RESEARCH ARTICLE

Accuracy of artificial intelligence-based mortality prediction models in COVID-19 patients: A systematic review and meta-analysis

Xi Li[1, †], Changyang Chen[1, †], Hongpiao Ma[1, †], Peng Luo[1], Junhua Wang[1], Jia Chen[2, *], Jiangping Zhang[1, 2, *]

[1]School of Public Health, The Key Laboratory of Environmental Pollution Monitoring and Disease Control, Ministry of Education, Guizhou Medical University, Guiyang, Guizhou, China. [2]Guiyang Public Health Clinical Center, Guiyang, Guizhou, China.

Received: January 6, 2026; **accepted:** May 6, 2026.

Accurate projection of the results of COVID-19 patients could optimize resource allocation and help formulate appropriate clinical management plans for patients, thereby reducing mortality and improving the overall health of the population. There are currently many modeling methods using artificial intelligence (AI), making it hard to evaluate the advantages and disadvantages of each model. This research undertook a meta-analysis of articles utilizing AI models to forecast the COVID-19 death rate to assess the predictive accuracy of such models by conducting literature search through the databases of Embase, PubMed, and Web of Science from December 1, 2019 to January 20, 2025 using "COVID-19" and "prediction model" as keywords. Meanwhile, sensitivity, specificity, accuracy, area under the curve (AUC), and other indicators were manually extracted. QUADAS-2 was used for quality evaluation, while Stata 16 was employed for data analysis. Ultimately, this study included 34 studies including 36 predictive models. The results indicated a sensitivity value of 0.86 (95%CI: 0.80, 0.91) and a specificity value of 0.89 (95%CI: 0.84, 0.93), demonstrating excellent discriminative ability with AUC of 0.94. The current evidence indicated that AI models could predict COVID-19 patients' mortality with acceptable performance. However, to improve the models' ability to generalize, more external validation works are necessary.

Keywords: COVID-19; mortality rate; predictive model; artificial intelligence; systematic review.

[*]**Corresponding authors:** Jia Chen and Jiangping Zhang, Guiyang Public Health Clinical Center, Guiyang, Guizhou 550004, China. Email: 2644152962@qq.com (Chen J), zhangjiangp23@163.com (Zhang J).

[†]These authors contributed equally to this work.

Introduction

Ever since the first confirmed patient of COVID-19 pneumonia was officially reported in Wuhan, Hubei, China on December 31, 2019, COVID-19 as a severe infectious disease had led to over 770 million infections and nearly 6 million deaths by the end of December 2023 worldwide [1]. The fast spread of COVID-19 worldwide had caused a serious shortage of medical and health resources. In many countries and regions, the supply of hospital beds, equipment, and medicines was tight and unable to meet the increasing demands of patients [2]. Meanwhile, healthcare providers who worked with long shifts and intense workloads were physically and

mentally exhausted and faced huge pressure and challenges [3]. This situation has clearly become a grievous global health problem.

Artificial intelligence (AI) is a technology that simulates human intelligence through computer algorithms and programs. Compared with traditional statistical methods, AI can conduct in-depth datamining and data analysis on large and more complex datasets to discover potential patterns and trends [4]. Rapid and accurate prediction of mortality rates among COVID-19 cases is crucial for optimizing healthcare resources and enabling individualized patient management. Currently, scientists have started employing various computational algorithms and statistical models to develop predictive tools aimed at assessing the mortality risk of COVID-19 patients, thereby achieving the goal of reducing the medical burden and supporting the implementation of effective treatment strategies [5]. However, many current studies have only used one or several artificial intelligence models, making it difficult to assess the strengths and weaknesses of each approach [6]. In addition, due to the differences in research regions, sample sources, data types, and modeling methods, predictive models often have a high risk of bias and lack certain universality. Furthermore, most current meta-analyses focus on the association between specific biochemical markers and mortality in COVID-19 patients or investigate mortality risks in certain groups [7, 8]. Few studies have compared the AI modeling methods used in predictions.

This research conducted a comprehensive meta-analysis of studies that predicted the mortality rate of COVID-19 using different artificial intelligence methods and systematically assessed the quality of the research and identified the common biases to effectively address the issue of low-quality models misleading prevention and control decisions due to the lack of methodological rigor in the early stages of the outbreak caused by time constraints. The results of this study provided valuable insights for predicting the risk of patient disease mortality and helped achieve the rational allocation of health and medical resources.

