

# Validation and Comparison of Four Mortality Prediction Models in a Geriatric Ward in China

Yuanyuan Li , Xiaohong Liu\*, Lin Kang\*, Jiaojiao Li 

Department of Geriatrics, Peking Union Medical College, Chinese Academy of Medical Sciences, Peking Union Medical College Hospital, Beijing, People's Republic of China

\*These authors contributed equally to this work

Correspondence: Xiaohong Liu; Lin Kang, Department of Geriatrics, Peking Union Medical College, Chinese Academy of Medical Sciences, Peking Union Medical College Hospital, Beijing, People's Republic of China, Email xhliu41@sina.com; kangl@pumch.cn

**Purpose:** The efficacy of mortality risk prediction models among older patients in China remains uncertain. We aimed to validate and compare the performances of the Walter Index, Geriatric Prognostic Index (GPI), Charlson Comorbidity Index (CCI), and FRAIL Scale in predicting 1-year all-cause mortality post-discharge in geriatric inpatients in China.

**Patients and Methods:** This study was conducted at a geriatric ward of a tertiary Hospital in Beijing, including patients aged 70 years or older with a documented comprehensive geriatric assessment, discharged between January 1, 2016, and December 31, 2021. Patients with a hospital stay  $\leq 24$  h or  $>60$  days were excluded. All-cause mortality data within one year of discharge were collected from medical files and telephone interviews between August 2022 and February 2023. Multiple imputation, Logistic regression analysis, Brier scores, C-statistics, Hosmer-Lemeshow goodness-of-fit-test, and calibration plots were employed for statistical analysis.

**Results:** We included 832 patients with a median (interquartile range) age of 77 (74–82) years. One-hundred patients (12.0%) died within one year. After adjusting for covariates—marital status, social support, cigarette use, length of stay, number of medications, hemoglobin levels, handgrip strength, and Short Physical Performance Battery—CCI scores of 3–4 and  $>4$ , and increased Walter Index, GPI, and FRAIL Scale scores were significantly associated with 1-year mortality risk. The Brier scores varied from 0.07 (Walter Index) to 0.10 (FRAIL Scale). The C-statistic ranged from 0.74 (95% confidence interval, 0.69–0.78) for FRAIL Scale to 0.88 (95% confidence interval, 0.84–0.91) for the Walter Index. Calibration curves showed that the Walter Index, GPI, and FRAIL Scale were well calibrated, while the CCI was poor.

**Conclusion:** Combining the Brier score, discrimination and calibration, the Walter Index was confirmed for the first time to be the best model to predict the 1-year mortality risk of geriatric inpatients in China among the four models.

**Keywords:** aged, care for older adults, comprehensive geriatric assessment, frail, prediction models

## Introduction

Population aging is a source of concern to health professionals and policymakers worldwide.<sup>1</sup> The disease-oriented approach to healthcare has proven insufficient to meet the healthcare and social needs of a fast-growing cohort of older adults.<sup>2,3</sup> Chronic diseases are not curable, and geriatrics aims to maintain functional capacity, enhance the quality of life, and extend the healthy life expectancy of older adults. Ignoring prognosis may contribute to overuse and low-value care.<sup>4</sup> It is often difficult to predict life expectancy for older adults only on empirical clinical judgment, muddling the differentiation of those who could benefit from disease screening and medical interventions from those who could benefit from hospice and palliative care.<sup>5</sup> The benefits-harms balance of clinical interventions has important clinical implications, especially in older adults. Mortality prediction models can help estimate individual prognosis and guide clinical decision-making, enabling physicians to tailor treatment according to the patient's condition, preferences, and prognosis.

Current mortality prediction models for the older adults include 6-month, 1-year, and more than 1-year mortality prediction models.<sup>6–13</sup> The variables included in these models consist of diseases only, functional status only, diseases combined with

functional status, and geriatric syndrome.<sup>6–13</sup> Some of these models have been validated in European and American countries, but there is a lack of studies to confirm their validity for the older adults in China. Based on the data available in this study, the Walter Index, GPI, CCI, and FRAIL Scale models were selected for validation and comparison. The Walter Index prediction variables include diseases and physical function, and is recommended as a 1-year mortality prediction model for older inpatients by the University of San Francisco.<sup>6</sup> The GPI is the only 3- and 5-year mortality prediction model developed with Korean patients, which centers on comorbidities and the comprehensive geriatric assessment (CGA).<sup>13</sup> The Charlson Comorbidity Index (CCI) is a 1-year mortality prediction model for older patients with multimorbidity, and high scores predict a significantly increased risk of death.<sup>14</sup> Frailty is a very important geriatric syndrome that significantly increases mortality risk.<sup>15,16</sup> The FRAIL Scale is a simple tool to assess frailty, comprising only five questions that can be assessed in 5 mins.<sup>17,18</sup> The Walter Index and GPI have not been validated in older adults in China, and the CCI and FRAIL Scale have not been compared with the Walter Index and GPI in China; therefore, the purpose of this study was to validate and compare the performance of the Walter Index, GPI, CCI, and FRAIL Scale in predicting 1-year mortality, providing evidence from geriatric inpatients in China.

## Materials and Methods

### Setting and Patients

This study was conducted in a geriatric assessment and management ward in the Peking Union Medical College, Chinese Academy of Medical Sciences, Peking Union Medical College Hospital, a tertiary care hospital in Beijing. The geriatric ward is a sub-acute management unit staffed by geriatricians, specialists, psychologists, nutritionists, rehabilitation physicians, pharmacists, nurses, and social workers, dedicated to maintaining the functional status of older patients while managing their diseases. We reviewed the records of patients discharged between January 1, 2016, and December 31, 2021. Patients aged 70 years or older with documented CGA were included; patients with a hospital stay  $\leq 24$  h or  $>60$  days were excluded. If the patients were hospitalized repeatedly, the first hospitalization was recorded. The data of all these patients were recorded, including demographics (age, sex, marital status, and cigarette use), CGA, length of stay, discharge diagnosis, and clinical laboratory test results (albumin, hemoglobin, and creatinine), to calculate the scores of the four models.

