Received: 13 December 2022 | Revised: 11 May 2023 | Accepted: 16 May 2023

DOI: 10.1002/hsr2.1309

ORIGINAL RESEARCH

Health Science Reports Open Access WILEY

Prediction models of all-cause mortality among older adults in nursing home setting: A systematic review and meta-analysis

Shengruo Zhang[1,2] | Kehan Zhang[2] | Yan Chen[2,3] | Chenkai Wu[2]

[1]Department of Epidemiology, University of Washington, Seattle, Washington, USA

[2]Global Health Research Center, Duke Kunshan University, Kunshan, Jiangsu, China

[3]School of Public Health, Wuhan University, Wuhan, Hubei, China

Correspondence
Chenkai Wu, Global Health Research Center, Duke Kunshan University, Academic Bldg 3038, 8 Duke Ave, Kunshan, Jiangsu 215316, China.
Email: chenkai.wu@dukekunshan.edu.cn

Funding information
Chinese Ministry of Science and Technology, Grant/Award Number: 2020YFC2005600

Abstract

Background and Aims: Few studies have meta-analyzed different prognostic models developed for older adults, especially nursing home residents. We aimed to systematically review and meta-analyze the performance of all published models that predicted all-cause mortality among older nursing home residents.

Methods: We systematically searched PubMed and EMBASE from the databases' inception to January 1, 2020 to capture studies developing and/or validating a prognostic/prediction model for all-cause mortality among nursing home residents. We then carried out both qualitative and quantitative analyses evaluating these models' risks of bias and applicability.

Results: The systematic search yielded 23,975 articles. We identified 28 indices that predicted the risk of all-cause mortality from 14 days to 39 months among older adults in nursing homes. The most used predictors were age, sex, body weight, swallowing problem, congestive heart failure, shortness of breath, body mass index, and activities of daily living. Of the 28 indices, 8 (29%) and 3 (11%) were internally and externally validated, respectively. None of the indices was validated in more than one cohort. Of the 28 indices, 22 (79%) reported the C-statistic, while only 6 (6%) reported the 95% confidence interval for the C statistic in the development cohorts. In the validation cohorts, 11 (39%) reported the C-statistic and 8 (29%) reported the 95% confidence interval. The meta-analyzed C statistic for all indices is 0.733 (95% prediction interval: 0.669–0.797). All studies/indices had high risks of bias and high concern for applicability according to PROBAST.

Conclusion: We identified 28 indices for predicting all-cause mortality among older nursing home residents. The overall quality of evidence was low due to a high degree of bias and poor reporting of model performance statistics. Before any prediction model could be recommended in routine care, future research is needed to rigorously validate existing prediction models and evaluate their applicability and develop new prediction models.

KEYWORDS

all-cause mortality, meta-analysis, nursing home residents, prediction models, systematic review

1 | INTRODUCTION

The COVID-19 pandemic has had a profound impact on lives across the globe, especially those of older adults. A study from Canada showed that the increasing age was associated with the development of critical illness and death among COVID-19 patients.[1] It has been well-documented that older adults bear disproportionally higher rates of hospitalization and mortality due to COVID-19.[2] Older adults were also found to have worse outcomes and raised rates of functional decline and social isolation when they were afflicted by the COVID-19 pandemic.[3] Nursing home residents, living in relatively confined spaces and usually in poorer health status with multiple comorbidities, were hit hardest by the pandemic. One-third to more than half of the fatalities during the early waves of the COVID-19 epidemic were nursing home patients.[4] Understanding the risk factors for death and developing and implementing a multivariable prediction model is critical for the early identification of nursing home residents at risk, which in turn, could inform the design of person-centered care plans and resource allocations.

Research on prediction models for death has grown rapidly ever since the introduction of the well-known Charlson Comorbidity Index, a model that predicts the 10-year mortality for a patient who may have a range of comorbid conditions, in 1994.[5] More prediction models have been developed more recently to encompass a wider variety of individual factors.[6] Older adults in a wide variety of settings have been studied, such as community, hospital, and nursing home. Despite the increasing amount of efforts to develop prediction models of death among nursing home residents, these models varies greatly in many features, such as population, sample size, length of study, and selection of predictors. In addition, little is known about the quality of prediction indices among this population, limiting their clinical use.

In the present study, we aimed to systematically review and meta-analyze the performance of all published models that predicted all-cause mortality among older nursing home residents. This endeavor would pave the way for the potential development of an accurate and clinically useful index, the use of which could guide and lead to better clinical choices, both diagnostic and therapeutic. This, in turn, could also be used in the decision-making regarding resource allocation in nursing homes.