Materials and methods

Literature search strategy

A systematic literature search was conducted in this research across three major scientific publication databases including PubMed (https://pubmed.ncbi.nlm.nih.gov), Embase (https://www.embase.com), and Web of Science (https://www.webofscience.com) to identify studies on COVID-19 predictive models published from December 1, 2019, to January 20, 2025. An expanded search using medical subject headings (MeSH) (https://www.nlm.nih.gov/mesh/meshhome.html) was performed in the database. Relevant MeSH terms for COVID-19 were first retrieved and combined with the Boolean operator "OR" followed by the retrieval and "OR" combination of MeSH terms related to predictive models. These two sets of expanded MeSH terms were then integrated with the Boolean operator "AND" to conduct a title and abstract search. All search results from the three databases were exported and uniformly imported into NoteExpress reference management software (https://www.inoteexpress.com/aegean/) for subsequent literature processing. The literature inclusion and exclusion criteria included that studies were eligible for inclusion if they were peer-reviewed publications in English, contained sufficient data to directly or indirectly reconstruct the confusion matrix for mortality prediction, reported the total COVID-19 patient count, reported sufficient predictive model performance metrics, and clearly reported data sources. Studies were excluded if they were not related to the topic of this study, represented duplicate publications, were review papers, secondary analyses, conference papers, animal studies, patents, non-peer-reviewed book chapters, or the full text unavailable.

Data selection and extraction

Two independent reviewers performed the screening study and data extraction based on predefined criteria. Inter-rater agreement was evaluated with Cohen's kappa during both phases. Each model was evaluated independently for inclusion when validated in distinct cohorts, while models sharing overlapping training or validation samples were excluded to prevent data duplication bias. Conflicts between the two reviewers were settled through discussion or by arbitration from an independent assessor. The types of data being extracted included author(s), publication year, region in the modeling dataset, study design, data source, dataset type of predictive models, AI techniques, and data type. A meta-analysis was then performed by using the processed data, which included four contingency table values of true positives [TP], false positives [FP], true negatives [TN], false negatives [FN] sourced from each study, either as reported or as calculated from available data.

Quality assessment

The study quality was evaluated by two independent researchers using the quality assessment of diagnostic accuracy studies 2 (QUADAS-2) criteria (https://www.bristol.ac.uk/population-health-sciences/projects/quadas/quadas-2) including the risk of bias in the areas of patient selection, indicator tests, reference criteria, procedure, time and the applicability concerns in three areas of patient selection, indicator tests, reference criteria. Domains with insufficient reporting were conservatively rated as high risk of bias.

Data analysis

STATA version 16.0 (StataCorp LLC, College Station, TX, USA) incorporating the MIDAS module was employed for statistical analysis of this research. Pooled estimates of sensitivity, specificity, and the area under the curve (AUC) values were computed to access diagnostic accuracy. Model heterogeneity was evaluated through forest plots and the I^2 statistic. In cases of substantial heterogeneity with $I^2 > 50\%$, a bivariate random-effects model was employed to incorporate between-study covariance in sensitivity and specificity. The hierarchical summary receiver operating characteristic (HSROC) model complemented AUC synthesis where appropriate. Subgroup analyses and meta-regression were adopted to explore heterogeneity. Furthermore, sensitivity analysis was conducted to identify studies exerting disproportionate influence on the diagnostic odds ratio (DOR) effect size. Deeks' funnel plot asymmetry test was applied to evaluate publication bias.

Results

Literature sifting process

The literature screening process resulted that a Boolean search with predefined keywords initially yielded 9,931 records. After removing 4,067 duplicates and 796 ineligible documents including meta-analyses, reviews, book chapters, and animal studies, additional 4,892 irrelevant records were eliminated based on the title and abstract screening, leaving 176 articles for full-text eligibility assessment. The full-text review process further excluded 142 articles including 110 non-target traditional models, 4 unavailable full texts, and 28 with incomplete data. Ultimately, 34 studies met all inclusion criteria and were included in this meta-analysis (Figure 1).

