Journal of Clinical Epidemiology 168 (2024) 111270

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

External validation of six COVID-19 prognostic models for predicting mortality risk in older populations in a hospital, primary care, and nursing home setting

Anum Zahra[a,*], Maarten van Smeden[a], Evertine J. Abbink[b], Jesse M. van den Berg[c,d,e], Marieke T. Blom[c,d], Carline J. van den Dries[a], Jacobijn Gussekloo[f], Fenne Wouters[g,h], Karlijn J. Joling[g,h], René Melis[i], Simon P. Mooijaart[j], Jeannette B. Peters[k], Harmke A. Polinder-Bos[l], Bas F.M. van Raaij[j], Brent Appelman[m], Hannah M. la Roi-Teeuw[a], Karel G.M. Moons[a], Kim Luijken[a], on behalf of the COOP Consortium

[a]*Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands*
[b]*Department of Internal Medicine, Radboud University Medical Center, Nijmegen, The Netherlands*
[c]*Department of General Practice, Amsterdam UMC Location Vrije Universiteit Amsterdam, Amsterdam, The Netherlands*
[d]*Amsterdam Public Health, Health Behaviors & Chronic Diseases, Amsterdam, The Netherlands*
[e]*PHARMO Institute for Drug Outcomes Research, Utrecht, The Netherlands*
[f]*Section Gerontology and Geriatrics, LUMC Center for Medicine for Older People & Department of Public Health and Primary Care & Department of Internal Medicine, Leiden University Medical Center, Leiden, The Netherlands*
[g]*Department of Medicine for Older People, Amsterdam UMC, Location Vrije Universiteit Amsterdam, De Boelelaan 1117, Amsterdam, The Netherlands*
[h]*Amsterdam Public Health Research Institute, Aging & Later Life, Amsterdam, The Netherlands*
[i]*Department of Geriatric Medicine, Radboud University Medical Center, Nijmegen, The Netherlands*
[j]*LUMC Center for Medicine for Older People, LUMC, Leiden, The Netherlands*
[k]*Department of Pulmonary Diseases, Radboud University Medical Center, Radboud Institute for Health Sciences, Nijmegen, The Netherlands*
[l]*Section of Geriatric Medicine, Department of Internal Medicine, Erasmus MC, University Medical Center Rotterdam, The Netherlands*
[m]*Amsterdam UMC Location University of Amsterdam, Center for Experimental and Molecular Medicine, Amsterdam, The Netherlands*

Accepted 26 January 2024; Published online 2 February 2024

Abstract

Objectives: To systematically evaluate the performance of COVID-19 prognostic models and scores for mortality risk in older populations across three health-care settings: hospitals, primary care, and nursing homes.

Study Design and Setting: This retrospective external validation study included 14,092 older individuals of ≥70 years of age with a clinical or polymerase chain reaction-confirmed COVID-19 diagnosis from March 2020 to December 2020. The six validation cohorts include three hospital-based (CliniCo, COVID-OLD, COVID-PREDICT), two primary care-based (Julius General Practitioners Network/Academisch network huisartsgeneeskunde/Network of Academic general Practitioners, PHARMO), and one nursing home cohort (YSIS) in the Netherlands. Based on a living systematic review of COVID-19 prediction models using Prediction model Risk Of Bias ASsessment Tool for quality and risk of bias assessment and considering predictor availability in validation cohorts, we selected six prognostic models predicting mortality risk in adults with COVID-19 infection (GAL-COVID-19 mortality, 4C Mortality Score, National Early Warning Score 2-extended model, Xie model, Wang clinical model, and CURB65 score). All six prognostic models were validated in the hospital cohorts and the GAL-COVID-19 mortality model was validated in all three healthcare settings. The primary outcome was in-hospital mortality for hospitals and 28-day mortality for primary care and nursing home settings. Model performance was evaluated in each validation cohort separately in terms of discrimination, calibration, and decision curves. An intercept update was performed in models indicating miscalibration followed by predictive performance re-evaluation.

Main Outcome Measure: In-hospital mortality for hospitals and 28-day mortality for primary care and nursing home setting.

Results: All six prognostic models performed poorly and showed miscalibration in the older population cohorts. In the hospital settings, model performance ranged from calibration-in-the-large −1.45 to 7.46, calibration slopes 0.24−0.81, and C-statistic 0.55−0.71 with 4C

Funding: This work was supported by Zorg Onderzoek Nederland en Medische Wetenschappen (ZonMw) to the COVID-19 Outcomes in Older People (COOP) study (project number 10430102110005) under the COVID-19 program.

* Corresponding author. PO Box 85500, Stratenum 6.131, 3508 GA Utrecht, The Netherlands. Tel.: +31638441115.
E-mail address: a.zahra@umcutrecht.nl (A. Zahra).

https://doi.org/10.1016/j.jclinepi.2024.111270

Mortality Score performing as the most discriminative and well-calibrated model. Performance across health-care settings was similar for the GAL-COVID-19 model, with a calibration-in-the-large in the range of −2.35 to −0.15 indicating overestimation, calibration slopes of 0.24−0.81 indicating signs of overfitting, and C-statistic of 0.55−0.71.

Conclusion: Our results show that most prognostic models for predicting mortality risk performed poorly in the older population with COVID-19, in each health-care setting: hospital, primary care, and nursing home settings. Insights into factors influencing predictive model performance in the older population are needed for pandemic preparedness and reliable prognostication of health-related outcomes in this demographic.

Keywords: COVID-19; COVID-19-Related mortality; External validation; Older population; Prognostic models; clinical prediction models

1. Introduction

The COVID-19 pandemic has had a significant impact on the social, economic, and health-care sectors with an estimated death toll of over 6.9 million lives globally (till December 2023) [1]. The older population faced the highest risk of severe COVID-19 disease leading to hospitalization, mortality, and morbidity [2]. In the Netherlands, 89% of all deaths caused by COVID-19 disease occurred in the older population (70 years or older) even though they make up only 14% of the total population (till December 2022) [3]. Similar trends in mortality were reported in other parts of the world like the United States [4], Europe [5], and low−middle-income countries (eg, India, Mexico, Pakistan, Philippines, and South Africa) [6].

In response to the COVID-19 pandemic, a large number of predictive tools were developed. Among these tools are clinical prognostic models, which aim to provide insight into COVID-19 patient risks for severe outcomes for both physicians and patients [7]. For the valid application of these prognostic models, an external validation initially developed in the general population, especially in a high-risk older population, is crucial.

A living systematic review assessing the quality of COVID-19 prediction models found most of these models of poor quality and high risk of bias [8]. Although some models appraised at low risk of bias have been externally validated in the general population [9−13], validation in an older population (including an age-based subgroup analysis during model development) has so far not been done. Extensive assessment including validation studies is needed to evaluate their quality, accuracy, and implementation feasibility for future infectious diseases, especially in this high-risk population. Validation of prognostic models in an older population with COVID-19 infection gives insight into the role and relevance of the included predictors like age, and comorbidities.

In this comprehensive external validation study, we assessed the predictive performance of six COVID-19 prognostic models for mortality risk in a population of ≥70 years of age presenting with COVID-19, across various settings that is, hospital, primary care, and nursing homes.

