0.6%)	10 (0.6%)	
Urban residence	6 972 (55.5%)	2 593 (55.1%)	1 857 (56.0%)	1 531 (55.8%)	991 (55.3%)	p = .862†
Living alone	2 035 (16.9%)	608 (13.3%)	736 (23.4%)	503 (19.6%)	188 (11.0%)	p < .01[b]
Self-rated health (score range 1–5; the higher the worse)	2.55 ± 0.89	2.53 ± 0.90	2.60 ± 0.89	2.54 ± 0.88	2.50 ± 0.89	p < .01*
Cognitive score‡ (score range 0–7; the higher the better)	6.42 ± 1.07	6.83 ± 0.48	6.53 ± 0.83	6.00 ± 1.35	5.45 ± 1.57	p < .01*
Informal care from family members (hours/day; range 0–24)	3.89 ± 6.53	1.12 ± 3.48	2.35 ± 5.25	4.81 ± 7.40	7.89 ± 8.92	p < .01*
Heavy informal care burden (>8 hours/day)	1 281 (13.6%)	118 (3.6%)	206 (8.4%)	436 (19.7%)	521 (34.6%)	p < .01*
Limited function (GALI)	4 231 (33.8%)	893 (19.0%)	1 034 (31.2%)	1 243 (45.4%)	1 061 (59.4%)	p < .01†
Feel not well regarding health	1 978 (15.9%)	696 (15.0%)	560 (17.1%)	427 (15.8%)	295 (16.7%)	p = .059†

Notes: Presented as mean ± *SD* or *N* (%). GALI = one-item Global Activity Limitation Index. Missing items: educational level = 1 771; current marital status = 120; living alone = 553; self-rated health = 717; cognitive score = 2 814; informal care from family members = 3 131; limited function = 30; feeling not well = 153.

*p Value based on ANOVA.

†p Value based on Chi-square test.

‡Cognitive score assessed by 4 items from Community Screening Instrument for Dementia (CIS-D) and 3 items from the Telephone Interview for Cognitive Status (TICS-M).

Table 3. Score Distributions and Internal Consistency Reliability of the Disability Index (DI; *n* = 12 559)

Disability Index	Sample	Mean score ± *SD*	Range	% of Min*	% of Max†	25th %tile	50th %tile‡	75th %tile	Cronbach's Alpha§
Disability Index	Total	19.67 ± 23.65	0–95.83	27.7	0	0	8.33	33.33	**0.94**
(DI; 21 items)	65–79 years	4.71 ± 10.00	0–95.83	55.0	0	0	0	4.69	**0.89**
	80–89 years	15.39 ± 18.26	0–95.83	21.0	0	1.56	8.33	23.00	**0.91**
	90–99 years	31.59 ± 23.52	0–93.75	6.0	0	11.38	28.04	49.92	**0.92**
	100–105 years	48.69 ± 22.66	0–95.83	1.5	0	32.21	51.04	65.58	**0.91**
Activities of daily living domain	Total	7.38 ± 18.10	0–100	78.2	0.4	0	0	0	**0.87**
(ADL; 6 items)	65–79 years	1.13 ± 6.92	0–100	95.7	<0.1	0	0	0	**0.83**
	80–89 years	4.37 ± 13.59	0–100	84.6	0.2	0	0	0	**0.84**
	90–99 years	11.53 ± 21.20	0–100	64.4	0.8	0	0	16.67	**0.86**
	100–105 years	23.03 ± 27.40	0–100	41.3	0.8	0	16.67	41.67	**0.87**
Instrumental activities of daily living domain	Total	32.60 ± 37.06	0–100	36.1	12.3	0	12.50	62.50	**0.95**
(IADL; 8 items)	65–79 years	7.92 ± 17.96	0–100	68.0	1.3	0	0	6.25	**0.90**
	80–89 years	27.16 ± 30.85	0–100	30.3	5.8	0	12.5	43.75	**0.92**
	90–99 years	53.30 ± 35.70	0–100	9.8	20.0	18.75	50.00	87.50	**0.93**
	100–105 years	75.84 ± 29.82	0–100	3.0	41.8	56.25	87.50	100	**0.92**
Physical performance domain	Total	12.59 ± 17.76	0–91.67	52.7	0	0	0	16.67	**0.72**
(6 items)	65–79 years	4.25 ± 10.24	0–83.33	77.7	0	0	0	0	0.66
	80–89 years	10.34 ± 15.38	0–83.33	54.9	0	0	0	16.67	0.67
	90–99 years	19.01 ± 19.13	0–91.67	32.5	0	0	16.67	33.33	0.68
	100–105 years	28.87 ± 20.43	0–91.67	14.3	0	8.33	33.33	41.67	0.68
Cognitive function domain	Total	26.11 ± 38.71	0–100	63.0	17.5	0	0	33.00	N/A
(1 item derived from the total score of MMSE‖)	65–79 years	5.52 ± 17.21	0–100	88.3	1.5	0	0	0	N/A
	80–89 years	19.69 ± 32.59	0–100	66.9	9.6	0	0	33.00	N/A
	90–99 years	42.52 ± 42.81	0–100	42.4	29.9	0	33.00	100	N/A
	100–105 years	67.00 ± 40.99	0–100	20.8	55.2	33.00	100	100	N/A