Characteristics of included studies

Based on the World Health Organization's (WHO) regional classification criteria that if more than 90% of the data in a study originates from countries within a specific region, the study is assigned to that region, 19 of the 36 predictive models (52.78%) analyzed were derived from studies conducted primarily within Asian geopolitical entities with China, Iran, and Saudi Arabia being the predominant contributors. Multi-continental collaborations were categorized as "Transregional" and analyzed separately. 33 out of 36 models (91.67%) were cohort studies, while 3 models used cross-sectional designs. 29 models (80.56%) used

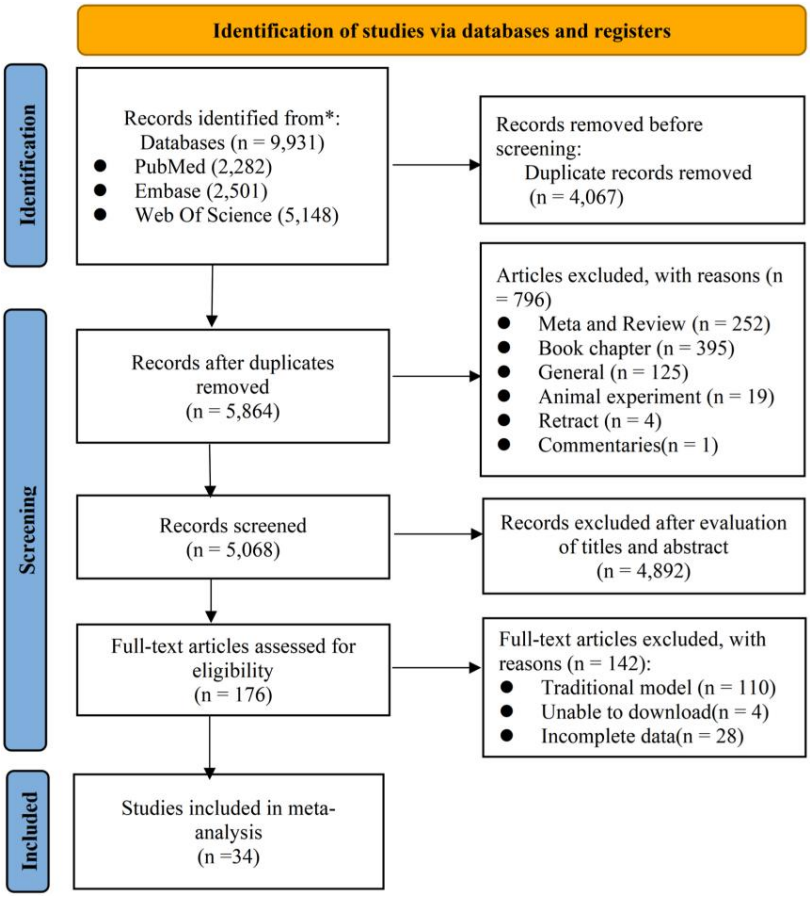

Figure 1. Study selection flowchart.

Table 1. Characteristics of included models (n = 36).

Characteristics	Total	Values	Frequency	%
Region	36	Asia	19	52.78
		Non-Asia	17	47.22
Study design	36	Cross-sectional study	3	8.33
		Cohort study	33	91.67
Data source	36	Multicenter	21	58.33
		Monocentric	15	41.67
Dataset	36	Internal validation	29	80.56
		External validation	7	19.44
AI techniques	36	Machine learning	31	86.11
		Deep learning	5	13.89
Data type	36	Clinical data	24	66.67
		Combing clinical and imaging data	12	33.33

internal validation, while 7 had external validation. Methodologically, 31 models (86.11%) used traditional machine learning like random forest or support vector machines, while 5 used deep learning. 24 models (66.67%) used only clinical data, while 12 combined clinical and imaging data (Table 1).

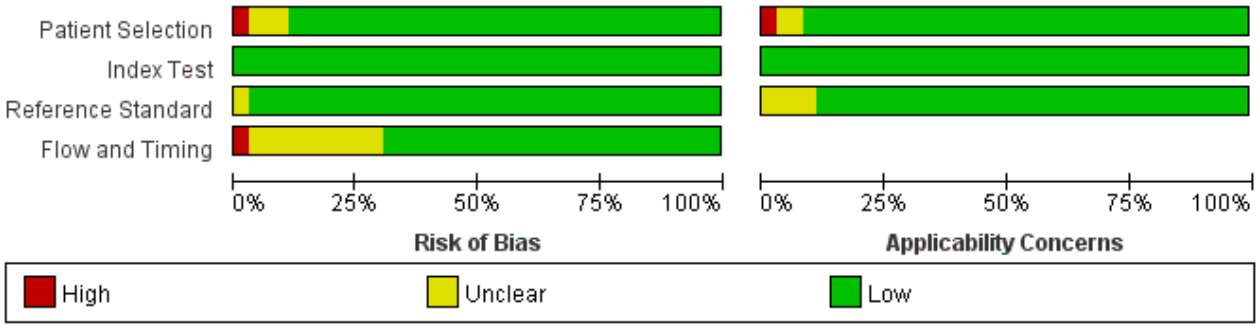

Figure 2. Summary of methodological evaluation through QUADAS-2.