### Comprehensive Geriatric Assessment

The CGA of all patients was performed by a trained physician when the disease condition was relatively stable after hospital admission. The functional status assessment was based on the activities of daily living (ADL) and instrumental activities of daily living (IADL). The Physical Self-Maintenance Scale (PSMS) was used to assess ADL by measuring the degree of dependence/independence in daily personal care activities (feeding, toileting, grooming, bathing, and physical ambulation).<sup>19</sup> The Lawton-IADL was used to evaluate IADL (independence for eight activities: ability to use telephone, food preparation, shopping, laundry, housekeeping, responsibility for own medication, mode of transportation, and capacity to handle finances).<sup>19</sup> The Mini-Nutritional Assessment-Short Form (MNA-SF) was used to assess nutritional status and was comprised of mobility, declining food intake, weight loss, psychological stress or acute illness, neuropsychological illness, and body mass index or calf circumference (cm).<sup>20</sup> The Mini-Mental State Examination (MMSE) was used to assess time and place orientation, registration, attention, recall, and language.<sup>21</sup> Mood was evaluated by the Geriatric Depression Scale (GDS-15).<sup>22</sup> Muscle strength was evaluated by handgrip strength, with low muscle strength identified as handgrip strength  $<28$  kg for males or  $<18$  kg for females.<sup>23</sup> A Short Physical Performance Battery (SPPB) score  $\leq 9$  was considered low physical performance.<sup>23,24</sup> The number of medications (including prescribed medications) was recorded, and  $\geq 5$  was defined as polypharmacy. The social support network was classified as living alone, living with family, or institutionalized.

### Mortality Prediction Models

The Walter Index includes five variables: sex, ADL (a modified version of Katz-ADL) dependence, comorbidities (congestive heart failure, cancer, metastatic cancer), albumin, and creatinine, with a total score of 20.<sup>6,25</sup> The GPI includes eight variables: age, sex, Korean ADL, Korean IADL, Mini-Nutritional Assessment (MNA), MMSE, GDS-15, and CCI, with a total score

of 8.<sup>13,21,22,26–28</sup> The CCI includes 19 diseases; each disease is assigned a different score of 1, 2, 3, or 6.<sup>28</sup> The FRAIL Scale consists of 5 variables: fatigue, resistance, ambulation, illness, and weight loss; a score of 0 is strong, 1–2 is pre-frail, and 3–5 is frail.<sup>17,18</sup> Model details of the original studies assessing the Walter Index, CCI, and GPI are summarized in [Table S1](#). The characteristics of the different ADL and IADL scales are shown in [Table S2](#).

## Outcome Measures

The primary outcome was all-cause mortality within one year of discharge, which was defined as death from any cause. Data was collected from medical files and telephone interviews conducted between August 2022 and February 2023.

## Statistical Analysis

Baseline patient characteristics are presented as medians (interquartile range [IQR]) or numbers (proportion). We report the number (proportion) of participants with missing data for each variable. To maximize statistical power and minimize potential bias by excluding patients with missing data from the analysis, we created 20 imputation datasets using multiple imputations based on the Markov chain Monte Carlo method in the SPSS program (version 26.0).<sup>29</sup> To ascertain whether post-imputation data differed significantly from pre-imputation data, we utilized sensitivity comparative analysis. Pre-imputation data were used to report the results of our analyses; the results from the 20 imputation datasets are reported in the [Supplementary Materials](#).

Logistic regression was used to assess the association between the four models and the 1-year risk of all-cause mortality. We adjusted for common covariates as categorical variables, which were not covered in any of the four models, including marital status (married, not married), cigarette use (Never, Former, Current), social support (living alone, living with family, institutionalized), length of stay ( $\leq 7$ d,  $> 7$ d), number of medications ( $< 5$ ,  $\geq 5$ ), hemoglobin level ( $\geq 110$  g/L,  $< 110$  g/L), handgrip strength (normal, low), and SPPB ( $> 9$ ,  $\leq 9$ ). To avoid the bias of missing data, multivariate logistic regression results based on the imputation data were combined according to Rubin's rule.<sup>30</sup>

We used the calculated model scores as the only independent variable to assess the performance of the models, including the: (1) Brier Score, an overall performance measure, ranging from 0 (excellent overall performance) to 0.25 (indicating a non-informative model)<sup>31,32</sup> and (2) C-statistic, corresponding to the area under the receiver operating characteristic (ROC) curve of the binary dependent variable, which served as the model discrimination measure. The C-statistic shows how well the prediction model differentiates those at higher and lower mortality risk and ranges from 0.5 to 0.59 for poor, 0.6 to 0.69 for moderate, 0.7 to 0.79 for good, 0.8 to 0.89 for very good, and  $> 0.9$  for excellent.<sup>33</sup> We performed pairwise comparisons of the ROC curves for each model with the DeLong method.<sup>34</sup> (3) The Hosmer-Lemeshow goodness-of-fit-test (HL test) and calibration plots were used to assess the calibration of the four models. The HL test distinguishes between a null hypothesis of an excellent fit, where the model assumes probabilities that are consistent with the actual probabilities, and a generalized alternative hypothesis of a non-perfect fit.<sup>35</sup> Calibration curves were plotted based on observed and predicted mortality, using 1000 bootstrap resamples to reduce overfit bias.

Analysis was performed with SPSS, version 26.0 (released 2019, IBM SPSS Statistics for Windows, Armonk, NY), R project, version 4.2.3 (R-project of the statistical calculation, Vienna, Austria) and EmpowerStats software, version 4.1 ([www.empowerstats.com](http://www.empowerstats.com), X&Y solutions, Boston, MA, USA). All statistical tests were two-sided, and  $P < 0.05$  was considered statistically significant.