2. Methods

The protocol for this study is publicly available [14]. We note that the protocol mentioned the validation of eight models, but Sepsis-related organ Failure assessment (SOFA) [15] and Acute Physiology and Chronic Health Evaluation-II score [16] could not be validated due to the unavailability of data on predictors included in these two models, which are predictors that are typically measured in an intensive care setting and are not recorded in the hospital cohorts included in this study (Supplementary file 1). We adhered to the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosisreporting guidelines [17] (Supplementary file 14).

2.1. Selection of COVID-19 prognostic models

Prediction models were systematically reviewed using Prediction model Risk Of Bias ASsessment Tool (a quality or risk of bias assessment tool) [18,19] in the fourth and final update of the living systematic review of diagnostic and prognostic prediction models for COVID-19 [8]. Using these findings, we identified all prognostic models that predict the risk of mortality in individuals with COVID-19 infection that were rated at uncertain or low risk of bias. Fifteen candidate models met this criterion of which seven prognostic models (Pandemic Respiratory Infection Emergency System Triage [20], Clinical frailty scale, Urea, Consolidation, Age, FiO_2, Sex, Respiratory rate [12], CUCA-SF [12], Acute Physiology and Chronic Health Evaluation-II [16], SOFA score [15], and QCOVID [21] for males and females) were not included for validation because data on certain predictors were unavailable in the six cohorts of older individuals. Two prognostic scores (Quick Sequential Organ Failure Assessment [22] and National Early Warning Score [NEWS] [23] were excluded because they express risk of mortality qualitatively rather than quantitatively using a formal probability or risk estimation [14] (Supplementary file 1).

This resulted in a selection of six prognostic models being externally validated in our older population cohorts across the three settings. Of these six models, five were COVID-19−specific (GAL COVID-19 mortality model [24], 4C Mortality Score [25], NEWS2-extended model

What is new?

Key findings

- External validation of existing COVID-19 models in older populations in hospital, primary care, and nursing home setting demonstrated poor prognostic performance.
- Overall, the 4C Mortality Score performed as the most discriminative and moderately calibrated prognostic model to predict in-hospital mortality.

What this adds to what was known?

- External validation of existing COVID-19 models in older populations across various health-care settings demonstrated poor prognostic performance.

What is the implication and what should change now?

- Most prognostic models demonstrated poor performance. Approaches for optimally developing accurate prediction models in the older population need to be further investigated.
- During prediction model development, modeling of the functional form of age as a predictor (eg, linear or nonlinear) requires extra consideration so that predictions are available to an older population as well.

[26], Xie model [27], Wang clinical model [28]). One prognostic sore (CURB65 [29]) existed before the COVID-19 pandemic and is used for the prediction of in-hospital mortality risk after admission for respiratory infections or sepsis (Table 1).

2.2. Validation cohorts

Data were collected from six cohorts (three hospital cohorts, two primary care cohorts, and one nursing home cohort) representing older individuals ($\geq$ 70 years of age) presenting with COVID-19 infection in the Netherlands. COVID-19 infection was defined as clinical diagnosis (when reverse transcription polymerase chain reaction testing was not yet available) or reverse transcription polymerase chain reaction confirmed COVID-19. The moment of inclusion was at first presentation with COVID-19 infection in the respective health-care settings from March 2020 to December 2021. Details on participant inclusion in each cohort are described in Supplementary files 2 to 7.

2.3. Outcome

The original outcome of all six prediction models was mortality with varied prediction horizons (Table 1). The outcome of our validation study was defined as in-hospital mortality in hospital cohorts and 28-day mortality in the primary care and nursing home cohorts.

2.4. Predictors

Prognostic models for validation included predictors such as demographics, comorbidities, and laboratory parameters. Predictor measurement definitions and timing of the measurement for the six prognostic models were extracted and matched, as closely as possible, to the original predictor measurement procedures outlined in the original publications (Supplementary files 2 to 7).

2.5. Statistical analysis

We externally validated the six selected COVID-19 prognostic models in six cohorts of older patients with COVID-19 from March 2020 to December 2020 to assess their predictive performance when transported from a general adult population to a specific older population. The GAL-COVID-19 model could be assessed across the three health-care settings (originally derived in a primary care setting) [30]. The 4C Mortality Score, NEWS2+ model, Xie model, Wang clinical model, and confusion, urea, respiratory rate, blood pressure, age above or below 65 score were developed in hospitalized populations and validated in the same health-care setting. Evaluation and assessment of the predictive performance were performed in each cohort separately. All statistical analyses were performed in R Statistical Software version 4.2.2 [31].

2.5.1. Descriptives

To assess similarities and differences between the derivation and validation study populations, descriptives of the validation population cohorts were compared to descriptives of the model development studies [32].

2.5.2. Missing data

Missing data was described to determine possible reasons for and patterns of missingness [33]. There were no missing data in primary care settings as the absence of a comorbidity record in electronic health records was interpreted as its absence. Predictor data in the nursing home cohort were missing for only four individuals in 2020, hence a complete case analysis was performed. Multiple imputation by chained equations using the Full Conditional Specification or Joint Modeling [34] was used to handle missing data in hospital cohorts in all available data at once (ie, in year 2020 and 2021 if available). All variables and outcomes in all six prognostic models were included in the imputation model to ensure compatibility. Based on the level of missingness, a total of 50 imputed datasets (with 50 iterations) were generated for CliniCo and COVID-PREDICT cohorts and 70 imputations were used for the COVID-OLD cohort for all relevant variables

Table 1. Overview of validated prognostic models

Model name	Derivation country	Pre-existing or COVID-19–specific	Derivation health-care setting	Derivation population	Age Median (IQR)	Intended moment of use	Predicted outcome	Predictors	Model type
GAL-COVID-19-mortality model	Spain	COVID-19–specific model	Primary care	Adults (≥18 yr) with confirmed COVID-19 diagnosis	58.0 (20.0)[a]	First presentation with COVID-19 infection at general practitioner	Mortality (no prediction horizon reported)	• Age • Sex • Lymphoma/leukemia • Liver disease • Dementia • Ischemic heart disease • Chronic obstructive pulmonary disease • Diabetes mellitus • Chronic kidney disease	Prediction model
4C-Mortality Score	United Kingdom	COVID-19–specific model	Hospital	Adults (≥18 yr) with confirmed COVID-19 diagnosis	73.0 (59.0–83.0)	At hospital admission for COVID-19 infection	In-hospital mortality	• Age • Sex • Respiratory rate • Peripheral oxygen saturation on room air • Glasgow Coma Scale • Urea • C-reactive protein • Number of comorbidities (counted as chronic cardiac disease, chronic respiratory disease (excluding asthma), chronic renal disease, liver disease, dementia, chronic neurological conditions, connective tissue disease, diabetes mellitus, HIV or AIDS, malignancy, obesity	Points-based score (Model equation available)
NEWS2+ model	United Kingdom	Pre-existing risk stratification score updated for COVID-19 patients	Hospital	Adults (≥18 yr) admitted to hospital with a confirmed COVID-19 diagnosis	71.5 (57.1–82.6)	• At hospital admission for non-nosocomial patients (ie, community acquired COVID infection) • At the date of symptom onset for nosocomial patients. If the date of onset	ICU admission or death within 14 days of admission	• Age • Peripheral oxygen saturation • Heart rate • Systolic blood pressure • Body temperature • Alertness • Supplemental oxygen flow rate • Urea • C-reactive protein • Estimated glomerular filtration rate • Neutrophil count	Prediction model