Quality assessment

The results of methodological quality evaluation showed that, for index test interpretation, 35 models (97%) were at low risk and only 1 model (3%) at high risk. Similar results were found for reference standard application. However, despite the predominantly low risk ratings, deeper examination revealed critical limitations. Inconsistent patient selection criteria across studies might introduce spectrum bias, while insufficient external validation cohorts potentially overestimated generalizability. Further, patient selection domains still required improvement to enhance reliability. The QUADAS-2 applicability assessment demonstrated high conformity in patient selection (92%), index test alignment (100%), and reference standard appropriateness (89%). Most studies demonstrated appropriate test accuracy and patient selection, though some reference standard issues persisted (Figure 2).

Pooled diagnostic accuracy

The AI models demonstrated excellent predictive performance for COVID-19 patient mortality with a pooled AUC of 0.94. The meta-analysis indicated a combined sensitivity of 0.86 (95% CI: 0.80 - 0.91) and specificity of 0.89 (95% CI: 0.84 - 0.93). The positive likelihood ratio (LR+) was 8.0 (95% CI: 5.4 - 11.6), indicating that surviving patients had approximately 8-fold higher odds of being classified as negative by AI models compared to deceased patients. The negative likelihood ratio (LR-) was 0.15 (95% CI: 0.10 - 0.23), and the diagnostic odds ratio (DOR) was 52 (95% CI: 29 - 93), collectively demonstrating strong discriminative ability. However, substantial heterogeneity was observed across studies with I^2 values exceeding 98% for both sensitivity and specificity ($P < 0.05$), suggesting significant difference between-study variation in accuracy measures (Table 2).

Table 2. Predictive performance of artificial intelligence in COVID-19 detection.

Metrics	Performance (95%CI)
AUC	0.94
Sensitivity	0.86 (0.80, 0.91)
Specificity	0.89 (0.84, 0.93)
LR+	8.0 (5.4, 11.6)
LR-	0.15 (0.10, 0.23)
DOR	52 (29, 93)

Subgroup and meta-regression analysis

To identify covariates affecting COVID-19 mortality prediction models, subgroup analyses and meta-regression for regions, study designs, data sources, data-set types, AI techniques, and data types were performed. The results demonstrated that significant heterogeneity was associated with geographic region for sensitivity ($P < 0.01$) with Asia studies showing higher pooled sensitivity (0.91, 95% CI: 0.85 - 0.96) compared to non-Asia studies (0.80, 95% CI: 0.70 - 0.91). No significant associations were found with study design, data source, dataset type, data type, or AI technique. For specificity, significant heterogeneity was observed across regions ($P < 0.05$), data sources ($P < 0.01$), dataset types ($P <$

Table 3. Sensitivity and specificity subgroup analyses.

Categories	Models (n)	Sensitivity (95% CI)	*P* value	Specificity (95% CI)	*P* value
Region					
Asia	19	0.91 (0.85 - 0.96)	< 0.01	0.89 (0.83 - 0.95)	< 0.05
Non-Asia	17	0.80 (0.70 - 0.91)		0.90 (0.84 - 0.95)	
Study framework					
Cross-sectional study	3	0.80 (0.54 - 1.00)	> 0.05	0.90 (0.85 - 0.94)	> 0.05
Cohort study	33	0.87 (0.81 - 0.93)		0.84 (0.65 - 1.00)	
Source of data					
Multi-center	21	0.82 (0.73 - 0.91)	> 0.05	0.90 (0.85 - 0.95)	< 0.01
Monocentric	15	0.91 (0.85 - 0.97)		0.88 (0.81 - 0.95)	
Verification set					
Internal validation	29	0.86 (0.80 - 0.93)	> 0.05	0.91 (0.87 - 0.95)	< 0.05
External validation	7	0.87 (0.75 - 0.99)		0.81 (0.67 - 0.96)	
AI techniques					
Machine learning	31	0.88 (0.82 - 0.93)	> 0.05	0.88 (0.84 - 0.93)	> 0.05
Deep learning	5	0.77 (0.54 - 0.99)		0.93 (0.86 - 1.00)	
Data type					
Clinical data	24	0.83 (0.75 - 0.91)	> 0.05	0.90 (0.85 - 0.95)	< 0.01
Combing clinical and imaging data	12	0.92 (0.85 - 0.98)		0.87 (0.79 - 0.95)	