## Results

### Baseline Characteristics of the Study Population

Of 1275 possible participants, 369 patients with recurrent admissions or that did not complete the CGA and two patients with a length of stay  $> 60$  d were excluded from this study. Of the remaining 904 patients, 24 (2.7%) patients' families refused information on the patients' condition and 48 (5.3%) patients were lost to follow-up. The study sample included 832 patients, with a median age of 77.0 (74.0–78.0) years (interquartile range, IQR) and 380 men (45.7%). The median (IQR) scores were 2.0 (1.0–4.0), 2.5 (1.5–4.0), 1.0 (0.0–2.0), and 1.0 (0.0–3.0) for the Walter Index, GPI, CCI, and FRAIL Scale, respectively. Within one year, 100 participants (12.0%) died. The baseline characteristics of the participants are shown in [Table 1](#).

**Table 1** Baseline Characteristics of Participants

Characteristic	Participants n=832	Patients with Missing Data n (%)
Age, median (IQR), y	77.0 (74.0–82.0)	0
Sex, n (%)		0
Female	452 (54.3)	
Male	380 (45.7)	
Length of hospital stay, median (IQR), d	15.0 (11.0–20.0)	0
No. of medications, median (IQR)	5.0 (3.0–7.0)	0
ADL, dependent, n (%)		0
Toilet	167 (20.1)	
Feeding	51 (6.1)	
Dressing	124 (14.9)	
Grooming	131 (15.7)	
Physical Ambulation	335 (40.3)	
Bathing	220 (26.4)	
IADL, median (IQR)	8.0 (4.0–8.0)	0
MNA-SF, median (IQR)	12.0 (9.0–13.5)	1 (0.1)
MMSE, median (IQR)	27.0 (24.0–29.0)	52 (6.3)
GDS-15, median (IQR)	3.0 (1.0–5.0)	79 (9.5)
Grip strength, median (IQR), kg	22.0 (17.8–28.7)	111 (13.3)
SPPB, median (IQR)	9.0 (4.0–11.0)	59 (7.1)
Living condition, n (%)		0
Living alone	103 (12.4)	
Living with family	721 (86.7)	
Institutionalized	8 (1.0)	
Cigarette use, n (%)		0
Never	596 (71.6)	
Former	113 (13.6)	
Current	123 (14.8)	
Marital, n (%)		0
Married	642 (77.2)	
Not married	190 (22.8)	
Laboratory values, median (IQR)		
Hemoglobin, g/L	123.0 (110.8–136.0)	0
Albumin, g/dL	3.8 (3.5–4.1)	1 (0.1)
Creatine, mg/dL	0.8 (0.7–1.0)	1 (0.1)
Walter Index, median (IQR)	2.0 (1.0–4.0)	2 (0.2)
GPI, median (IQR)	2.5 (1.5–4.0)	101 (12.1)
CCI, median (IQR)	1.0 (0.0–2.0)	0
FRAIL Scale, median (IQR)	1.0 (0.0–3.0)	13 (1.6)

**Abbreviations:** IQR, interquartile range; ADL, Activities of daily living; IADL, Instrumental activities of daily living; GDS-15, Geriatric Depression Scale –15; MMSE, Mini-Mental State Examination; MNA-SF, Mini Nutritional Assessment Short Form; SPPB, Short-Physical Performance Battery, GPI, Geriatric Prognostic Index, CCI, Charlson Comorbidity Index.

Of the 832 participants, 194 (23.3%) had missing data and two (0.2%) from the Walter Index, 101 (12.1%) from GPI, and 13 (1.6%) from the FRAIL Scale. There were no missing data for CCI (Table 1). Sensitivity comparative analysis results indicated no significant differences between pre- and post-imputation data (Table S3).

## Prediction of 1-Year of Mortality Risk

The results of the logistic regression analysis are presented in Table 2. With univariate logistic regression and adjustment for common covariates (length of hospitalization, marital status, social support, number of medications, and hemoglobin

**Table 2** Logistic Regression of Associations Between the 4 Models and 1-Year Mortality Risk

	Non-Adjusted		Adjustment I		Adjustment II	
	OR (95% CI)	P -value	OR (95% CI)	P- value	OR (95% CI)	P -value
Walter Index risk groups, scores						
0-1	Reference		Reference		Reference	
2-3	5.58 (1.54, 20.23)	0.009*	5.21 (1.42, 19.06)	0.013*	3.89 (1.03, 14.68)	0.045*
4-6	31.85 (9.63, 105.39)	<0.001 <sup>†</sup>	23.86 (7.02, 81.07)	<0.001 <sup>†</sup>	17.36 (4.96, 60.84)	<0.001 <sup>†</sup>
>6	163.81 (48.56, 552.54)	<0.001 <sup>†</sup>	111.02 (31.73, 388.39)	<0.001 <sup>†</sup>	81.43 (22.54, 294.18)	<0.001 <sup>†</sup>
CCI risk groups, scores						
0	Reference		Reference		Reference	
1-2	2.58 (1.11, 5.99)	0.027*	2.52 (1.07, 5.96)	0.035*	2.22 (0.93, 5.32)	0.07
3-4	6.28 (2.59, 15.21)	<0.001 <sup>†</sup>	5.43 (2.14, 13.74)	<0.001 <sup>†</sup>	3.89 (1.50, 10.08)	0.005*
>4	48.30 (19.88, 117.30)	<0.001 <sup>†</sup>	31.23 (12.24, 79.67)	<0.001 <sup>†</sup>	27.06 (10.35, 70.76)	<0.001 <sup>†</sup>
FRAIL Scale risk groups, scores						
0	Reference		Reference		Reference	
1-2	3.47 (1.41, 8.51)	0.007*	3.34 (1.33, 8.42)	0.011*	2.76 (1.07, 7.15)	0.036*
3-5	14.06 (5.96, 33.18)	<0.001 <sup>†</sup>	10.76 (4.33, 26.70)	<0.001 <sup>†</sup>	7.97 (2.82, 22.56)	<0.001 <sup>†</sup>
GPI, continuous variable, scores						
	1.73 (1.48, 2.03)	<0.001 <sup>†</sup>	1.66 (1.38, 1.98)	<0.001 <sup>†</sup>	1.63 (1.30, 2.04)	<0.001 <sup>†</sup>