(*Continued*)

Table 1. Continued

Model name	Derivation country	Pre-existing or COVID-19–specific	Derivation health-care setting	Derivation population	Age Median (IQR)	Intended moment of use	Predicted outcome	Predictors	Model type
						was unavailable the date of positive SARS-CoV-2 RT-PCR minus 4 days was used instead		• Neutrophil/lymphocyte ratio	
Xie model	China	COVID-19–specific model	Hospital	Adults (≥18 yr) admitted to hospital with a confirmed COVID-19 diagnosis	65.0 (54.0–73.0)	At hospital admission for COVID-19 infection	In-hospital mortality	• Age • Lactate dehydrogenase • Lymphocyte count • Oxygen saturation	Prediction model
Wang clinical model	China	COVID-19–specific model	Hospital	Adults (≥18 yr) admitted to hospital with a confirmed COVID-19 diagnosis. Pregnant women were excluded	47.3 (15.0)[a]	At hospital admission for COVID-19 infection	In-hospital mortality	• Age • History of hypertension • History of coronary heart disease	Prediction model
CURB-65	United Kingdom, New Zealand, The Netherlands	Pre-existing risk stratification score	Hospital	Patients with community-acquired pneumonia	64.1 (NR)	For triage at the emergency department	Mortality (30 days)	• Age • Alertness (new confusion) • Urea • Respiratory rate • Systolic blood pressure • Diastolic blood pressure	Points-based score

Abbreviations: IQR, interquartile range; ICU, intensive care unit; RT-PCR, reverse transcription polymerase chain reaction; NR; not reported; CURB-65, confusion, urea, respiratory rate, blood pressure, age above or below 65; NEWS, National Early Warning Score.

[a] Age given as mean (standard deviation).

[35]. Individuals with a missing outcome were excluded after imputation [36].

2.5.3. Assessment of predictive performance

Each prognostic model was applied in accordance with the descriptions provided by the original authors. Predictive performance was evaluated in terms of discrimination (the model's ability to distinguish individuals who died after presentation with COVID-19 diagnosis from those who did not) [37] and calibration (the agreement between predicted and observed mortality risks) in each cohort [38]. Discrimination was assessed by quantifying the area under the receiver operating characteristic curve that is, the C-statistic [39], and where applicable pooled over imputed datasets on a log scale using Rubin's rules [35].

Calibration was assessed by visualizing calibration of expected vs. observed risk using locally estimated scatterplot smoothing-smoothed plots, where applicable on stacked imputed data sets [38]. The 4C Mortality model, GAL-COVID-19 mortality model, NEWS2+ model, Xie model, and Wang clinical model are model equations. For these models, calibration was assessed in terms of the calibration-in-the-large coefficient and calibration slope [38]. The coefficients were again pooled over imputed datasets on a log scale using Rubin's rules where applicable [35].

2.5.4. Decision curve analysis

Decision curve analyses were performed to quantify the pooled net benefit (across all imputed datasets) achieved by each model for predicting the originally intended endpoint across a range of risk thresholds ranging from zero to one [40].

2.5.5. Updating

Prediction models showing miscalibration were adjusted using an intercept update from March 2020 to December 2021. The recalibrated model was reassessed from the period of March 2020 to December 2020.

2.5.6. Additional analyses

Two additional analyses were performed in all the settings except the CliniCo hospital cohort (unavailability of data for the year 2021) to assess the change in the predictive performance of the six COVID-19 prognostic models over time. The external validation was performed using data from March 2020 to December 2021 and January 2021 to December 2021. Additionally, predictive performance in estimating the 90-day mortality risk was evaluated in primary care and nursing home cohorts.

2.6. Patient and public involvement

The *COVID-19 Outcomes in Older People* consortium is a national collaboration in the Netherlands between researchers, health-care professionals, and a seniors advisory board with members of the public. The seniors advisory board was involved in the research consortium from the moment of grant writing onwards to discuss preferred research topics and study designs. The ongoing discussions take place at regular consortia meetings.

3. Results

3.1. Description of the study cohorts

The study comprised of six validation cohorts with a total of 14,092 older participants presenting with COVID-19 infection from March to December 2020 in the primary analysis. This included 6,203 participants from the hospital settings, 6,171 from general practices, and 1,718 from the nursing home settings. The median age ranged from 77 to 79 years in the hospital cohorts, from 77 to 78 years in the primary care cohorts, and nursing home participants had a median age of 89 years. The mortality fraction ranged from 41% (YSIS nursing home cohort) to 3% (PHARMO primary care cohort). Participant characteristics are presented in Table 2 (hospital cohorts) and Table 3 (primary care and nursing home cohorts) and in Supplementary files 2 to 7 per cohort.

3.2. Comparison between development and validation population

The median age in the validation cohorts was often higher than in the derivation cohorts, except for the development population of the 4C Mortality Score and NEWS2+ model, where the median age (>70 years) was comparable (Supplementary files 2 to 4). The incidence of mortality and comorbidities was generally higher in the validation cohorts compared to the derivation cohorts.

3.3. Missing data in study cohorts

In the hospital cohorts, missingness in predictor values ranged from 0% to 26% in CliniCo, 0% to 64% in COVID-OLD, and 0% to 45% in COVID-PREDICT (Supplementary files 2 to 7). Multiple imputation was performed as planned. In the CliniCo cohort, the outcome mortality was not recorded in 36 participants who were excluded from the analysis (after multiple imputation) [36]. In the nursing home cohort, 35 participants were excluded from the analysis due to missing values for mortality outcome ($n = 31$) and absence of a value for all comorbidities ($n = 4$).

3.4. Predictive performance

Overall, most models displayed poor calibration across the settings with a systematic overestimation of the average mortality risk (calibration-in-the-large coefficient less than zero) and exhibited extremes in the distribution of estimated risks compared to observed risks (calibration slope

Table 2. Characteristics of the older population in the imputed hospital validation cohorts in the year 2020