Notes: SD = standard deviation.

*Percentage of respondents with the lowest possible score (floor).

†Percentage of respondents with the highest possible score (ceiling).

‡Median.

§A value of Cronbach's alpha of 0.7 or above represented satisfactory internal consistency reliability (29); **The value of Cronbach's alpha ≥0.7 in bold.**

‖The internal consistency of the Chinese version of the Mini-Mental State Examination (MMSE) can be found in Supplementary Table 3.

Table 4. Convergent and Divergent Validity: Correlations of the Disability Index (DI) and its Subdomains with the Alternative Measures

Score of Alternative Measures	Sample	Full DI Score		ADL Domain Score		IADL Domain Score		Physical Performance Domain Score		Cognitive Function Domain Score	
		r*	r†	r*	r†	r*	r†	r*	r†	r*	r†
Self-rated health	Total	0.193	N/A	0.112	N/A	**0.185**	N/A	0.176	N/A	0.123	N/A
(Score range 1–5)	65–79 years	0.255	N/A	0.131	N/A	**0.276**	N/A	0.176	N/A	0.106	N/A
(The higher the	80–89 years	0.277	N/A	0.170	N/A	**0.266**	N/A	0.227	N/A	0.163	N/A
worse)	90–99 years	0.245	N/A	0.135	N/A	**0.214**	N/A	0.206	N/A	0.181	N/A
	100–105 years	0.252	N/A	0.166	N/A	0.175	N/A	**0.255**	N/A	0.173	N/A
Informal care from	Total	N/A	0.475	N/A	0.454	N/A	**0.461**	N/A	0.388	N/A	0.335
family members	65–79 years	N/A	0.239	N/A	**0.223**	N/A	0.216	N/A	0.187	N/A	0.139
(hours/day)	80–89 years	N/A	0.371	N/A	**0.394**	N/A	0.351	N/A	0.279	N/A	0.212
(Score range 0–24)	90–99 years	N/A	0.361	N/A	**0.388**	N/A	0.358	N/A	0.291	N/A	0.173
	100–105 years	N/A	0.324	N/A	**0.333**	N/A	0.312	N/A	0.239	N/A	0.140
Cognitive score‡	Total	N/A	−0.589	N/A	−0.366	N/A	−0.469	N/A	−0.379	N/A	**−0.597**
(Score range 0–7)	65–79 years	N/A	−0.296	N/A	−0.139	N/A	−0.190	N/A	−0.144	N/A	**−0.343**
(The higher the	80–89 years	N/A	−0.411	N/A	−0.166	N/A	−0.285	N/A	−0.191	N/A	**−0.461**
better)	90–99 years	N/A	−0.542	N/A	−0.324	N/A	−0.357	N/A	−0.321	N/A	**−0.565**
	100–105 years	N/A	−0.473	N/A	−0.259	N/A	−0.268	N/A	−0.265	N/A	**−0.506**

Notes: ADL = Activities of daily living; DI = Disability Index; IADL = Instrumental activities of daily living domain. Missing items: self-rated health = 717; informal care from family members = 3 131; cognitive score = 2 814.