2 | METHOD

2.1 | Literature search

We systematically searched PubMed and EMBASE from the databases' inception to January 1, 2020 to capture studies developing and/or validating a prognostic/prediction model for all-cause mortality among nursing home residents. In the search algorithm, we included keywords relevant to prediction models and all-cause mortality among older adults. The keywords and detailed search algorithm used were shown in the Supporting Information. Furthermore, we manually searched the reference lists, reviews, and citations of each eligible article to identify additional studies. No restriction regarding publication, period, sex, race/ethnicity, or country was applied in our searches. The language of our searched articles was limited to English. The protocol for this systematic review and meta-analysis was registered on PROSPERO with CRD of 2020 CRD42020165261.

2.2 | Eligibility criteria

We included all studies reporting at least one multivariable model, tool, or index that has been proposed for predicting the risk of all-cause mortality among older adults. Studies were eligible for inclusion if they met both of the following criteria: (1). the average age of the study participants was at least 60 years old; (2). the main outcome of the predictive models, tools, or indices was all-cause mortality. Studies were excluded from our analysis if they fell into any of the following categories: (1) in-vitro or animal model studies; (2) studies that examined intensive care unit (ICU), in-hospital, or disease-specific mortality; (3) case reports, case series, letters, editorials, thesis, reviews, protocols, conferences, or news; and (4) abstracts of articles or book chapters. In each model, we defined the group not developing the outcome of interest—all-cause mortality—as the control/reference group. The type of studies we considered were observational studies, excluding cross-sectional studies, examining the predictive value of a multivariable model for all-cause mortality.

2.3 | Screening process

Two investigators (S. Zhang and Y. Chen) independently screened titles and abstracts and then texts of full-length articles passing the title and abstract screening. Disagreement was resolved by consensus involving a third investigator (C. Wu). We downloaded all the references into the reference management software EndNote (Clarivate Analytics, Philadelphia, and Pennsylvania). To obtain full texts of the articles, we searched through the holdings of Duke University libraries and on the rare occasions where we were not able to find the full texts of the articles this way, we attempted to contact the relevant study authors. Two independent reviewers screened titles and abstracts and examined the full texts as well as any potentially related references and citations for inclusion. During the entire process of the review, any disagreement arisen was discussed and resolved through consensus or adjudicated by a third reviewer.

2.4 | Data extraction

Two investigators (S. Zhang and Y. Chen) conducted the data extraction independently. We constructed a standardized data extraction form based on the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modeling Studies checklist

(CHARMS checklist), which listed relevant items to extract from individual studies in a systematic review of prediction models.[7] The following information was extracted: author name, publication year, study design, study population, geographical location, predicted outcome, prediction horizon (i.e., length of follow-up), modeling method and method of internal and external validation, sample size, basic demographic, and clinical characteristics of study participants (age, sex, race/ethnicity, chronic conditions, and disability), number and type of predictors in the final models, model presentations, and measures of predictive performance. Data were first extracted onto an electronic case record Excel form and were later consolidated and prepared into a data frame in R (R, version 4.0.2). If there were any missing data or unclear information, we attempted to contact the relevant study authors for clarification.

2.5 | Risk of bias assessment and critical appraisal

We used the Prediction Model Risk of Bias Assessment Tool (PROBAST),[8] a standardized tool for assessing multivariable prediction models, to evaluate the risk of bias in four domains: (1) selection of study participants; (2) selection and measurement of predictors; (3) definition and determination of modeling outcome; and (4) statistical analysis. We considered the quality of evidence downgraded if we detected a high risk of bias through PROBAST.

We extracted key data reflecting the predictive performances of models/indices in our chosen studies in the domains of model discrimination, calibration, and classification to conduct quantitative meta-analysis. C-statistic was used to assess model discrimination with a C-statistic < 0.6 considered poor discrimination, 0.6–0.75 considered good discrimination and, >0.75 considered excellent discrimination. Classification, if reported in our chosen studies, were presented in sensitivity, specificity, and predictive values. Calibration, if reported in our chosen studies, were summarized in either calibration plots, Hosmer–Lemeshow statistics or observed to expected ratios.

2.6 | Statistical analysis

Meta-analysis was conducted for all included indices. We first performed a descriptive analysis to capture key study characteristics, to assess the methodology of eligible studies and to summarize the predictive performances of all included models/indices. We implemented the random effects meta-analysis model where weights in the model are based on the within-study error variance rather than the number of events in each study. We applied the 2-sided restricted maximum likelihood estimation for the calculation of summary C-statistics at 0.05 significance level. For the approximate 95% prediction intervals, we used the Hartung-Knapp-Sidik-Jonkman method to calculate the 95% confidence intervals for the average performance. We described the statistical heterogeneity across studies using both the I^2 measure and results from the Q tests. Neither classification nor calibration could be analyzed across studies due to lack of data (only five out of the twenty-one included studies reported statistics for classification or calibration).

To examine whether the study results were distorted by low quality evidence, we performed a subgroup analysis by excluding those studies at high risk of bias (e.g., study populations with pre-existing underlying conditions), to evaluate the robustness of the results. All statistical analyses were performed in R (R, version 4.0.2).