**Notes:** Adjustment I covariates: marital status; social support network, cigarette use, length of stay, number of medications, hemoglobin. Adjustment II covariates: Adjustment I covariates + handgrip strength and Short Physical Performance Battery. \*P<0.05. <sup>†</sup> P<0.001. **Abbreviations:** OR, Odds Ratio; CI, Confidence Interval; CCI, Charlson Comorbidity index; GPI, Geriatric Prognostic Index.

levels), the 1- year mortality risk increased with increasing scores in all four models. After adjusting for handgrip strength and SPPB (except for CCI 1-2 scores that were no longer associated with 1-year mortality), and CCI 3-4 and >4 scores, increased scores of the Walter Index, FRAIL Scale, and GPI showed significant associations with 1-year mortality. The results of logistic regression analyses based on the imputed datasets, combined with Rubin’s rules, are shown in [Table S4](#), and indicated no significant differences between pre- and post-imputation.

### Performance of the Mortality Prognostic Models

According to the Brier scores, the Walter Index (0.07), GPI (0.08), CCI (0.08), and FRAIL Scale (1.00) performed well, with the Walter Index being the best. The C-statistic showed good discriminatory power with the Walter Index 0.88 (95% confidence interval, CI: 0.84-0.91), CCI 0.78 (95% CI: 0.73-0.83), GPI 0.75 (95% CI: 0.69-0.80), and FRAIL Scale 0.74 (0.69-0.78) ([Table 3](#), [Figure S1](#)). In the pairwise comparison, the discrimination of the Walter index outperformed those of the GPI, CCI, and FRAIL Scale (all P <0.001). The differences among the GPI, CCI, and FRAIL Scale were not

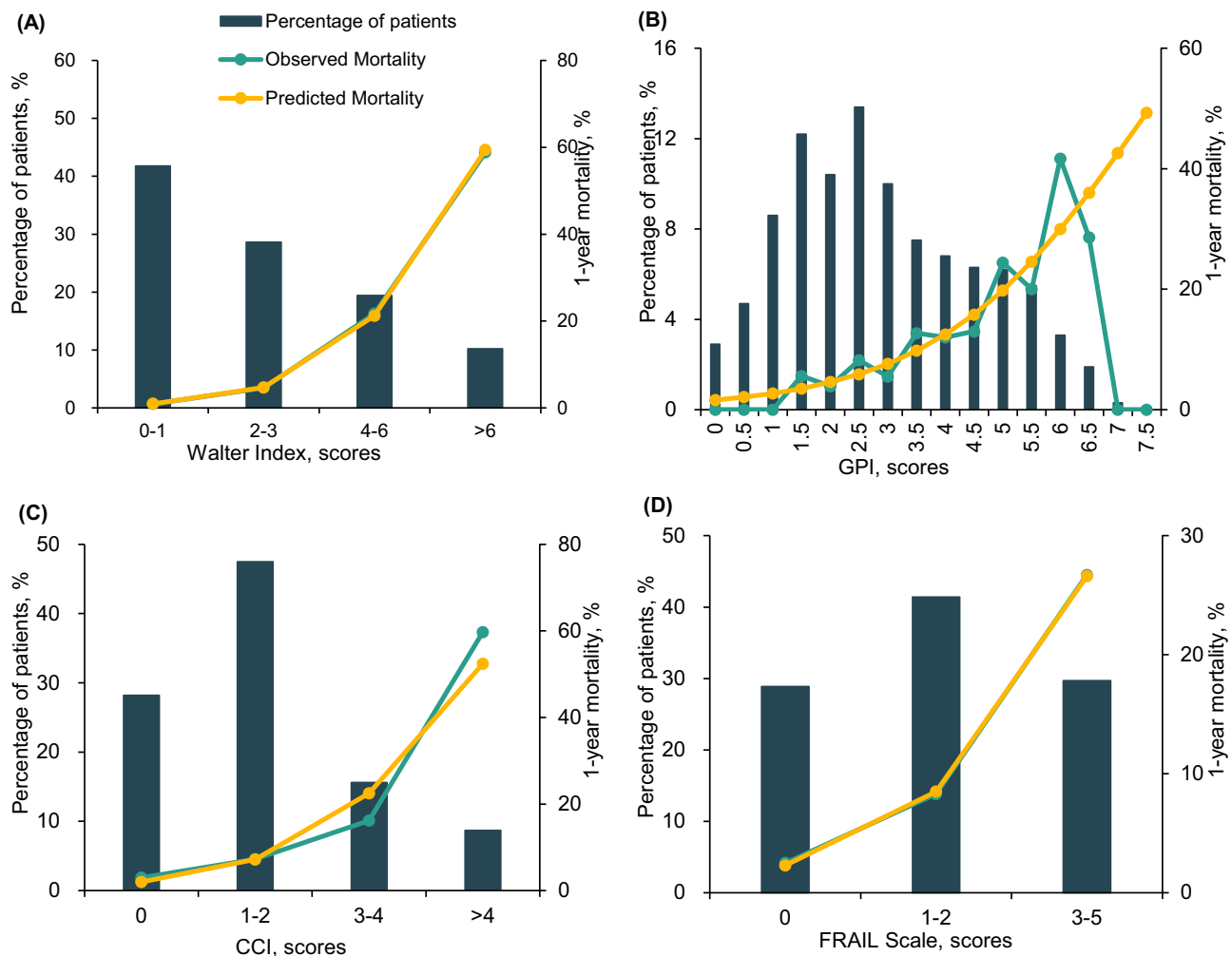
**Table 3** The Performance of 4 Prediction Models in Predicting 1-Year Mortality