Demographics	CliniCo	COVID-PREDICT	COVID-OLD
Total participants	591	3,115	2,497
Mortality (%)	239 (40)	723 (23)	833 (34)
Age (yr), median (IQR)	77 (73–82)	78 (74–84)	79 (74–84)
Male, *n* (%)	384 (65)	1,842 (59)	1,518 (61)
Comorbidities			
Chronic cardiac disease, *n* (%)	444 (75)	1,174 (38)	-
Chronic kidney disease, *n* (%)	83 (14)	479 (16)	1,074 (43)
Chronic liver disease, *n* (%)	13 (2)	21 (<1)	35 (1)
Chronic neurological disease, *n* (%)	103 (17)	530 (17)	-
Chronic pulmonary disease, *n* (%)	161 (27)	-	-
Malignancy, *n* (%)	138 (23)	306 (10)	459 (18)
History of diabetes, *n* (%)	156 (26)	976 (31)	794 (32)
History of dementia, *n* (%)	30 (5)	232 (7)	222 (9)
Clinical frailty scale, *n* (%)			
Nonfrail	247 (42)	1,336 (43)	1,101 (44)
Prefrail	217 (37)	978 (31)	677 (27)
Frail	127 (21)	795 (26)	719 (29)
Disease severity indicators			
Immunocompromised, *n* (%)	127 (21)	120 (4)	-
Temperature (°C), median (IQR)	37.0 (37.0–38.5)	37.5 (36.7–38.0)	37.7 (37.0–38.5)
Respiratory rate (breaths/min), median (IQR)	24 (19–27)	22 (18–26)	21 (18–26)
Oxygen saturation (%), median (IQR)	94 (92–96)	95 (92–97)	93 (89–96)
Lymphocytes (10^9/L), median (IQR)	0.8 (0.6–1.1)	0.9 (0.6–1.3)	0.9 (0.6–1.3)
Creatinine (micromole/L), median (IQR)	92 (70–122)	92 (72–122)	94 (75–130)
Lactic acid dehydrogenase (U/L), median (IQR)	359 (278–457)	308 (244–399)	315 (238–422)
C-reactive protein (mg/L), median (IQR)	92 (47–150)	72 (28–130)	75 (37–133)

Abbreviation: IQR; interquartile range.
"-" indicating variables not recorded in the hospital cohorts.

Table 3. Characteristics of the older population in the primary care and nursing home cohorts in the year 2020 (No imputation performed)

	Primary care setting		Nursing home setting
Demographics	JGPN/ANH/AHA	PHARMO	YSIS
All participants	1,444	4,727	1,718
Mortality (%)	212 (9)	163 (3)	699 (41)
Age (yr), median (IQR)	78 (73–83)	77 (73–82)	89 (84–94)
Male, *n* (%)	657 (46)	2,154 (46)	588 (34)
Comorbidities			
Chronic kidney disease, *n* (%)	340 (24)	906 (19)	326 (19)
Chronic liver disease, *n* (%)	32 (2)	95 (2)	10 (<1)
History of diabetes, *n* (%)	509 (35)	1,256 (27)	321 (19)
History of dementia, *n* (%)	93 (6)	399 (8)	1,249 (73)
Frailty index, median (IQR)	0.3 (0.22–0.38)	0.1 (0.08–0.2)	-
Lymphoma/leukemia, *n* (%)	31 (2)	85 (2)	7 (<1)
Ischemic heart disease, *n* (%)	382 (27)	1,000 (21)	269 (16)
Chronic obstructive pulmonary disease, *n* (%)	203 (14)	552 (12)	243 (14)

Abbreviations: IQR; interquartile range; JGPN, Julius General Practitioners Network; ANH, Academisch network huisartsgeneeskunde; AHA, Network of Academic general Practitioners.
"-" indicating variables not recorded in the nursing home cohort.

less than one). Fig. 1 summarizes the predictive performance of the prognostic models in the validation cohorts in the year 2020. The 4C Mortality Score showed the best predictive performance in comparison to other prognostic models in hospital settings (calibration-in-the-large: −0.78 to 0.03, calibration slopes: 0.82−1.15, and C-statistic: 0.66−0.74) (Figs. 1 and 2).

3.4.1. Calibration

Calibration-in-the-large varied across models and settings. Most models overestimated the overall risk of mortality (calibration-in-the-large coefficient less than zero) except the NEWS2+ and GAL COVID-19 models (in hospital settings). The calibration-in-the-large coefficient of the NEWS2+ model indicated extreme underestimation of the average mortality risk (7.04, 95% confidence interval: 6.62−7.45).

A low calibration slope estimate was observed in most validation cohorts implying a trend of too extreme predicted risks compared to observed risks (too low for low-outcome risks and too high for high-outcome risks) (Fig. 1). This is suggestive of model overfitting. Smooth calibration curves showed poor calibration indicating that the predicted probabilities did not align effectively with the observed probabilities of mortality in the validation population. This is shown in Fig. 3 for the GAL COVID-19 model and in Supplementary files 2 to 7 for all models.

3.4.2. Discrimination

The discrimination varied from poor to moderate across the hospital cohorts, where the C-statistic ranges were: Wang model (0.57−0.60), GAL mortality model (0.57−0.61), confusion, urea, respiratory rate, blood pressure, age above or below 65 (0.60−0.65), Xie model (0.64−0.66) and 4C Mortality Score (0.66−0.74). In primary care cohorts, the GAL COVID-19 mortality model displayed moderate discriminative ability with a C-statistic ranging from 0.70 to 0.71, while the C-statistic was 0.55 in the nursing home cohort, indicating poor discriminatory ability.

3.4.3. Decision curve analysis

Decision curve analysis showed that most models performed worse than the treat-all strategy, especially across a lower (clinically relevant) threshold range (0−20%) (Supplementary file 8).

3.5. Additional analyses

3.5.1. Performance in January 2021 to December 2021 and March 2020 to December 2021

The additional analysis in which the models were validated in data from January 2021 to December 2021 was performed with a total of 9,574 participants (1,956 participants from the hospital settings, 6,531 from general practices, and 1,087 from the nursing home settings). 4C Mortality Score remained the most discriminative model while prognostic performance did not improve in all cohorts and models persistently showed miscalibration and low discrimination (Supplementary Files 9 and 10).

3.5.2. Performance with 90-day mortality in primary care settings

Additional analysis with 90-day mortality as the outcome in the primary care and nursing home setting showed that the predictive performance of the GAL-COVID-19 model remained poor (Supplementary File 11).

3.6. Model updating

The model intercept was updated as planned but predictive performance was not restored in most models (except the GAL COVID-19 model in Julius General Practitioners Network/Academisch network huisartsgeneeskunde/ Network of Academic general Practitioners and PHARMO cohort) as indicated by updated smooth calibration curves (see Supplementary File 12).

4. Discussion

In our comprehensive external validation study assessing the predictive performance of six existing COVID-19 prognostic models in older patients with COVID-19, all prognostic models performed poorly irrespective of the healthcare settings, particularly in terms of calibration. The 4C Mortality Score appeared as the most discriminative (C-statistic in hospital cohorts: 0.66−0.74) and moderately calibrated (calibration slopes: 0.82−1.15, calibration-in-the-large: −0.78 to 0.03) for predicting in-hospital mortality after COVID-19 infection among the validated models.

A decrease in prognostic performance at external validation in the older population was expected, due to differences in the incidence of baseline mortality risk and predictors between the development and validation populations [41], as well as homogeneity in ranges of important predictors such as age and comorbidities within the validation population. Similarly, additional analysis in the year 2021 showed a further deterioration in the performance of all models that can be attributed to newer (less severe) COVID-19 variants, targeted vaccination in older populations, and changes in treatment protocols.