3 | RESULTS

3.1 | Selection process

The PRISMA flow chart of our search process that generated the finalized sample of included studies is shown in Figure 1. The first step of our search process returned 23,975 articles. After multiple rounds of title and abstract screening, 503 papers were eligible for full-text screening, resulting in 241 articles to be included in systematic literature review. We conducted systematic literature review on all 241 articles and identified 21 studies focusing specifically on nursing home residents. From these 21 studies, we identified 28 unique indices, of which 79% ($n = 22$) reported C-statistics for the development cohorts while only 21% ($n = 6$) reported 95% confidence intervals for the C-statistics in the development cohorts. Out of the 28 included indices, 39% ($n = 11$) reported C-statistics for their validation cohorts but only 29% ($n = 8$) reported 95% confidence intervals for their C-statistics in the validation cohorts.

We conducted a sensitivity analysis on a sub-sample of the finalized sample of included indices. This sub-sample is chosen as a group of indices focusing on the general populations, defined as cohorts of populations not chosen for one or more certain specified pre-existing conditions/comorbidities in the studies. We deemed this sub-sample of the indices to be less biased than the total finalized sample based on generalizability. This sub-sample consisted of 64% ($n = 18$) of the total finalized sample. Out of the total number of indices in the sub-sample, 67% ($n = 12$) reported C-statistics for the development cohorts while only 22% ($n = 4$) of the indices reported 95% confidence intervals for the C-statistics in the development cohorts. Out of the included indices in the sub-sample, 33% ($n = 6$) reported C-statistics for their validation cohorts but only 28% ($n = 5$) reported 95% confidence intervals for their C-statistics in the validation cohorts.

3.2 | Summary of findings

3.2.1 | Study designs and index characteristics

Table 1 shows the summary information of each included study. Of the 28 included prediction indices, there were 12 different types: 7-day ($n = 1$, 3.6%), 14-day ($n = 1$, 3.6%), 1-month ($n = 1$,

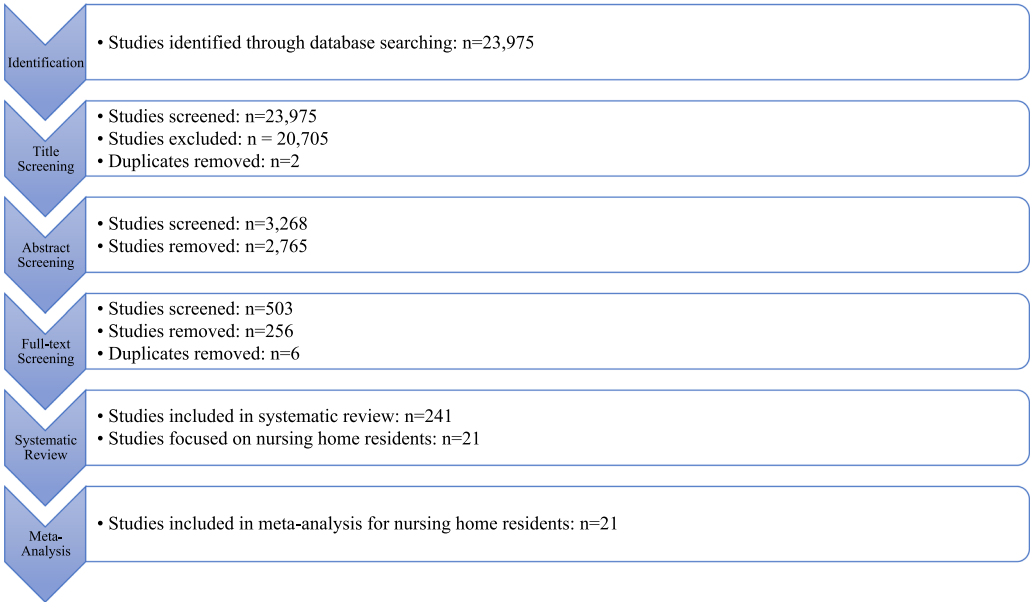

FIGURE 1 Flowchart of literature search for prognostic models for all-cause mortality among older adults in nursing home. The PRISMA flow chart of our search process that generated the finalized sample of included studies.

TABLE 1 Study details and patient characteristics.