	Overall Performance	Discrimination	Calibration By HOSMER-Lemeshow Goodness-of-Fit Test	
	Brier Score	C-Statistic (95% CI)	$\chi^2$	P-value
Walter Index	0.07	0.88 (0.84-0.91)	0.05	0.98
GPI	0.08	0.75 (0.69-0.80)	5.86	0.66
CCI	0.08	0.78 (0.73-0.83)	0.49	0.78
FRAIL Scale	0.10	0.74 (0.69-0.78)	0.01	0.92

**Abbreviations:** GPI, Geriatric Prognostic Index; CCI, Charlson Comorbidity Index; CI, Confidence Interval.

statistically significant ( $P > 0.05$ ) (Table S5). The Brier scores and C-statistics of the four models pre- and post-imputation are shown in Tables S6 and S7 and show no significant changes.

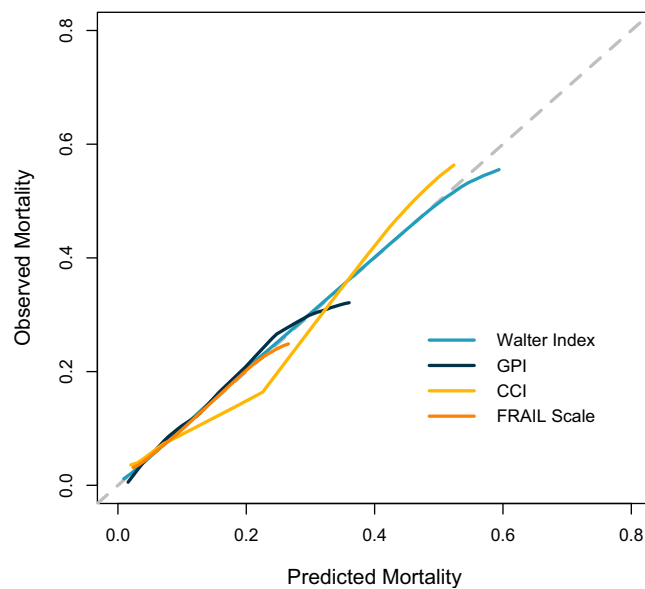
The  $P$  values for the four models in the HL test ranged from 0.66 to 0.98, indicating acceptable calibration (Table 3). The distribution of the calculated risk scores and the relationship between the 1-year observed and predicted mortalities are shown in Figure 1 and Table S8. The calibration curves are shown in Figure 2. The predicted 1-year mortality, based on the Walter Index and FRAIL Scale, for each risk group was approximated to the observed mortality, and the calibration curves almost corresponded to the ideal calibration curve (Figures 1A, 1D and 2). The 1-year mortality risk range predicted by the Walter Index was much greater than that predicted by the FRAIL Scale. GPI scores of 0–5.5 corresponded to predicted mortality risk of 1.6% to 24.6%, with a small difference between predicted and observed mortality risk, with a larger difference for GPI scores of 6–7.5 (Figure 1B). However, the patient sample for GPI scores of 6–7.5 was insufficient. After resampling Bootstrap 1000 repetitions, the calibration curve almost corresponded to the ideal curve for 1-year mortality  $< 20\%$  and deviated slightly from the ideal curve when 1-year mortality increased (Figure 2). The CCI calibration was good when the 1-year mortality rate was below 7% but poorer as the 1-year mortality rate increased (Figure 1C). Despite resampling Bootstrap 1000 repetitions, the calibration curve significantly deviated from the ideal calibration curve (Figure 2). The calibration curves for the pre- and post-imputation data of the Walter



**Figure 1** Distribution of calculated risk scores and the relationship between observed and predicted 1-year mortality.

**Notes:** The columns show the percentage of patients in different subgroups, and the lines show the 1-year mortality, with the green line is the observed mortality and the orange line is the predicted mortality. Sub-figures **A–D** represent distribution of calculated risk scores of the Walter Index, GPI, CCI, and FRAIL Scale and the relationship between observed and predicted 1-year mortality, respectively.

**Abbreviations:** CCI, Charlson Comorbidity Index; GPI, Geriatric Prognostic Index.



**Figure 2** Calibration plot for 1-year mortality.

**Notes:** Calibration plot with predicted mortality on the x-axis and observed mortality on the y-axis. A perfect prediction corresponds to the 45° dotted line (ideal calibration line). Calibration curves were bias-corrected using bootstrapping (B=1000 repetitions), indicating the performance of the observed models.

**Abbreviations:** CCI, Charlson Comorbidity Index; GPI, Geriatric Prognostic Index.

Index, GPI, and FRAIL Scale are shown in [Figures S2–S4](#), indicating no significant differences between pre- and post-imputation. For the two prediction models with 1-year mortality data from the original study, the Walter Index and CCI, we further compared the observed 1-year mortality based on the original development cohort with our study population and found that the observed mortality risk per risk strata of the Walter Index and CCI was generally lower in our study than in the development cohorts ([Figures S5, S6](#)).

## Discussion

Prediction models should be validated at various times and locations to better guide clinical practice. For the first time, we provide data from older inpatients in China to validate the Walter Index and GPI and compare them with conventional models of CCI and FRAIL. The 1-year mortality risk in older patients increased progressively with increases in the four model scores. All the models exhibited good discrimination in predicting the 1-year mortality risk, the best of which was the Walter Index. In terms of calibration, the Walter Index also performed the best, followed by the FRAIL Scale and GPI, and the CCI was poor.