One other explanation for the suboptimal performance of the validated prognostic models can be the modeling decisions made regarding the functional form of the predictor age. The prognostic models evaluated in this study predominantly modeled age using a linear term only, rather than a more flexible functional form such as a quadratic term or splines. This is reflected in a large proportion of estimated in-hospital or 28-day mortality risks having (unrealistic) extreme values close to 100% (Figs. 2 and 3). These observations highlight the significance of including suitable

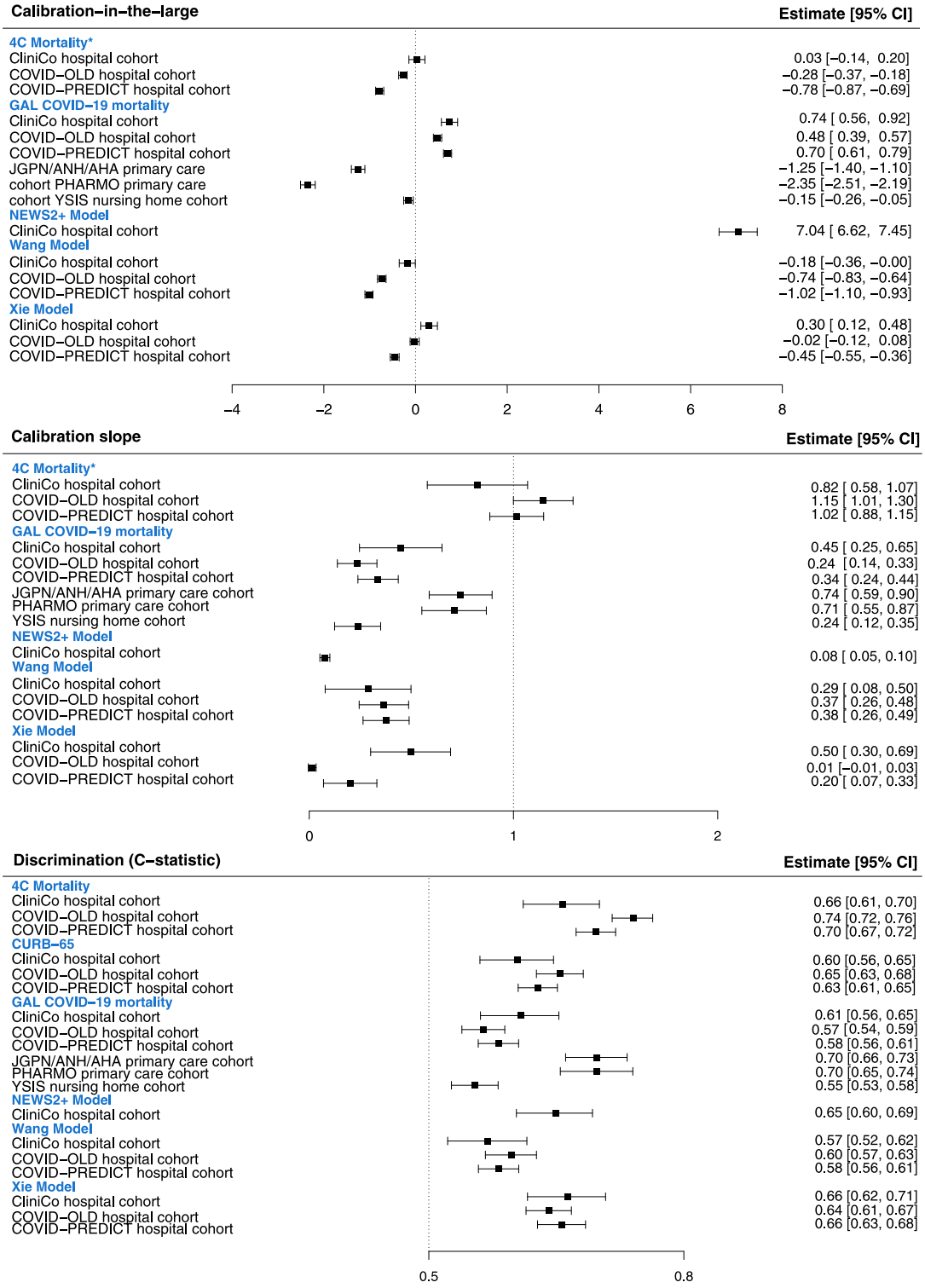

Fig. 1. Predictive performance of the prognostic models in the year 2020 (‘‘*’’ 4C Mortality model used for calculating Calibration-in-the-large and Calibration slope.

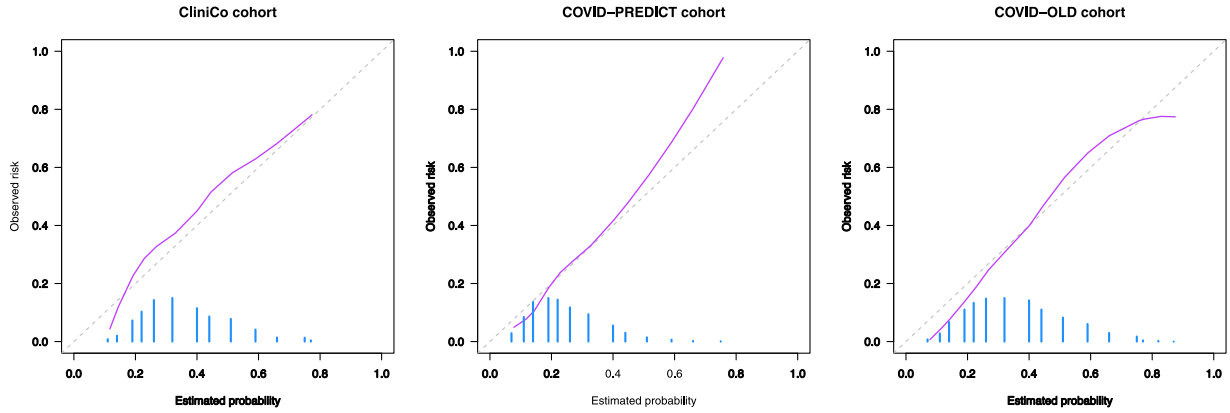

Fig. 2. Performance of 4C mortality score in hospital cohorts in validation cohorts in 2020 (Calibration is shown using locally weighted smoothing (locally estimated scatterplot smoothing) across stacked.

functional forms of predictors like age during prediction model development, particularly those designed for prognostication in an older population [42].

In the NEWS2+ model, the presence of multicollinearity among predictors (urea, glomerular filtration rate, neutrophil count, and neutrophil-lymphocyte ratio) impacted predictor effects (coefficients and variances), potentially contributed to miscalibration as the differences in levels of collinearity between the development and validation dataset can influence the predictive performance [42,43].