*Spearman rank order correlation coefficients; all correlation coefficients are significant ($p < .01$); **the highest value of correlation coefficient in the 4 domains of the DI in bold.**

†Pearson correlation coefficient; all correlation coefficients are significant ($p < 0.01$); **the highest value of correlation coefficient in the 4 domains of the DI in bold.**

‡Cognitive score assessed by 4 items from Community Screening Instrument for Dementia (CIS-D) and 3 items from the Telephone Interview for Cognitive Status (TICS-M).

Convergent and Divergent Validity

Table 4 presents the convergent and divergent validity of the DI subdomains. In the total sample and in each age subgroup, the IADL and physical performance domains correlated significantly with self-rated health. These correlations were higher than those between the ADL or cognitive function domains and the self-rated health.

In the total sample and in each age subgroup, the ADL and IADL domains correlated significantly with the informal care from family members. These correlations were higher than those between the physical performance or cognitive function domains and the informal care.

In the total sample and in each age subgroup, the cognitive function domain correlated significantly with the cognitive score. These correlations were higher than those between the ADL, IADL, or physical performance domain versus the cognitive score.

Known-group Validity

Supplementary Table 4 presents the known-group validity of the DI and its 4 subdomains. In the total sample and in each age subgroup, the full DI and each subdomain scores were higher in group of *Heavy informal care burden* (>8 hours/day) compared with group ≤8 hours/day, higher in group of *Limited function* compared with group of not limited, and higher in group of *Feel not well regarding health* compared with group of feel well. All the differences were statistically significant ($p < .01$).

Concurrent Validity

Table 5 presents the concurrent validity of the DI. In the total sample, and in the subgroups 80–89 years old and 90–99 years old groups, the AUCs of the full DI using *Heavy informal care burden* as the criterion were acceptable to excellent. In 65–79 years old and 100–105 years old, the AUCs varied between 0.64 and 0.69. In the total sample and all age subgroups, the AUCs of the full DI using *Limited function* as the criterion were acceptable to excellent.

Discussion

In the current study, we report the development of the 21-item DI covering 4 subdomains of ADL, IADL, physical performance, and cognitive function that indicate care needs. We found satisfactory internal consistency (Cronbach's alpha ≥0.70) of the full DI and its 4 subdomains among older adults aged 65–105 years. Our results support the convergent and divergent validity of the 4 subdomains, the known-group validity of the full DI and its 4 subdomains, and the concurrent validity of the full DI among older adults aged 65–105 years.

The Cronbach's alpha of the full DI was 0.94 in the total sample and varied between 0.89 and 0.92 in each age subgroup, which represented satisfactory internal consistency reliability (29). The Cronbach's alpha of subdomains including ADL and IADL was satisfactory with an internal consistency of Cronbach's alpha >0.7 in the total sample and each age subgroup, which were similar to other instruments with ADL and/or IADL like BI (34), GARS (35), and LB-IADL (36). The Cronbach's alpha of physical performance domain was satisfactory in the total sample with Cronbach's alpha of 0.72 and varied between 0.66 and 0.68 in each age subgroup. However, a value of Cronbach's alpha >0.6 is also acceptable according to a previous methodological study on internal consistency reliability (37). We additionally assessed the internal consistency for

Table 5. Concurrent Validity of the Disability Index (DI)