0.01), and data types ($P < 0.01$). Specifically, multicenter studies demonstrated higher specificity (0.90, 95% CI: 0.85 - 0.95) than monocentric studies (0.88, 95% CI: 0.81 - 0.95), internal validation yielded higher specificity (0.91, 95% CI: 0.87 - 0.95) than external validation (0.81, 95% CI: 0.67 - 0.96), and clinical data alone achieved higher specificity (0.90, 95% CI: 0.85 - 0.95) than combined clinical and imaging data (0.87, 95% CI: 0.79 - 0.95). Study design and AI technique showed no significant effects on specificity (Table 3).

Sensitivity analysis

To evaluate the robustness of the pooled results, a sensitivity analysis was conducted by sequentially omitting each individual study and recalculating the pooled diagnostic odds ratio. The analysis demonstrated that no single study exerted excessive influence on the overall effect size as all point estimates remained within the 95% confidence interval of the combined estimate (Figure 3). The results indicated that meta-analysis results of this study were statistically robust and not disproportionately driven by any particular study.

Publication bias

Deeks' regression analysis was employed to assess publication bias across all 34 included studies comprising a total of 36 predictive models meeting the screening criteria by examining the asymmetry of funnel plots. A *P* value of less than 0.1 was considered indicative of publication bias (Figure 4). The symmetrical point distribution around the true effect size in funnel plots indicated minimal publication bias [9]. There was no evidence of publication bias being found in this study ($P > 0.05$).

Discussion

To ensure the equitable allocation of limited healthcare resources and to provide targeted, personalized care aimed at improving patient outcomes, a comprehensive evaluation of the predictive performance of artificial intelligence models in forecasting the mortality risk associated with COVID-19 infection was conducted. The results of meta-analysis demonstrated that AI models exhibited high diagnostic accuracy in recognizing patients at the