COVID-19 prognostic models (like Xie model [27], Wang model [28], and 4C Mortality Score [25] that previously performed well in certain population-wide external validation studies [10,13] displayed a reduction in predictive accuracy when validated in an older population in the current study. In the hospital cohorts, the 4C Mortality Score demonstrated higher discrimination and calibration

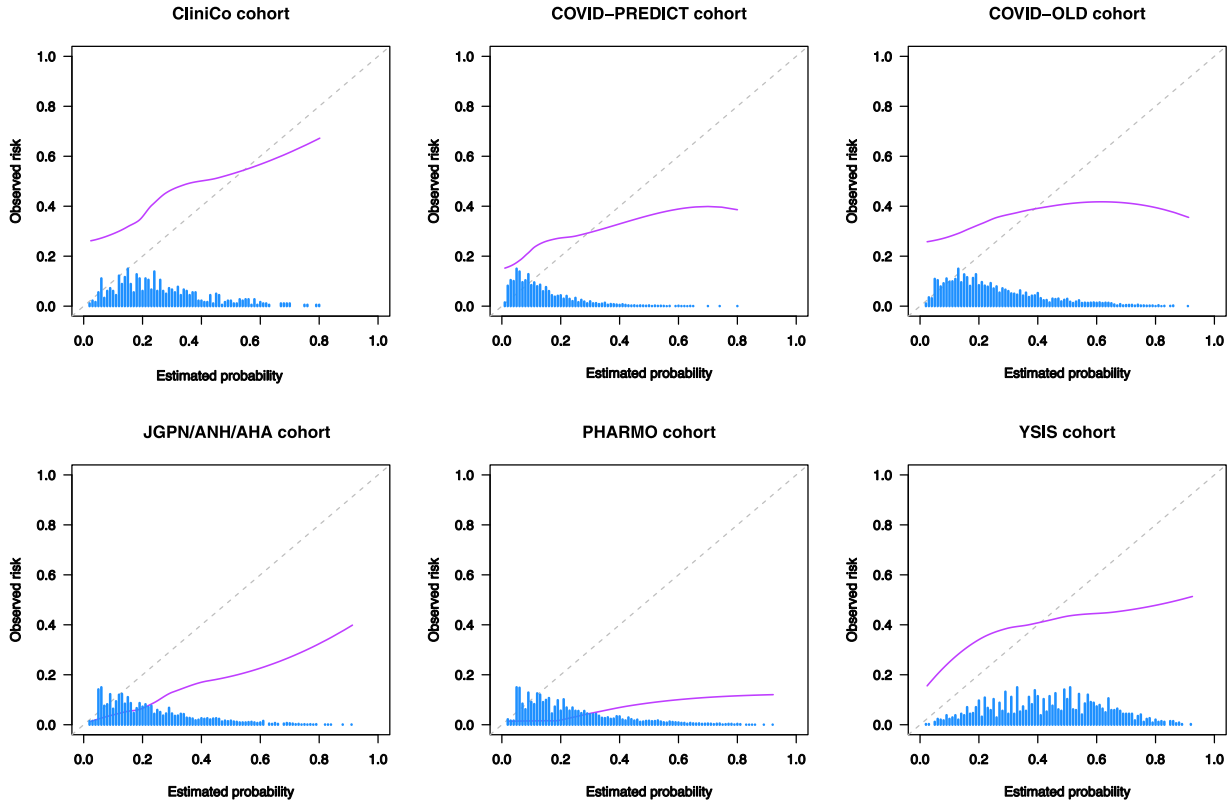

Fig. 3. Predictive performance of the GAL COVID-19 mortality model in all validation cohorts in 2020 (Calibration is shown using locally weighted.

among the other validated models in year 2020 (under similar conditions of model development) and 2021 (after vaccinations were introduced). Systematic reviews of COVID-19 prognostic models [8,44] and external validation studies in other countries like the United States [45], Canada [46], Europe [10,47,48], and (older adults) in Italy [49] have also found 4C Mortality Score as a promising tool for predicting mortality risk in COVID-19 patients. One possible reason for the 4C Mortality Score's better performance relative to other models is its robust development, which involved a large population, half of whom were represented by older individuals [25]. Therefore, of all the COVID-19 models, the 4C Mortality Score has the most potential to be considered for implementation in clinical settings after further targeted (temporal) validation [50] and model updating [51].

A limitation of the current study is that it could not validate all the low-risk-of-bias COVID-19 prognostic models identified in the living systematic review [8] and only one model could be validated in all three settings. This limitation was encountered because predictor information, primarily related to intensive care, was often unavailable in the validation cohorts, especially primary care and nursing home settings. This limitation also highlights the clinical relevance of including widely available clinical tests as predictors when developing prognostic models intended for clinical use in different health-care settings, given the variation in predictor availability across settings.

Another limitation of the study is related to the underreporting of certain predictors and outcome mortality. In the nursing home cohort, comorbidities were identified using free text searches in the medical history of electronic health records, potentially underestimating disease prevalence when comorbidities were not documented. This comorbidity underreporting may create nondifferential misclassification of comorbidities that potentially impairs calibration. Additionally, mortality may have been underreported in the PHARMO cohort. The mortality fraction was lower compared to the Julius General Practitioners Network/Academisch network huisartsgeneeskunde/Network of Academic general Practitioners primary care cohort. This underreporting might have affected predictive performance in the PHARMO cohort.

5. Conclusion

External validation of existing COVID-19 models in older populations across various health-care settings demonstrated poor prognostic performance. In the future, researchers could direct their efforts toward determining the optimal approach for developing prognostic models intended for clinical use in the older population. The current study has highlighted the importance of appropriately incorporating age during model development to generate reliable predictions, particularly for the older population. It is currently unclear whether the development of prognostic models should prioritize more generalizable models (applicable to a broader range of infectious diseases or age groups) or whether future prognostic models should be exclusively developed within the older population. Future studies should further explore if a shift toward age-dependent factors can improve prognostication in the older population. This can be done by including geriatrics-focused indicators and predictors, such as multimorbidity [52] and frailty [53,54], in future prognostic models.

Patient consent for publication

Not applicable.

Ethics approval

This study will be conducted in accordance with the EU General Data Protection Regulation. Since this study does not require direct patient or physician involvement, the need for formal ethical reviewing was waived by the local medical research ethics committee (MREC Utrecht, the Netherlands), according to Dutch law.

CRediT authorship contribution statement

Anum Zahra: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation. **Maarten van Smeden:** Writing – review & editing, Supervision, Project administration, Methodology, Conceptualization. **Evertine J. Abbink:** Writing – review & editing, Data curation. **Jesse M. van den Berg:** Writing – review & editing, Investigation, Data curation. **Marieke T. Blom:** Writing – review & editing, Investigation, Data curation. **Carline J. van den Dries:** Writing – review & editing, Investigation, Data curation. **Jacobijn Gussekloo:** Writing – review & editing. **Fenne Wouters:** Writing – review & editing, Data curation. **Karlijn J. Joling:** Writing – review & editing, Data curation. **René Melis:** Writing – review & editing, Investigation, Data curation. **Simon P. Mooijaart:** Project administration, Funding acquisition. **Jeannette B. Peters:** Writing – review & editing, Data curation. **Harmke A. Polinder-Bos:** Writing – review & editing, Investigation, Data curation. **Bas F.M. van Raaij:** Writing – review & editing, Investigation, Data curation. **Brent Appelman:** Writing – review & editing, Investigation, Data curation. **Hannah M. la Roi-Teeuw:** Writing – review & editing, Investigation, Data curation. **Karel G.M. Moons:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Kim Luijken:** Writing – review & editing, Supervision, Methodology, Investigation, Conceptualization.

Declaration of Generative artificial intelligence and artificial intelligence-assisted technologies in the writing process

During the preparation of this work, the author(s) used GTP-3.5 to improve readability and language. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Data availability

Data will be made available on request.