References	Design	*N*	Average age	Men, %	Follow-up, months	Mortality rate, %
Falcone et al.[9]	Prospective observational study	446	80[a]	42.4	30 days	28.7
Flacker et al.[10]	Retrospective cohort study	780	88.3	24.1	1 year	7.1–85.7
Harrold et al.[11]	Prospective cohort study	466	78	46	6 months	6–96
Mehr et al.[12]	Prospective cohort study	975	60–90[b]	32.9	30 days	14.7
Mitchell et al.[13]	Prospective/retrospective cohort study	22,405	84.5	23.0	12 months	40.6
Naughton et al.[14]	Retrospective chart review	378	83	34.1	30 days	21.4
Porock et al.[15]	Retrospective cohort study	32,599	65–85+[b]	26.44	6 months	23
Rauh et al.[16]	Cohort study	380	84.2	43	14 days	14
Schoufour et al.[17]	Longitudinal observational study	982	62	51	3 years	14.5
Sharifi et al.[18]	Prospective cohort study	247	65	N/A	39 months	30
Bont et al.[19]	Cohort study	821	64.1	51	30 days	9.5
Chan et al.[20]	Prospective cohort study	585	85.6	34.7	2 years	32.1
Flacker et al.[21]	Retrospective cohort study	52,402	N/A	31	1 year	32.1
Flacker et al.[21]	Prospective cohort study	19,812	N/A	25.9	1 year	21.4
Frisoni et al.[22]	Cohort study	104	82	19.2	18 months	20
Ogarek et al.[23]	Cohort study	1,297,117	81.6	35.7	30 days	0.9–73.7
Ogarek et al.[23]	Cohort study	1,297,117	81.6	35.7	60 days	2.5–78.8
Ogarek et al.[23]	Cohort study	1,297,117	81.6	35.7	1 year	14.5–91.8
Woo et al.[24]	Cohort study	208	78.4	27.9	3 months	13.9

[a]Median age.

[b]Age range.

3.6%), 30-day (*n* = 6, 21.4%), 60-day (*n* = 2, 7.1%), 3-month (*n* = 2, 7.1%), 6-month (*n* = 6, 21.4%), 1-year (*n* = 5, 17.9%), 12-month (*n* = 1, 3.6%), 18-month (*n* = 1, 3.6%), 2-year (*n* = 1, 3.6%) and 3-year (*n* = 1, 3.6%). As for the sub-sample of indices on general population, there were nine different types: 7-day (*n* = 1, 4.8%), 1-month (*n* = 1, 4.8%), 30-day (*n* = 6, 28.6%), 60-day (*n* = 2, 9.5%), 3-month (*n* = 2, 9.5%), 6-month (*n* = 2, 9.5%), 1-year (*n* = 5, 23.8%), 18-month (*n* = 1, 4.8%) and 2-year (*n* = 1, 4.8%).

The number of predictors ranged from 4 to 51. Predictors were generally clinical signs, measurements or observations that could be collected easily and noninvasively.

3.2.2 | Risk of bias

Table 2 shows the PROBAST we used to assess the risk of bias of both the development and validation of the included studies and indices. The PROBAST showed that both the development and validation of all studies and indices were at high risk of bias. We also used the PROBAST to identify several major sources of bias, including: (1). Insufficient, unclear, or nonreporting of key performance statistics and measures (e.g., missing discrimination, calibration, classification); (2). Censoring resulting in an absence of accounting for model complexity; (3). Issues regarding modelling methods (e.g., unclear justification of the selection of candidate variables, complete model selection); (4). Poor/unclear handling of missing data (e.g., not described in writing clearly or in details in the literature of the studies); (5). Lack of either internal or external validation.

The PROBAST also showed that all studies/indices had high concern for applicability due to the selection of study populations with restrictive inclusion criteria and unclear exclusion criteria, as well as the selection of a specified timeline for the studies.

3.2.3 | Prediction model performance

Indices reported their discrimination performances using the C-statistics. C-statistics reported in the total finalized sample of included indices ranged from good to excellent with no indices in the poor range, suggesting satisfactory discrimination performances across all studies and indices (Tables 3 and 4). As for the C-statistics reported in the sub-sample of included indices on general population, discrimination performances also ranged from good to excellent with no indices in the poor range.

We cannot analyze either classification or calibration due to lack of data and insufficient reporting from the studies. Only five out of the 21 eligible studies reported measures of classification or calibration.

TABLE 2 Risk of bias assessment using the PROBAST tool.

			Risk of bias				Applicability			
Study	Year	Population	Predictors	Outcome	Analysis	Overall	Population	Predictors	Outcome	Overall
Falcone et al.[9]	2018	+	−	−	+	+	+	−	−	+
Flacker et al.[10]	1998	−	−	−	+	+	−	−	−	+
Harrold et al.[11]	2005	−	−	−	+	+	−	−	−	+
Mehr et al.[12]	2001	+	−	−	+	+	+	−	−	+
Mitchell et al.[13]	2010	+	?	+	+	+	+	?	+	+
Naughton et al.[14]	2000	+	−	−	+	+	+	−	−	+
Porock et al.[15]	2005	−	−	−	+	+	−	−	−	+
Rauh et al.[16]	2019	+	−	−	+	+	+	−	−	+
Schoufour et al.[17]	2015	+	+	+	+	+	+	+	+	+
Sharifi et al.[18]	2012	−	?	+	+	+	−	?	+	+
Bont et al.[19]	2008	−	+	+	+	+	−	+	+	+
Chan et al.[20]	2012	+	−	−	+	+	+	−	−	+
Flacker et al.[21]	2003	−	−	−	+	+	−	−	−	+
Frisoni et al.[22]	1994	−	−	−	+	+	−	−	−	+
Ogarek et al.[23]	2018	−	−	−	+	+	−	−	−	+
Woo et al.[24]	1989	−	−	−	+	+	−	−	−	+

Note: −, indicates low risk of bias/low concern regarding applicability; +, indicates high risk of bias/high concern regarding applicability;?, Indicates unclear risk of bias/unclear concern regarding applicability.