Prediction models could provide the determinants of poor prognosis and assist in clinical decision-making and prioritization of care goals. This could help clinicians and patients choose invasive and/or conservative treatments and facilitate shared decision-making between physicians and patients.<sup>36</sup> This would be important for older patients with limited life expectancy to avoid useless life-prolonging measures.<sup>36</sup> In our study, the Walter Index, GPI, and FRAIL Scale showed good discrimination and calibration, with commonality in the predictive variables used, with both disease and functional status being assessed, compared with CCI, which only assessed disease. Additionally, the Walter Index had better discrimination than the CCI in predicting 1-year mortality in older inpatients in a study from six European countries.<sup>32</sup> The results, analyzed from a large European countries database, emphasize that for older patients, poor prognosis is affected by disease patterns, cognition, functional status, and socioeconomic support.<sup>2</sup> A comprehensive assessment should be conducted to determine different care options and to change the traditional disease-oriented treatment model towards maintaining functional status and improving quality of life.<sup>2</sup> Therefore, mortality prediction models that simultaneously assess disease and functional status would be more appropriate for older patients.

In our study, the Walter Index was the best-performing 1-year mortality prediction model, with a total score of 20. Analysis of the predictive variables and the weights assigned to each variable of the Walter Index revealed that

for comorbidities, it focused more on diseases with high 1-year mortality, such as heart failure, tumors, and metastatic tumors, with a weight of 8 points. It also focused more on disabled patients, with all ADLs dependence assigned five points. However, each variable was assigned the same weight in the GPI and FRAIL Scale, not considering the different effects of each variable on mortality risk. The Walter Index contains laboratory variables, with low serum albumin indicating malnutrition and chronic inflammation, which is an independent predictor of 1-year mortality.<sup>37–39</sup> The Walter Index contains risk factors in four domains: demographics, medical diagnoses, functional status, and laboratory test results, consistent with clinical practice, where the causes of death are multifactorial in older adults.<sup>6,40</sup>

The Walter Index showed very good calibration, but the 1-year mortality in our study population was substantially lower than that in the original development cohort, possibly because our study population was younger, less ADL-dependent, had a lower incidence of decreased albumin and increased creatinine levels, and had a lower incidence of comorbidities (other than diabetes and tumors) than that of the Walter index development cohort population (Table S9). The lower mortality may be related to public health developments, such as increased public health funding, improved health care access, decreased health inequities, better educational levels, and improved nutrition, which have contributed to decreased mortality due to chronic diseases and cancer and an increased life expectancy in China.<sup>41–47</sup> For the CCI, we also observed lower mortality for each risk stratum compared to the original development cohort. Further validation of the Walter Index is needed for more accurate prediction of 1-year mortality in older inpatients from different regions of China and other developing countries.

In addition, the GPI and FRAIL Scale were also demonstrated to present good discrimination and calibration in this study, especially for the FRAIL Scale, which can be scored in a short period of time, which is very convenient. In contrast, GPI involves multidimensional assessment, which is relatively time-consuming and laborious, may not be suitable for some scores for very frail patients, such as MMSE and GDS-15. But different interventions can be given according to the results of different dimensions, such as malnutrition, cognitive decline, and depression, which is more beneficial to the comprehensive management of older inpatients. Therefore, clinicians can choose more appropriate scoring methods according to different scenarios.

We evaluated four models using a multidimensional method to provide evidence from older inpatients in China. This study has a few limitations. It was retrospective, with some missing data, thus, the Walter Index, GPI, and FRAIL Scale were not assessed for all participants. We performed multiple imputations, and the results showed no significant differences between the pre- and post-imputation data. Our study was not identical to the scales used to assess functional status (ADL and IADL scales) in the GPI and Walter Index, which may have biased the results; however, no studies comparing the modified Katz ADL, PSMS ADL, and Korean ADL and the Lawton IADL and Korean IADL in predicting poor prognosis differences were found. The content of the different scales is summarized in Table S2. An unauthorized version of the Chinese MMSE was used by the study team without permission and has been rectified with PAR. The MMSE is a copyrighted instrument and may not be used or reproduced in whole or in part, in any form or language, or by any means without written permission of PAR ([www.parinc.com](http://www.parinc.com)). Our study population was single-center and small-sample with low 1-year mortality, and the performance of the four models should be further validated in different regions in China and other developing countries.

## Conclusions

The four mortality prediction models exhibited good discriminatory power in predicting the 1-year mortality risk of geriatric inpatients in a geriatric ward in China. However, compared with the CCI, which only assessed disease, the Walter Index, GPI, and FRAIL Scale, which assessed both disease and function, showed better Calibration. Combining the Brier score, discrimination and calibration, the Walter Index was confirmed for the first time to be the best model to predict the 1-year mortality risk of geriatric inpatients in China among the four models.

## Abbreviations

ADL, Activities of daily living; CCI, Charlson Comorbidity Index; CGA, Comprehensive geriatric assessment; CI, Confidence Interval; GDS-15, Geriatric depression scale with 15 items; GPI, Geriatric Prognostic Index; HL test, Hosmer-Lemeshow



goodness-of-fit-test; IADL, Instrumental activities of daily living; IQR, Interquartile range; MMSE, Mini-Mental State Examination; MNA, Mini-Nutritional Assessment; MNA-SF, Mini-Nutritional Assessment Short Form; MPI, Multidimensional prognostic index; OR, Odds ratio; PSMS, Physical Self-Maintenance Scale; ROC, Receiver operating characteristic; SPPB, Short Physical Performance Battery.

## Data Sharing Statement

The datasets generated or analyzed in the current study are not publicly available in order to protect the privacy of the participants, but are available from the corresponding authors on reasonable request.

## Ethics Approval and Informed Consent

Because our study entailed no intervention in patient care, informed consent was waived; however, we informed participants of the purpose of our follow-up and sought consent from them or their families via telephone. The research protocol complied with the Declaration of Helsinki and was certified by the Ethical Committee for Clinical Research of the Peking Union Medical College Hospital (Project number: I-22YJ095).