Adverse Outcomes	Sample	Cutoff Point* of DI	Sensitivity	Specificity	AUC (95% CI)†
Heavy informal care burden (>8 hours/day)	Total	≥30.17	0.74	0.76	**0.81 (0.80, 0.82)**
	65–79 years	≥13.76	0.35	0.90	0.64 (0.59, 0.70)
	80–89 years	≥22.40	0.66	0.76	**0.76 (0.73, 0.80)**
	90–99 years	≥42.23	0.61	0.71	**0.71 (0.68, 0.74)**
	100–105 years	≥54.17	0.68	0.61	0.69 (0.66, 0.72)
Limited function (GALI)	Total‡	≥13.02	0.73	0.72	**0.80 (0.79, 0.80)**
		≥13.24	0.73	0.72	
	65–79 years	≥2.60	0.62	0.73	**0.72 (0.70, 0.74)**
	80–89 years	≥12.98	0.66	0.72	**0.76 (0.74, 0.77)**
	90–99 years‡	≥34.64	0.62	0.75	**0.76 (0.74, 0.78)**
		≥34.81	0.62	0.75	
	100–105 years‡	≥54.21	0.62	0.82	**0.78 (0.76, 0.80)**
		≥54.25	0.62	0.82	

Notes: AUC = area under the receiver operating characteristic curve (ROC); CI = confidential interval; DI = Disability Index; GALI = one-item Global Activity Limitation Index. Missing items: informal care from family members = 3 131; limited function = 30.

*The Youden's index (YI, YI = sensitivity + specificity − 1) was used to determine the optimum cutoff point(s).

†0.7 ≤ AUC < 0.8 is considered acceptable concurrent validity; 0.8 ≤ AUC < 0.9 excellent; AUC ≥ 0.9 outstanding. **The value of AUC ≥ 0.7 in bold**.

‡More than one cutoff points had the maximum value according to the Youden's index, all potential cutoff points as well as corresponding sensitivity and specificity were provided.

the MMSE and found satisfactory reliability with an internal consistency of Cronbach's alpha ≥0.7 in the total sample and each age subgroup (Supplementary Table 3), which is consistent with previous studies (38,39).

The convergent and divergent validity of the 4 subdomains of DI were supported by correlations with corresponding alternative measurements in the total sample and each age subgroup. The known-group validity of the full DI and its 4 subdomains were supported by clear discriminative ability in the total sample and each age subgroup. Additionally, the concurrent validity of the full DI was excellent with AUCs ≥0.8 in the total sample and acceptable with most AUCs ≥0.7 in each age subgroup. However, the AUCs using *Heavy informal care burden* as the criterion in groups of 65–79 years old and 100–105 years old varied between 0.64 and 0.69 that were slightly lower than 0.7. The results indicated good concurrent validity of the full DI among general older populations aged 65–105 years old, but the DI still needs further improvement on specific age subgroups. We recommend future studies on the adaption of items of the DI or the development of different versions for age subgroups to meet the specific requirements of different target populations.

Additionally, we also conducted the analyses by sex subgroups, and the results supported the reliability and validity of the full DI and its 4 subdomains among both male and female older adults aged 65–105 years old (Supplementary Tables 5–8).

To the best of our knowledge, the DI is the first measurement developed specifically for older adults containing both self-reported and objectively-tested measures to assess the disability status from the perspective of care needs. Furthermore, we assessed the reliability and validity of the DI by age subgroups using the largest data set of oldest–old cohorts in the World (7), which provides important evidences on psychometric properties of the DI for further generalization among older population.

However, some limitations need to be discussed. First, we did not assess the DI's consistency over time (test–retest reliability). However, because disability status is not anticipated to remain constant over time, a poor test–retest correlation may be present with regard to the long follow-up period in the CLHLS (24–36 months). We propose to evaluate the test–retest reliability in a separate study in the future, in addition to the evaluation of the DI's consistency across items (internal consistency). Second, there is no gold standard for selecting alternative measures of DI for measuring the convergent, divergent, know-group, and concurrent validity. We adopted the alternative measures that have been validated and widely used to measure similar domains of the DI by previous studies, for instance, the self-rated health, CIS-D, TICS-M, GALI, and health during the past 2 weeks. Furthermore, considering the aim of DI was to assess the disability status of older adults to indicate care needs, we adopted the alternative measure that could represent the care needs (for instance, the informal care from family members). However, the number of alternative measures was restricted by the CLHLS project's data availability, and we recommend future research with more alternative measurements to confirm our findings. Third, we conducted the current study among Chinese older population only, and we recommend further studies on the application of the DI in other countries to contribute to the generalizability of the DI to other local contexts. Finally, we draw our conclusion mainly based on statistical analyses, and we cannot prove whether the use of the DI could provide clinically meaningful outcomes. We recommend future studies to explore the application of the DI in clinical practice.