TABLE 3 Discrimination performances and heterogeneity assessment.

	Development		Validation	
	All indices	Indices on general population	All indices	Indices on general population
Discrimination performances				
Poor	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Good	13 (47%)	8 (45%)	6 (21%)	3 (17%)
Excellent	9 (32%)	4 (22%)	5 (18%)	3 (17%)
NA	6 (21%)	6 (33%)	17 (61%)	12 (67%)
C-statistic, CI	0.733 (0.669–0.797)	0.709 (0.652–0.766)		
I^2 measure	99.89%	99.87%	99.89%	99.90%
Q Test	860.2548, p < 0.0001	829.1577, p < 0.0001	5562.8, p < 0.0001	5249.0167, p < 0.0001

3.3 | Statistical analysis

Of the 28 indices reported in 21 studies, 8 (28.6%) and 3 (10.7%) were internally and externally validated, respectively. None of the indices was validated in more than one cohort.

In the assessment of summary measures and heterogeneity for the development and validation cohorts, we have found that, for the development cohort, the summary C-statistic calculated for all indices [0.733; Approximate 95% Prediction Interval: (0.669–0.797)] outperformed than that calculated for indices on general population only [0.709; Approximate 95% Prediction Interval: (0.652–0.766)], with high heterogeneity (I^2 > 99%). As for the validation cohort, the summary C-statistic calculated for all indices [0.719; Approximate 95% Prediction Interval: (0.675–0.762)] did not outperform than that calculated for indices on general population only [0.722; Approximate 95% Prediction Interval: (0.678–0.766)], also with high heterogeneity (I^2 > 99%).

4 | DISCUSSION

4.1 | Main findings and interpretations

Our systematic review and meta-analysis included all prognostic models and indices for all-cause mortality developed or validated for nursing home residents from 1989 to 2020. 28 unique prognostic indices from 21 different studies were considered and analyzed in this systematic review and meta-analysis. Very few (n = 3) indices were validated externally and none of the 28 included prognostic indices were validated in more than one cohort. Discrimination performances of the individual all-cause mortality prognostic indices for nursing home residents are overall satisfactory, ranging from good to excellent with no indices in the poor range. Calibration for the included prognostic indices was poorly reported.

A half number of the papers included in this literature review were developed in global settings, such as Italy, Hong Kong, Iran, and the Netherlands. The Italian population was selected from Italian Long-term care facilities, and nursing homes.[9,22] The research studies from Hong Kong enrolled older Chinese residents from Hong Kong's nursing homes or chronic care institutions.[20,24] Three studies from the Netherlands used data from the PneuMonitor Study, or collected data from Dutch care provider services and general practitioners.[16,17,19] An Iran study used older adults' data from Kahrizak Elderly Study in Kahrizak Charity Foundation.[18]

The Minimum Data Set (MDS), a clinical and administrative data set that is legally required in all US nursing homes, was used or examined by nine indices. To identify mortality-related factors to establish long-term care residents, the MDS information was examined to estimate mortality.[10] In the multivariate proportional hazard regression, the study's authors discovered eight variables were linked to 1-year mortality. These included swallowing issues, congestive heart failure, old age, male gender, weight loss, shortness of breath, and functional impairment. Another study also obtained and measured potential risk factors from MDS, such as depression and delirium, to predict 30-day all-cause mortality.[12] With the Advanced Dementia Prognostic Tool, a study identified the residents with dementia using MDS assessments and obtained mortality from Medicare records to determine 12-month survival.[13] MDS as an instrument was used in a study from Missouri to evaluate to predict 6-month deaths involving death certificate data.[15] A few other studies used a similar way to get both MDS and National Death Index or Medicare Master Beneficiary Summary File to develop a tool to estimate risk for mortality.[21,23]

4.2 | Challenges and opportunities

Our systematic review and meta-analysis yielded useful results that could assist clinicians and researchers in their future efforts of

TABLE 4 Summary of the performance of the multicomponent prognostic models.