## Acknowledgments

We acknowledge all the patients, researchers, and staff who were involved in the study.

## Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

## Funding

This work was supported by National High Level Hospital Clinical Research Funding (project number 2022-PUMCH-B-132).

## Disclosure

The authors declare no conflicts of interest in this work.

---

## References

1. Mazzaglia G, Roti L, Corsini G, et al. Screening of older community-dwelling people at risk for death and hospitalization: the Assistenza Socio-Sanitaria in Italia project. *J Am Geriatr Soc.* 2007;55(12):1955–1960. doi:10.1111/j.1532-5415.2007.01446.x
2. Onder G, Palmer K, Navickas R, et al. Time to face the challenge of multimorbidity. A European perspective from the joint action on chronic diseases and promoting healthy ageing across the life cycle (JA-CHRODIS). *Eur J Internal Med.* 2015;26(3):157–159. doi:10.1016/j.ejim.2015.02.020
3. Hov J, Alteren J, Kvigne K. Rehabilitation of the frail older adults in primary healthcare in rural areas: a scoping review protocol. *BMJ open.* 2021;11(6):e048820. doi:10.1136/bmjopen-2021-048820
4. Brownlee S, Korenstein D. Better understanding the downsides of low value healthcare could reduce harm. *BMJ.* 2021;372:n117.
5. Carey E, Covinsky K, Lui L, Eng C, Sands L, Walter L. Prediction of mortality in community-living frail elderly people with long-term care needs. *J Am Geriatr Soc.* 2008;56(1):68–75. doi:10.1111/j.1532-5415.2007.01496.x
6. Walter L, Brand R, Counsell S, et al. Development and validation of a prognostic index for 1-year mortality in older adults after hospitalization. *JAMA.* 2001;285(23):2987–2994. doi:10.1001/jama.285.23.2987
7. Pilotto A, Ferrucci L, Franceschi M, et al. Development and validation of a multidimensional prognostic index for one-year mortality from comprehensive geriatric assessment in hospitalized older patients. *Rejuvenation Res.* 2008;11(1):151–161. doi:10.1089/rej.2007.0569
8. Fischer S, Gozansky W, Sawaia A, Min S, Kutner J, Kramer A. A practical tool to identify patients who may benefit from a palliative approach: the CARING criteria. *J Pain Symptom Management.* 2006;31(4):285–292. doi:10.1016/j.jpainsymman.2005.08.012
9. Teno J, Harrell F, Knaus W, et al. Prediction of survival for older hospitalized patients: the HELP survival model. *J Am Geriatr Soc.* 2000;48(S1):S16–24. doi:10.1111/j.1532-5415.2000.tb03126.x
10. Levine SK, Sachs GA, Jin L, Meltzer D. A prognostic model for 1-year mortality in older adults after hospital discharge. *Am J Med.* 2007;120(5):455–460. doi:10.1016/j.amjmed.2006.09.021
11. Di Bari M, Balzi D, Roberts A, et al. Prognostic stratification of older persons based on simple administrative data: development and validation of the “Silver Code”, to be used in emergency department triage. *J Gerontol a Biol Sci Med Sci.* 2010;65(2):159–164. doi:10.1093/gerona/glp043