Conclusion

Our study supported the reliability and validity of the DI with 4 subdomains among older adults aged 65–105 years. The DI may be utilized among older adults with a wide range of ages for the purpose of large-scale population studies regarding the changes in disability status summarized by the DI, in addition to analyses of different aspects of disability. Furthermore, we can use the DI as a summary measure to project future care needs for older adults, while may not use each of the 4 subdomains of the disabilities and their 16 (=4 × 4) combinations in our future trends projections to avoid overcomplications. We also recommend more validation studies on the DI among older populations in different age subgroups and different countries in the future.

Supplementary Material

Supplementary data are available at *The Journals of Gerontology, Series A: Biological Sciences and Medical Sciences* online.

Funding

This work was supported by the National Key R&D Program of China (grant number 2018YFC2000400), National Natural Science Foundation of China (grant number 72061137004), and the U.S. National Institute of Aging/National Institute of Health (grant number P01AG031719); X.Z. was supported by 2020 International Postdoctoral Exchange Fellowship Program (Talent-Introduction Program) of the Office of China Postdoc Council (OCPC) for her Postdoctoral research at Peking University, Beijing, China.

Sponsor's Role: The National Key R&D Program of China, National Natural Science Foundation of China, the U.S. National Institute of Aging/National Institute of Health, and the Office of China Postdoc Council (OCPC) had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of Interest

None declared.

Acknowledgments

We thank all participants and all organizations and professionals involved in the Chinese Longitudinal Healthy Longevity Survey (CLHLS).

Author Contributions

X.Z. and Y.Z.: study concept and design. X.L., G.G.L., and Y.Z.: acquisition of subjects and data. X.Z., L.L., X.S., and H.R.: analysis and interpretation of data. X.Z., L.L., and H.R.: preparation of manuscript. All authors: critical revision and final approval of the manuscript.