Declaration of competing interest

All authors have completed the International Committee of Medical Journals Editors uniform disclosure form at https://www.icmje.org/disclosure-of-interest/ and declare: funding from the Netherlands Organisation for Scientific Research (ZonMw). KJ and FW have received grants from the *program Leren van Data by the Dutch Ministry of Health, Welfare and Sport (grant number: 329517)*; all declare no support from any organization for the submitted work; no financial relationships with any organizations that might have an interest in the submitted work in the previous 3 years; no other relationships or activities that could appear to have influenced the submitted work.

Acknowledgments

The COVID-19 Outcomes in Older People (COOP)-consortium is a national collaboration in the Netherlands between stakeholders from different care settings (hospitals, primary care practices, and nursing homes) and a Seniors Advisory Board (Ouderenraad). The authors acknowledge COVID-19 outcomes in older people (COOP) consortium investigators: Evertine J. Abbink, Wilco P. Achterberg, Sifra H. van de Beek, Marian Beekman, Ludo F.M Beenen, Jesse *M. van* den Berg, Marieke T. Blom, Bram van den Borst, Sebastiaan J.H. Bredie, Frederiek van den Bos, Virgil A.S.H Dalm, Carline J. van den Dries, Petra J.M. Elders, Miriam C. Faes, Jan Festen, Geert-Jan Geersing, Jacobijn Gussekloo, Miriam L. Haaksma, Vanessa C. Harris, Ron M.C Herings, Cees M.P.M Hertogh, Jacobien J Hoogerwerf, Jeannette B. Peters, Steffy Jansen, Karlijn J. Joling, Anneke G.Julien, Veerle M.G.T.H van der Klei, Anna Kuranova, P. Hugo *M. van* der Kuy, Carolien M.J van der Linden, Anouk *M. van* Loon, Kim Luijken, Josephine S. van de Maat, Francesco U.S Mattace Raso, René J.F. Melis, Julia Minnema, Simon P. Mooijaart, Dennis O. Mook-Kanamori, Karel G.M. Moons, Mihai G. Netea, Geeske Peeters, Harmke A. Polinder-Bos, Bas F.M van Raaij, Roos S.G Sablerolles, P. Eline Slagboom, Maarten van Smeden, Rosalinde A.L. Smits, Annemieke Smorenberg, Lisanne Tap, Hannah M. la Roi-Teeuw, Lisa S. van Tol, Hanna C. Willems, and Anum Zahra. We are grateful to the Seniors Advisory Board (Ouderenraad) for their diverse participation as representatives of older persons and for their helpful feedback and insights throughout the entire project. Additionally, collaborative efforts within the study cohorts are recognized, including the CliniCo cohort represented by Evertine Abbink, Ton Dofferhoff, Jacobien Hoogerwerf, Angele Kerckhoffs, Josephine van der Maat, Mihai Netea, Kitty Slieker, Karin Veerman. The PHARMO cohort contributors include Jesse van den Berg, Karin Swart-Polinder and Ron Herings. For the COVID-PREDICT cohort contributors are Brent Appelman, Daisy Rusch, Niels C Gritters van den Oever, Suat Simsek, Frits H.*M. van* Osch, Martijn D de Kruif, Renée Douma, Hazra Moeniralam, Kees Brinkman, Nejma Bokhizzou, Helen Leavis, Martijn Beudel. The COVID-OLD cohort contributors, include Simon Mooijaart, Jan Festen, Harmke A. Polinder−Bos, Carolien van der Linden, Steffy Jansen, Hanna Willems, Jessica van der Bol.

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jclinepi.2024.111270.

References

[1] WHO. WHO Coronavirus(COVID-19) dashboard 2022 [Overview]. Available at: https://covid19.who.int. Accessed December 19, 2023.
[2] De Pue S, Gillebert C, Dierckx E, Vanderhasselt MA, De Raedt R, Van den Bussche E. The impact of the COVID-19 pandemic on well-being and cognitive functioning of older adults. Sci Rep 2021;11(1):4636.
[3] Rijksoverheid. Sterfte 2022. Available at: https://coronadashboard.rijksoverheid.nl/landelijk/sterfte. Accessed December 19, 2024.
[4] (U.S.) NCHS. COVID-19 mortality in adults aged 65 and over: United States, 2020 2022: Report on CDC (Centres of Disease Control and Prevention) website NCHS (National Center of Health Statistics) Data brief number 446- October 2022.
[5] Eurofound. COVID-19 and older people: Impact on their lives, support and care. Luxembourg: Publications Office of the European Union; 2022.
[6] Lloyd-Sherlock P, Sempe L, McKee M, Guntupalli A. Problems of data availability and quality for COVID-19 and older people in low- and middle-income countries. Gerontol 2021;61:141−4.
[7] van Smeden M, Reitsma JB, Riley RD, Collins GS, Moons KG. Clinical prediction models: diagnosis versus prognosis. J Clin Epidemiol 2021;132:142−5.
[8] Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. BMJ 2020;369:m1328.
[9] Gupta RK, Marks M, Samuels THA, Luintel A, Rampling T, Chowdhury H, et al. Systematic evaluation and external validation of 22 prognostic models among hospitalised adults with COVID-19: an observational cohort study. Eur Respir J 2020;56(6):2003498.
[10] de Jong VMT, Rousset RZ, Antonio-Villa NE, Buenen AG, Van Calster B, Bello-Chavolla OY, et al. Clinical prediction models for