Prognostic tool/model	Variables	Study	N	Discriminative ability	Period	Mortality
Multivariable logistic regression	Malnutrition, bilateral pneumonia, acute mental status deterioration, hypotension, PaO2/FiO2 ratio	Falcone et al.[9]	446	AUROC	30 days	28.7
Mortality score	functional impairment, weight loss, shortness of breath, male gender, low body mass index, swallowing problems, congestive heart failure, and advanced age	Flacker et al.[10]	780	AUROC	1 year	7.1–85.7
PPS	Site of care, diagnosis	Harrold et al.[11]	466	AUROC	6 months	6–96
Point scale	Serum urea nitrogen, white blood cell count, body mass index, pulse rate, activities of daily living status, absolute lymphocyte count, male sex, deterioration in mood	Mehr et al.[12]	1406	c-statistic	30 days	14.7
ADEPT score	Length of stay, age, male, dyspnea, pressure ulcers, total functional dependence, bedfast, insufficient intake, bowel incontinence, body mass index, weight loss, and congestive heart failure	Mitchell et al.[13]	2,333,662	AUROC	12 months	40.6
Scoring system	Respiratory rate, pulse, altered mental status, history of dementia	Naughton et al.[14]	378	AUC	30 days	21.4
CPS score	Demographics, diseases, clinical signs and symptoms, adverse events	Porock et al.[15]	43,510	Kaplan–Meier survival curves	6 months	23
Hosmer-Lemeshow statistics	Nutritional status, dehydration, bowel incontinence, increase in eating dependency, cardiovascular history, dressing dependency, mobility dependency, bedfast, coughing, aspiration in the month before diagnosis of pneumonia, chronic Obstructive Pulmonary Disease diagnosis	Rauh et al.[16]	380	AUC	14 days	14
FI	risk of death, independent of sex, age, level of ID, and Down syndrome	Schoufour et al.[17]	982	ROC curve	3 years	14.5
OPMI score	BI, age, hemoglobin, mid-arm circumference	Sharifi et al.[18]	247	AUC	39 months	30
CRB-65 Score severity assessment tool	New mental confusion, respiratory rate, Blood pressure, age	Bont et al.[19]	1135	PPV, NPV, sensitivity, specificity, AUC	30 days	9.5
Score system	Age, Charlson comorbidity index, Barthel Index	Chan et al.[20]	1120	AUROC	2 years	32.1
Mortality risk index score	Cancer, shortness of breath, congestive heart failure, bedfast, male, unstable conditions, food uneaten, low functional ability score, swallowing problem, bowel incontinence, body mass index	Flacker et al.[21]	52,402 and 19,812	AUROC	1 year	32.1 and 21.4
A dimentional measure	Cholesterol, lymphocytecount, mid-arm circumference, hemoglobin, age, and gender	Frisoni et al.[22]	104	sensitivity, specificity	18 months	20
MDS-CHESS 3.0 scores	Clinician determination of life expectancy, ADL, Cognitive Function Scale, Aggressive Behavior Scale, moderate impairment in daily decision-making, remaining health conditions, pressure ulcers	Ogarek et al.[23]	1,297,117	Survival curve	30 days 60 days 1 year	0.9–73.7 2.5–78.8 14.5–91.8
Discriminant function score	Plasma fructosamine, transferrin, glycosylated haemoglobin, prealbumin, haemoglobin	Woo et al.[24]	208	Sensitivity, Specificity, Predictive value	3 months	13.9

studying the diverse prognostic models and indices for nursing home residents as well as giving them a comprehensive guideline of tools to try to identify, in different settings with different objectives and time lengths, nursing home residents at serious risks of potential adverse outcomes. Our systematic review and meta-analysis also pointed out, given the high levels of heterogeneity and risks of bias for all the prognostic indices of all-cause mortality currently developed and validated for nursing home residents, the difficulties in identifying these nursing home residents at risks of potential adverse outcomes both reliably and consistently.

4.3 | Strengths and limitations

Our systematic review and meta-analysis have several notable strengths. First, to our knowledge, this is the first ever systematic review and meta-analysis conducted for the prognostic models of all-cause mortality specifically for older nursing home residents. Given the generally poorer health conditions of nursing home residents and their relatively bigger risks of developing adverse health outcomes, our systematic review and meta-analysis provide insights particularly meaningful for any future development of similar prognostic models as well as for the design and implementation of evidence-based interventions in protecting this vulnerable population group in our rapidly aging world. Second, we conducted a broad search of existing literatures pertaining to the prognostic models of nursing home residents, encompassing models developed using various sets of different clinical risk factors identified by said models. Third, we then carried out detailed examination of the evidence and comprehensive meta-analyses, both qualitative and quantitative, across these different multivariable prognostic models, mapping out their characteristics, examining their development and validation methodologies as well as synthesizing their statistical modelling performances. Fourth, we used the most updated tools for our analyses including the recently published PROBAST risk of bias tool. Our systematic review and meta-analysis made possible the future potential development of a composite index for nursing home residents, drawing on the strengths and addressing the limitations and weaknesses of the existing prognostic models and indices.