References

1. Smith SK, Rayer S, Smith E, Wang Z, Zeng Y. Population aging, disability and housing accessibility: implications for sub-national areas in the United States. *Hous Stud*. 2012;27:252–266. doi:10.1080/02673037.2012.649468
2. Fong JH, Feng J. Patterns of functional disability in the oldest adults in China. *J Am Geriatr Soc*. 2016;64(9):1890–1894. doi:10.1111/jgs.14264
3. Zeng Y, Chen H, Wang Z, Land KC. Implications of changes in households and living arrangements for future home-based care needs and costs for disabled elders in China. *J Aging Health*. 2015;27(3):519–550. doi:10.1177/0898264314552690
4. Edemekong P, Bomgaars D, Sukumaran S, Levy S. *Activities of daily living*. Updated June 26, 2020. In: StatPearls [Internet]. Treasure Island (FL). StatPearls Publishing; 2020.
5. Foster ER. Instrumental activities of daily living performance among people with Parkinson's disease without dementia. *Am J Occup Ther*. 2014;68(3):353–362. doi:10.5014/ajot.2014.010330
6. Burwell B, Jackson B. *The Disabled Elderly and Their Use of Long-Term Care*. U.S. Department of Health and Human Services; 1994.
7. Zeng Y, Feng Q, Hesketh T, Christensen K, Vaupel JW. Survival, disabilities in activities of daily living, and physical and cognitive functioning among the oldest-old in China: a cohort study. *Lancet*. 2017;389(10079):1619–1629. doi:10.1016/S0140-6736(17)30548-2
8. Mlinac ME, Feng MC. Assessment of activities of daily living, self-care, and independence. *Arch Clin Neuropsychol*. 2016;31(6):506–516. doi:10.1093/arclin/acw049
9. Yang M, Ding X, Dong B. The measurement of disability in the elderly: a systematic review of self-reported questionnaires. *J Am Med Dir Assoc*. 2014;15(2):150.e1–150.e9. doi:10.1016/j.jamda.2013.10.004
10. Theou O, Cann L, Blodgett J, Wallace LM, Brothers TD, Rockwood K. Modifications to the frailty phenotype criteria: systematic review of the current literature and investigation of 262 frailty phenotypes in the Survey of Health, Ageing, and Retirement in Europe. *Ageing Res Rev*. 2015;21:78–94. doi:10.1016/j.arr.2015.04.001
11. Purser JL, Feng Q, Yi Z, Hoenig H. A new classification of function and disability in China: subtypes based on performance-based and self-reported measures. *J Aging Health*. 2012;24(5):779–798. doi:10.1177/0898264312444310
12. Reuben DB, Seeman TE, Keeler E, et al. Refining the categorization of physical functional status: the added value of combining self-reported and performance-based measures. *J Gerontol A Biol Sci Med Sci*. 2004;59(10):1056–1061. doi:10.1093/gerona/59.10.m1056
13. Cosco TD, Stephan BC, Brayne C. Validation of an a priori, index model of successful aging in a population-based cohort study: the successful aging index. *Int Psychogeriatr*. 2015;27(12):1971–1977. doi:10.1017/S1041610215000708
14. Searle SD, Mitnitski A, Gahbauer EA, Gill TM, Rockwood K. A standard procedure for creating a frailty index. *BMC Geriatr*. 2008;8:24. doi:10.1186/1471-2318-8-24
15. Cosco TD, Prina AM, Perales J, Stephan BC, Brayne C. Operational definitions of successful aging: a systematic review. *Int Psychogeriatr*. 2014;26(3):373–381. doi:10.1017/S1041610213002287
16. Yang F, Gu D. Predictability of frailty index and its components on mortality in older adults in China. *BMC Geriatr*. 2016;16:145. doi:10.1186/s12877-016-0317-z
17. Zhang Y, Xiong Y, Yu Q, Shen S, Chen L, Lei X. The activity of daily living (ADL) subgroups and health impairment among Chinese elderly: a latent profile analysis. *BMC Geriatr*. 2021;21:30. doi:10.1186/s12877-020-01986-x
18. Zhang PD, Lv YB, Li ZH, et al. Age, period, and cohort effects on activities of daily living, physical performance, and cognitive functioning impairment among the oldest-old in China. *J Gerontol A Biol Sci Med Sci*. 2020;75(6):1214–1221. doi:10.1093/gerona/glz196
19. Xie J, Brayne C, Matthews FE; Medical Research Council Cognitive Function and Ageing Study Collaborators. Survival times in people with dementia: analysis from population based cohort study with 14 year follow-up. *BMJ*. 2008;336(7638):258–262. doi:10.1136/bmj.39433.616678.25
20. Yao Y, Chen H, Chen L, et al. Type of tea consumption and depressive symptoms in Chinese older adults. *BMC Geriatrics*. 2021;21:331. doi:10.1186/s12877-021-02203-z
21. Zeng Y, Feng Q, Gu D, Vaupel JW. Demographics, phenotypic health characteristics and genetic analysis of centenarians in China. *Mech Ageing Dev*. 2017;165(Pt B):86–97. doi:10.1016/j.mad.2016.12.010
22. Poston DL, Min H. The effects of sociodemographic factors on the hazard of dying among chinese oldest old. In: Yi Z, Poston DL, Vlosky DA, Gu D, eds. *Healthy Longevity in China: Demographic, Socioeconomic, and Psychological Dimensions*. Dordrecht: Springer Netherlands; 2008:121–132.
23. Hall KS, Gao S, Emsley CL, Ogunniyi AO, Morgan O, Hendrie HC. Community screening interview for dementia (CSI 'D'); performance in five disparate study sites. *Int J Geriatr Psychiatry*. 2000;15(6):521–531. doi:10.1002/1099-1166(200006)15:6<521::aid-gps182>3.0.co;2-f
24. de Jager CA, Budge MM, Clarke R. Utility of TICS-M for the assessment of cognitive function in older adults. *Int J Geriatr Psychiatry*. 2003;18(4):318–324. doi:10.1002/gps.830
25. Lin W. The relationship between formal and informal care among Chinese older adults: based on the 2014 CLHLS dataset. *BMC Health Serv Res*. 2019;19:323. doi:10.1186/s12913-019-4160-8.
26. Jowsey T, McRae I, Gillespie J, Banfield M, Yen L. Time to care? Health of informal older carers and time spent on health related activities: an Australian survey. *BMC Public Health*. 2013;13:374. doi:10.1186/1471-2458-13-374