mortality in patients with covid-19: external validation and individual participant data meta-analysis. BMJ 2022;378:e069881.
[11] Knight SR, Gupta RK, Ho A, Pius R, Buchan I, Carson G, et al. Prospective validation of the 4C prognostic models for adults hospitalised with COVID-19 using the ISARIC WHO Clinical Characterisation Protocol. Thorax 2022;77(6):606−15.
[12] Bradley P, Frost F, Tharmaratnam K, Wootton DG, NW Collaborative Organisation for Respiratory Research. Utility of established prognostic scores in COVID-19 hospital admissions: multicentre prospective evaluation of CURB-65, NEWS2 and qSOFA. BMJ Open Respir Res 2020;7(1):e000729.
[13] Wickstrom KE, Vitelli V, Carr E, Holten AR, Bendayan R, Reiner AH, et al. Regional performance variation in external validation of four prediction models for severity of COVID-19 at hospital admission: an observational multi-centre cohort study. PLoS One 2021;16:e0255748.
[14] Zahra A, Luijken K, Abbink EJ, van den Berg JM, Blom MT, Elders P, et al. A study protocol of external validation of eight COVID-19 prognostic models for predicting mortality risk in older populations in a hospital, primary care, and nursing home setting. Diagn Progn Res 2023;7(1):8.
[15] Vincent JL, Moreno R, Takala J, Willatts S, De Mendonça A, Bruining H, et al. The SOFA (Sepsis-related organ failure assessment) score to describe organ dysfunction/failure. On behalf of the working group on sepsis-related problems of the European society of intensive care medicine. Intensive Care Med 1996;22:707−10.
[16] Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. Crit Care Med 1985;13(10):818−29.
[17] Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. BMC Med 2015;13:1.
[18] Moons KGM, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. Ann Intern Med 2019;170:W1−33.
[19] Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med 2019;170:51−8.
[20] Goodacre S, Thomas B, Sutton L, Burnsall M, Lee E, Bradburn M, et al. Derivation and validation of a clinical severity score for acutely ill adults with suspected COVID-19: the PRIEST observational cohort study. PLoS One 2021;16:e0245840.
[21] Clift AK, Coupland CAC, Keogh RH, Diaz-Ordaz K, Williamson E, Harrison EM, et al. Living risk prediction algorithm (QCOVID) for risk of hospital admission and mortality from coronavirus 19 in adults: national derivation and validation cohort study. BMJ 2020;371:m3731.
[22] Seymour CW, Liu VX, Iwashyna TJ, Brunkhorst FM, Rea TD, Scherag A, et al. Assessment of clinical criteria for sepsis: for the third international consensus definitions for sepsis and septic shock (Sepsis-3). JAMA 2016;315:762−74.
[23] London R. National early warning score (NEWS): standardising the assessment of acute-illness severity in the NHS—report of a working party. London: Royal College of Physicians; 2012.
[24] Gude-Sampedro F, Fernandez-Merino C, Ferreiro L, Lado-Baleato O, Espasandin-Dominguez J, Hervada X, et al. Development and validation of a prognostic model based on comorbidities to predict COVID-19 severity: a population-based study. Int J Epidemiol 2021;50:64−74.
[25] Knight SR, Ho A, Pius R, Buchan I, Carson G, Drake TM, et al. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score. BMJ 2020;370:m3339.
[26] Carr E, Bendayan R, Bean D, Stammers M, Wang W, Zhang H, et al. Evaluation and improvement of the national early warning score (NEWS2) for COVID-19: a multi-hospital study. BMC Med 2021;19(1):23.
[27] Xie J, Hungerford D, Chen H, Abrams ST, Li S, Wang G, et al. Development and external validation of a prognostic multivariable model on admission for hospitalized patients with COVID-19. medRxiv 2020. Available at: medrxiv.org/content/medrxiv/early/2020/04//07/2020.03.28.20045997.full.pdf. Accessed February 19, 2024.
[28] Wang K, Zuo P, Liu Y, Zhang M, Zhao X, Xie S, et al. Clinical and laboratory predictors of in-hospital mortality in patients with coronavirus disease-2019: a cohort study in wuhan, China. Clin Infect Dis 2020;71:2079−88.
[29] Lim WS, van der Eerden MM, Laing R, Boersma WG, Karalus N, Town GI, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. Thorax 2003;58(5):377−82.
[30] Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Annals of internal medicine. Annals of internal medicine 1999;130:515−24.
[31] Team Rc. R: a language and environment for statistical computing. Available at:. Vienna, Austria: R Foundation for Statistical Computing; 2022. https://www.r-project.org. Accessed December 19, 2023.
[32] Debray TP, Vergouwe Y, Koffijberg H, Nieboer D, Steyerberg EW, Moons KG. A new framework to enhance the interpretation of external validation studies of clinical prediction models. J Clin Epidemiol 2015;68:279−89.
[33] Donders AR, van der Heijden GJ, Stijnen T, Moons KG. Review: a gentle introduction to imputation of missing values. J Clin Epidemiol 2006;59:1087−91.
[34] Carpenter J, Kenward M. Multiple imputation and its application 1st edition. Hoboken, NJ: John Wiley & Sons, Ltd; 2012.
[35] Rubin DB. Multiple imputation for nonresponse in surveys. Hoboken, NJ: John Wiley & Sons, Inc.; 1987.
[36] Von Hippel PT. 4. Regression with missing Ys: an improved strategy for analyzing multiply imputed data. Socio Methodol 2007;37(1):83−117.
[37] Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010;21:128−38.
[38] Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW, Topic Group 'Evaluating Diagnostic Tests and Prediction Models' of the STRATOS Initiative. Calibration: the Achilles heel of predictive analytics. BMC Med 2019;17(1):230.
[39] Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ 2009;338:b606.
[40] Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. Diagn Progn Res 2019;3:18.
[41] Sauerbrei W, Perperoglou A, Schmid M, Abrahamowicz M, Becher H, Binder H, et al. State of the art in selection of variables and functional forms in multivariable analysis-outstanding issues. Diagn Progn Res 2020;4:3.
[42] Leeuwenberg AM, van Smeden M, Langendijk JA, van der Schaaf A, Mauer ME, Moons KGM, et al. Performance of binary prediction models in high-correlation low-dimensional settings: a comparison of methods. Diagn Progn Res 2022;6(1):1.
[43] Kundu S, Mazumdar M, Ferket B. Impact of correlation of predictors on discrimination of risk models in development and external populations. BMC Med Res Methodol 2017;17:63.
[44] Buttia C, Llanaj E, Raeisi-Dehkordi H, Kastrati L, Amiri M, Mecani R, et al. Prognostic models in COVID-19 infection that predict severity: a systematic review. Eur J Epidemiol 2023;38(4):355−72.

[45] Riley JM, Moeller PJ, Crawford AG, Schaefer JW, Cheney-Peters DR, Venkataraman CM, et al. External validation of the COVID-19 4C mortality score in an urban United States cohort. Am J Med Sci 2022;364(4):409−13.

[46] Jones A, Pitre T, Junek M, Kapralik J, Patel R, Feng E, et al. External validation of the 4C mortality score among COVID-19 patients admitted to hospital in Ontario, Canada: a retrospective study. Sci Rep 2021;11(1):18638.

[47] Hassan S, Ramspek CL, Ferrari B, van Diepen M, Rossio R, Knevel R, et al. External validation of risk scores to predict in-hospital mortality in patients hospitalized due to coronavirus disease 2019. Eur J Intern Med 2022;102:63−71.

[48] Meijs DAM, van Kuijk SMJ, Wynants L, Stessel B, Mehagnoul-Schipper J, Hana A, et al. Predicting COVID-19 prognosis in the ICU remained challenging: external validation in a multinational regional cohort. J Clin Epidemiol 2022;152:257−68.

[49] Covino M, De Matteis G, Burzo ML, Russo A, Forte E, Carnicelli A, et al. Predicting in-hospital mortality in COVID-19 older patients with specifically developed scores. J Am Geriatr Soc 2021;69(1):37−43.

[50] Sperrin M, Riley RD, Collins GS, Martin GP. Targeted validation: validating clinical prediction models in their intended population and setting. Diagn Progn Res 2022;6(1):24.

[51] Jenkins DA, Martin GP, Sperrin M, Riley RD, Debray TPA, Collins GS, et al. Continual updating and monitoring of clinical prediction models: time for dynamic prediction systems? Diagn Progn Res 2021;5(1):1.

[52] Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet 2012;380(9836):37−43.

[53] Hewitt J, Carter B, Vilches-Moraga A, Quinn TJ, Braude P, Verduri A, et al. The effect of frailty on survival in patients with COVID-19 (COPE): a multicentre, European, observational cohort study. Lancet Public Health 2020;5(8):e444−51.

[54] Hoogendijk EO, Afilalo J, Ensrud KE, Kowal P, Onder G, Fried LP. Frailty: implications for clinical practice and public health. Lancet 2019;394:1365−75.