Despite the numerous strengths listed, we identified several challenges and limitations for our systematic review and meta-analysis. For our systematic review, most notably, our review is limited by both a high degree of bias and the poor reporting of available data. Most studies failed in the complete reporting of one or more of the following categories: (a). the justification or rationale for the selection of candidate predictors; (b). the justification or rationale for the development and validation methodologies; (c). the justification or rationale for the selection of modeling predictors included in the finalized model; (d). the justification or rationale for the handling of missing data. In addition, there is no identified authorized prognostic tool to predict older adults' mortality. The prognostic models were used inconsistently in the literature. The heterogeneity of included studies was high. The pooled results should be interpreted cautiously. Nevertheless, the present study was among the first to provide an overview of prediction model for mortality among nursing home residents. These results provided a solid foundation for future studies.

As for our meta-analysis, the most prominent challenge we observed was the poor reporting of performance characteristics, including both calibration and classification measures. This lack of available data, caused by the omissions in prognostic models' performance statistics, has led to a low or very low grading of the overall quality of the evidence. Only 3 indices were validated externally. None of the included prognostic indices was validated in more than one cohort, internal or otherwise. Moreover, a high degree of heterogeneity was observed among the prognostic models ($I^2 > 99\%$), which could be due to the variations in country of cohort, study inclusion criteria as well as the prediction horizon. Different countries of cohorts represent different health care and nursing home systems as well as different nursing home populations, contributing to this high degree of heterogeneity across prognostic models. We were not able to carry out a meta-regression analysis to further examine the effects of different study inclusion criteria on modeling performances due to the insufficient number of studies. This high degree of heterogeneity posed difficulties and obstacles for the generalization and applicability of the prediction models.

4.4 | Implications for policy and future research

To address these challenges and limitations, future studies should first determine a unified scoring system as the tools and focus on the development of new prognostic models. The current level of the total number of studies ($n = 21$) remains low in this niche space. More research attention needs to be brought upon this vulnerable group of population, nursing home residents, in this world of global aging. When developing new models, inclusion criteria need to be broad and recalibration should be used to ensure the generalizability of study results. When reporting for discrimination, calibration and classification measures, researchers should make sure to follow the general practice and usual recommendations regarding the conduct and evaluation of prognostic model validation studies to include complete reporting of the performance statistics. Besides the development of new prognostic models, future studies could also be carried out to validate, update and expand existing models, both rigorously and extensively, with good performance statistics, especially discrimination.

5 | CONCLUSIONS

In summary, overall quality of evidence was low due to a combination of factors ranging from a high degree of bias to the poor reporting of performance statistics. As a result, further research and study is necessary before any prognostic model for nursing home residents could be recommended and used for research or clinical practice.

AUTHOR CONTRIBUTIONS

Shengruo Zhang: Data curation; formal analysis; investigation; methodology; writing—original draft; writing—review and editing. **Kehan Zhang**: Formal analysis; investigation; methodology; writing—original draft; writing—review and editing. **Yan Chen**: Formal analysis; methodology; writing—review and editing. **Chenkai Wu**: Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; supervision; writing—original draft; writing—review and editing.

ACKNOWLEDGMENTS

All authors have read and approved the final version of the manuscript. Chenkai Wu had full access to all of the data in this study and takes complete responsibility for the integrity of the data and the accuracy of the data analysis. This research was supported by the Chinese Ministry of Science and Technology (Grant number: 2020YFC2005600. The sponsor had no role in the design of this study and will not have any role in study design, collection, analysis, and interpretation of data; writing of the report; the decision to submit the report for publication.

DATA AVAILABILITY STATEMENT

Data available on request from the authors

TRANSPARENCY STATEMENT

The lead author Chenkai Wu affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