27. Berger N, Van Oyen H, Cambois E, et al. Assessing the validity of the Global Activity Limitation Indicator in fourteen European countries. *BMC Med Res Methodol*. 2015;15:1. doi:10.1186/1471-2288-15-1
28. Raat H, Bonsel GJ, Essink-Bot ML, Landgraf JM, Gemke RJ. Reliability and validity of comprehensive health status measures in children: The Child Health Questionnaire in relation to the Health Utilities Index. *J Clin Epidemiol*. 2002;55(1):67–76. doi:10.1016/s0895-4356(01)00411-5
29. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ*. 2011;2:53–55. doi:10.5116/ijme.4dfb.8dfd
30. Zhang X, Tan SS, Bilajac L, et al. Reliability and validity of the tilburg frailty indicator in 5 European Countries. *J Am Med Dir Assoc*. 2020;21(6):772–779.e6. doi:10.1016/j.jamda.2020.03.019
31. Raat H, Botterweck AM, Landgraf JM, Hoogeveen WC, Essink-Bot ML. Reliability and validity of the short form of the child health questionnaire for parents (CHQ-PF28) in large random school based and general population samples. *J Epidemiol Community Health*. 2005;59(1):75–82. doi:10.1136/jech.2003.012914
32. Raat H, Landgraf JM, Oostenbrink R, Moll HA, Essink-Bot ML. Reliability and validity of the Infant and Toddler Quality of Life Questionnaire (ITQOL) in a general population and respiratory disease sample. *Qual Life Res*. 2007;16(3):445–460. doi:10.1007/s11136-006-9134-8
33. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol*. 2010;5(9):1315–1316. doi:10.1097/JTO.0b013e3181ec173d
34. Ohura T, Higashi T, Ishizaki T, Nakayama T. Assessment of the validity and internal consistency of a performance evaluation tool based on the Japanese version of the modified Barthel Index for elderly people living at home. *J Phys Ther Sci*. 2014;26(12):1971–1974. doi:10.1589/jpts.26.1971
35. Suurmeijer TP, Doeglas DM, Moum T, et al. The Groningen Activity Restriction Scale for measuring disability: its utility in international comparisons. *Am J Public Health*. 1994;84(8):1270–1273. doi:10.2105/ajph.84.8.1270
36. Isik EI, Yilmaz S, Uysal I, Basar S. Adaptation of the Lawton instrumental activities of daily living scale to Turkish: validity and reliability study. *Ann Geriatr Med Res*. 2020;24(1):35–40. doi:10.4235/agmr.19.0051
37. Taber KS. The use of Cronbach's Alpha when developing and reporting research instruments in science education. *Res Sci Educ*. 2018;48:1273–1296. doi:10.1007/s11165-016-9602-2
38. Marioni RE, Chatfield M, Brayne C, Matthews FE; Medical Research Council Cognitive Function and Ageing Study Group. The reliability of assigning individuals to cognitive states using the Mini Mental-State Examination: a population-based prospective cohort study. *BMC Med Res Methodol*. 2011;11:127. doi:10.1186/1471-2288-11-127
39. Kabatova O, Putekova S, Martinkova J. Analysis of psychometric features of the Mini-Mental State Examination and the Montreal Cognitive Assessment methods. *Clin Soc Work J*. 2016;7:62–69. doi:10.22359/CSWHI_7_2_08