12. Dramé M, Novella J, Lang P, et al. Derivation and validation of a mortality-risk index from a cohort of frail elderly patients hospitalised in medical wards via emergencies: the SAFES study. *Eur J Epidemiol.* 2008;23(12):783–791. doi:10.1007/s10654-008-9290-y
13. Jung H, Kim J, Han J, et al. Multidimensional Geriatric Prognostic Index, Based on a Geriatric Assessment, for Long-Term Survival in Older Adults in Korea. *PLoS One.* 2016;11(1):e0147032. doi:10.1371/journal.pone.0147032
14. Soh C, Ul Hassan S, Sacre J, Maier A. Morbidity Measures Predicting Mortality in Inpatients: a Systematic Review. *J Am Med Directors Assoc.* 2020;21(4):462–468.e467. doi:10.1016/j.jamda.2019.12.001
15. Morley J, Malmstrom T, Miller D. A simple frailty questionnaire (FRAIL) predicts outcomes in middle aged African Americans. *J Nutr Health Aging.* 2012;16(7):601–608. doi:10.1007/s12603-012-0084-2
16. Gong S, Qian D, Riazi S, et al. Association Between the FRAIL Scale and Postoperative Complications in Older Surgical Patients: a Systematic Review and Meta-Analysis. *Anesthesia Analgesia.* 2023;136(2):251–261. doi:10.1213/ANE.0000000000006272
17. Abellan van Kan G, Rolland Y, Bergman H, Morley J, Kritchevsky S, Vellas B. The I.A.N.A Task Force on frailty assessment of older people in clinical practice. *J Nutr Health Aging.* 2008;12(1):29–37. doi:10.1007/BF02982161
18. Abellan van Kan G, Rolland YM, Morley JE, Vellas B. Frailty: toward a clinical definition. *J Am Med Dir Assoc.* 2008;9(2):71–72. doi:10.1016/j.jamda.2007.11.005
19. Lawton M, Brody E. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist.* 1969;9(3):179–186. doi:10.1093/geront/9.3\_Part\_1.179
20. Rubenstein L, Harker J, Salvà A, Guigoz Y, Vellas B. Screening for undernutrition in geriatric practice: developing the short-form mini-nutritional assessment (MNA-SF). *J Gerontol a Biol Sci Med Sci.* 2001;56(6):M366–372. doi:10.1093/gerona/56.6.M366
21. Folstein M, Folstein S, McHugh P. “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. *J Psychiatric Res.* 1975;12(3):189–198. doi:10.1016/0022-3956(75)90026-6
22. Haworth Press. *Geriatric Depression Scale (GDS): Recent Evidence and Development of a Shorter Version.* US: Haworth Press; 1986.
23. Chen LK, Woo J, Assantachai P, et al. Asian Working Group for Sarcopenia: 2019 Consensus Update on Sarcopenia Diagnosis and Treatment. *J Am Med Dir Assoc.* 2020;21(3):300–307 e302. doi:10.1016/j.jamda.2019.12.012
24. Guralnik J, Simonsick E, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol.* 1994;49(2):M85–94. doi:10.1093/geronj/49.2.M85
25. Katz S, Ford A, Moskowitz R, Jackson B, Jaffe M. studies of illness in the aged. the index of adl: a standardized measure of biological and psychosocial function. *JAMA.* 1963;185:914–919. doi:10.1001/jama.1963.03060120024016
26. Won YK, Rho YG, Kim SY, et al. The Development of Korean Activities of Daily Living(K-ADL) and Korean Instrumental Activities of Daily Living(K-IADL) Scale. *Korean Geriatr Soc.* 2002;6.
27. Posner B, Jette A, Smith K, Miller D. Nutrition and health risks in the elderly: the nutrition screening initiative. *Am J Public Health.* 1993;83(7):972–978. doi:10.2105/AJPH.83.7.972
28. Charlson M, Pompei P, Ales K, MacKenzie C. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis.* 1987;40(5):373–383. doi:10.1016/0021-9681(87)90171-8
29. Schafer JL. *Analysis of Incomplete Multivariate Data.* CRC press. 1997;444.
30. Rubin DB. Multiple Imputation for Nonresponse in Surveys. *Wiley Series Probability Statistics.* 1987.
31. Rufibach K. Use of Brier score to assess binary predictions. *J Clin Epidemiol.* 2010;63(8):938–939. doi:10.1016/j.jclinepi.2009.11.009
32. Schneider C, Aubert C, Del Giovane C, et al. Comparison of 6 Mortality Risk Scores for Prediction of 1-Year Mortality Risk in Older Adults With Multimorbidity. *JAMA network open.* 2022;5(7):e2223911. doi:10.1001/jamanetworkopen.2022.23911
33. Hanley J, McNeil B. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 1982;143(1):29–36. doi:10.1148/radiology.143.1.7063747
34. DeLong E, DeLong D, Clarke-Pearson D. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics.* 1988;44(3):837–845. doi:10.2307/2531595
35. Liu I, Fernández D. Discussion on Assessing the goodness of fit of logistic regression models in large samples: a modification of the Hosmer-Lemeshow test. *Biometrics.* 2020;76(2):564–568. doi:10.1111/biom.13251
36. Monacelli F, Tafuro M, Molfetta L, et al. Evaluation of prognostic indices in elderly hospitalized patients. *Geriatrics Gerontology Int.* 2017;17(6):1015–1021. doi:10.1111/ggi.12801
37. Eckart A, Struja T, Kutz A, et al. Relationship of Nutritional Status, Inflammation, and Serum Albumin Levels During Acute Illness: a Prospective Study. *Am J Med.* 2020;133(6):713–722.e717. doi:10.1016/j.amjmed.2019.10.031
38. Sullivan D, Johnson L, Dennis R, et al. The Interrelationships among albumin, nutrient intake, and inflammation in elderly recuperative care patients. *J Nutr Health Aging.* 2011;15(4):311–315. doi:10.1007/s12603-010-0297-1
39. Cabrerizo S, Cuadras D, Gomez-Busto F, Artaza-Artabe I, Marin-Ciancas F, Malafarina V. Serum albumin and health in older people: review and meta analysis. *Maturitas.* 2015;81(1):17–27. doi:10.1016/j.maturitas.2015.02.009
40. Jackson J. Cause of death in very old people. *JAMA.* 1983;249(19):2637. doi:10.1001/jama.1983.03330430019009
41. Bai R, Liu Y, Zhang L, Dong W, Bai Z, Zhou M. Projections of future life expectancy in China up to 2035: a modelling study. *Lancet Public Health.* 2023. doi:10.1016/S2468-2667(22)00338-3
42. Chen H, Qian Y, Dong Y, et al. Patterns and changes in life expectancy in China, 1990–2016. *PLoS One.* 2020;15(4):67.
43. Health financing in China. Available from: <https://www.who.int/china/health-topics/health-financing>. Accessed May 20, 2022.
44. Yu H. Universal health insurance coverage for 1.3 billion people: what accounts for China’s success? *Health Policy.* 2015;119(9):1145–1152. doi:10.1016/j.healthpol.2015.07.008
45. The educational level of China’s population has improved significantly in the past ten years. Available from: [http://www.moe.gov.cn/jyb\\_xwfb/s5147/202105/t20210512\\_530993.html](http://www.moe.gov.cn/jyb_xwfb/s5147/202105/t20210512_530993.html). Accessed June 14, 2022.
46. Gao C, Xu J, Liu Y, Yang Y. Nutrition Policy and Healthy China 2030 Building. *Eur. J. Clin. Nutr.* 2021;75(2):238–246. doi:10.1038/s41430-020-00765-6
47. Rodriguez-Martinez A. Height and body-mass index trajectories of school-aged children and adolescents from 1985 to 2019 in 200 countries and territories: a pooled analysis of 2181 population-based studies with 65 million participants. *Lancet.* 2020;396(10261):1511–1524. doi:10.1016/S0140-6736(20)31859-6

Clinical Interventions in Aging

Dovepress

### Publish your work in this journal

Clinical Interventions in Aging is an international, peer-reviewed journal focusing on evidence-based reports on the value or lack thereof of treatments intended to prevent or delay the onset of maladaptive correlates of aging in human beings. This journal is indexed on PubMed Central, MedLine, CAS, Scopus and the Elsevier Bibliographic databases. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/clinical-interventions-in-aging-journal>