ORCID

Chenkai Wu http://orcid.org/0000-0002-2256-0653

REFERENCES

1. Murthy S, Archambault PM, Atique A, et al. Characteristics and outcomes of patients with COVID-19 admitted to hospital and intensive care in the first phase of the pandemic in Canada: a national cohort study. *CMAJ Open*. 2021;9(1):E181-E188. doi:10.9778/cmajo.20200250
2. Shahid Z, Kalayanamitra R, McClafferty B, et al. COVID-19 and older adults: what we know. *J Am Geriatr Soc*. 2020;68(5):926-929. doi:10.1111/jgs.16472
3. Kaushik R, Ferrante LE. Long-term recovery after critical illness in older adults. *Curr Opin Crit Care*. 2022;28(5):572-580. doi:10.1097/MCC.0000000000000981
4. Dyer AH, Fallon A, Noonan C, et al. Managing the impact of COVID-19 in nursing homes and long-term care facilities: an update. *J Am Med Dir Assoc*. 2022;23(9):1590-1602. doi:10.1016/j.jamda.2022.06.028
5. Charlson ME, Pompei P, Ales KL, MacKenzie CR. 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
6. Yourman LC, Lee SJ, Schonberg MA, Widera EW, Smith AK. Prognostic indices for older adults: a systematic review. *JAMA*. 2012;307(2):182-192. doi:10.1001/jama.2011.1966
7. Moons KGM, de Groot JAH, Bouwmeester W, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med*. 2014;11(10):e1001744. doi:10.1371/journal.pmed.1001744
8. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med*. 2019;170(1):51-58. doi:10.7326/M18-1376
9. Falcone M, Russo A, Gentiloni Silverj F, et al. Predictors of mortality in nursing-home residents with pneumonia: a multicentre study. *Clin Microbiol Infect*. 2018;24(1):72-77. doi:10.1016/j.cmi.2017.05.023
10. Flacker JM, Kiely DK. A practical approach to identifying mortality-related factors in established long-term care residents. *J Am Geriatr Soc*. 1998;46(8):1012-1015. doi:10.1111/j.1532-5415.1998.tb02759.x
11. Harrold J, Rickerson E, Carroll JT, et al. Is the palliative performance scale a useful predictor of mortality in a heterogeneous hospice population. *J Palliat Med*. 2005;8(3):503-509. doi:10.1089/jpm.2005.8.503
12. Mehr DR, Binder EF, Kruse RL, et al. Predicting mortality in nursing home residents with lower respiratory tract infection: the Missouri LRI study. *JAMA*. 2001;286(19):2427-2436. doi:10.1001/jama.286.19.2427
13. Mitchell SL, Miller SC, Teno JM, Davis RB, Shaffer ML. The advanced dementia prognostic tool: a risk score to estimate survival in nursing home residents with advanced dementia. *J Pain Symptom Manage*. 2010;40(5):639-651. doi:10.1016/j.jpainsymman.2010.02.014
14. Naughton BJ, Mylotte JM, Tayara A. Outcome of nursing home-acquired pneumonia: derivation and application of a practical model to predict 30 day mortality. *J Am Geriatr Soc*. 2000;48(10):1292-1299. doi:10.1111/j.1532-5415.2000.tb02604.x
15. Porock D, Parker Oliver D, Zweig S, et al. Predicting death in the nursing home: development and validation of the 6-month minimum data set mortality risk index. *J Gerontol A Biol Sci Med Sci*. 2005;60(4):491-498. doi:10.1093/gerona/60.4.491
16. Rauh SP, Heymans MW, van der Maaden T, et al. Predicting mortality in nursing home residents with dementia and pneumonia treated with antibiotics: validation of a prediction model in a more recent population. *The Journals of Gerontology: Series A*. 2019;74(12):1922-1928. doi:10.1093/gerona/gly260
17. Schoufour JD, Mitnitski A, Rockwood K, Evenhuis HM, Echteld MA. Predicting 3-year survival in older people with intellectual disabilities using a frailty index. *J Am Geriatr Soc*. 2015;63(3):531-536. doi:10.1111/jgs.13239
18. Sharifi F, Ghaderpanahi M, Fakhrzadeh H, et al. Older people's mortality index: development of a practical model for prediction of mortality in nursing homes (Kahrizak Elderly Study): older people's mortality index. *Geriat Gerontol Internat*. 2012;12(1):36-45. doi:10.1111/j.1447-0594.2011.00724.x
19. Bont J, Hak E, Hoes AW, et al. Predicting death in elderly patients with community-acquired pneumonia: a prospective validation study reevaluating the CRB-65 severity assessment tool. *Arch Intern Med*. 2008;168(13):1465-1468. doi:10.1001/archinte.168.13.1465
20. Chan TC, Shea YF, Luk KH, Chan HW, Chu LW. Development and validation of a prognostic index for 2-year mortality in Chinese older residents living in nursing homes. *Geriat Gerontol Internat*. 2012;12(3):555-562. doi:10.1111/j.1447-0594.2011.00804.x
21. Flacker JM, Kiely DK. Mortality-related factors and 1-year survival in nursing home residents. *J Am Geriatr Soc*. 2003;51(2):213-221. doi:10.1046/j.1532-5415.2003.51060.x
22. Frisoni GB, Franzoni S, Rozzini R, Ferrucci L, Boffelli S, Trablucchi M. A nutritional index predicting mortality in the nursing home. *J Am*

Geriatr Soc. 1994;42(11):1167-1172. doi:10.1111/j.1532-5415.1994.tb06983.x

23. Ogarek JA, McCreedy EM, Thomas KS, Teno JM, Gozalo PL. Minimum data set changes in health, End-Stage disease and symptoms and signs scale: a revised measure to predict mortality in nursing home residents. *J Am Geriatr Soc*. 2018;66(5):976-981. doi:10.1111/jgs.15305
24. Woo J, Chan SM, Mak YT, Swaminathan R. Biochemical predictors of short term mortality in elderly residents of chronic care institutions. *J Clin Pathol*. 1989;42(12):1241-1245. doi:10.1136/jcp.42.12.1241

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Zhang S, Zhang K, Chen Y, Wu C. Prediction models of all-cause mortality among older adults in nursing home setting: a systematic review and meta-analysis. *Health Sci Rep*. 2023;6:e1309. doi:10.1002/hsr2.1309