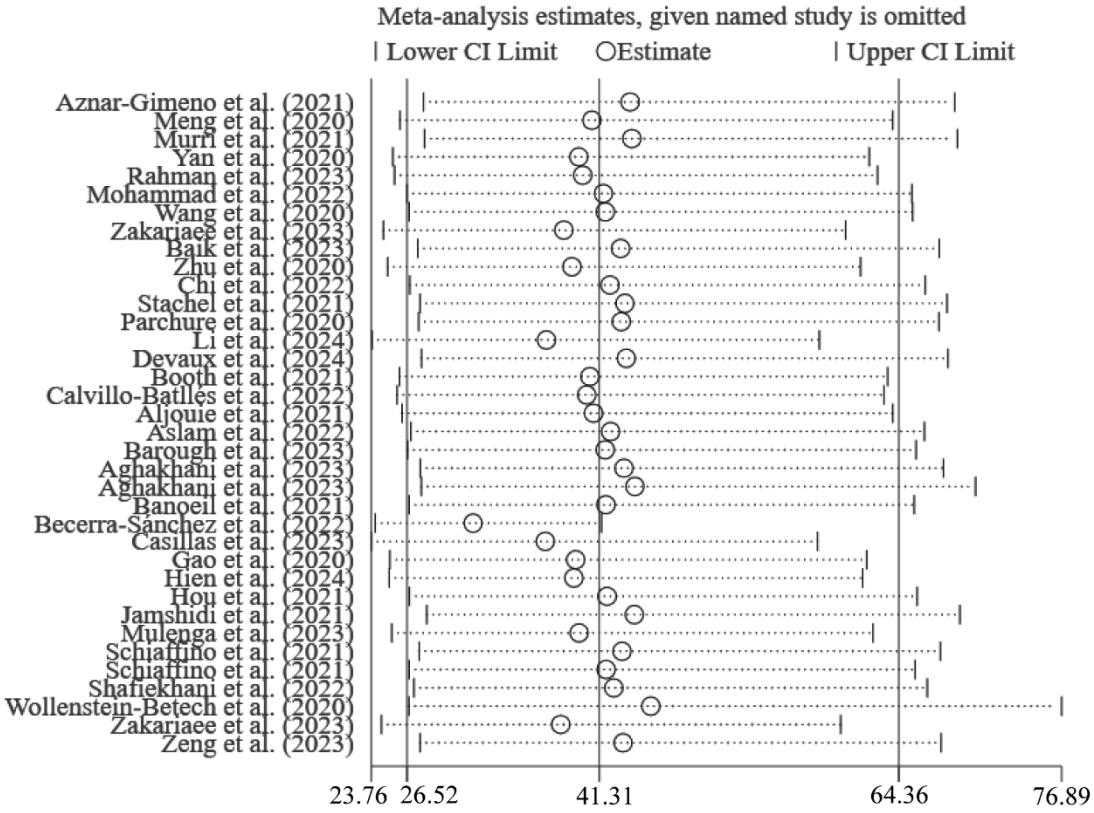

Figure 3. Sensitivity analysis.

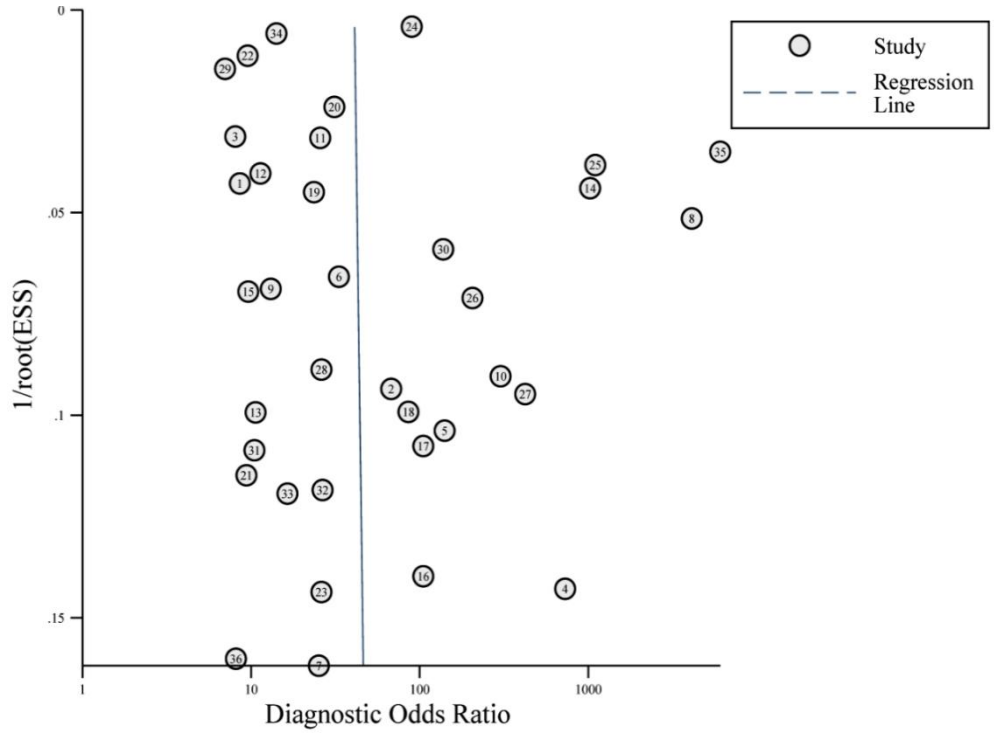

Figure 4. Deek's funnel plot results.

death threats of COVID-19 and suggested that AI-driven prediction models had promising performance in mortality risk stratification for COVID-19. This comprehensive evaluation was based on literature with its conclusions being more clinically practical. However, due to the study subjects differed in age, gender, and basic health status, and the modeling methods used varied, the quality of the literature was highly inconsistent, which introduced some instability into the combined results. Therefore, it is essential to further verify through prospective investigations. The results of subgroup analysis revealed that AI technologies achieved greater sensitivity and specificity when forecasting COVID-19 mortality rates in the Asian population compared to non-Asian groups, which might be attributed to the fact that Asia was the initial epicenter of SARS-CoV-2 with widespread transmission, prompting extensive local research. Further, disparities in health resources including medical facilities and equipment between the two regions might result in inconsistencies in the quality of data incorporated into the model. In addition, Asia previously had rich clinical experience and mature diagnostic techniques in managing the outbreak of infectious diseases such as SARS, which led to higher quality clinical data. Therefore, it had resulted in a wealth of clinical data available for model training. Consequently, the model demonstrated higher sensitivity and specificity in predicting mortality rates among the Asian population [10]. Unlike other studies, this research observed that the machine learning model demonstrated higher sensitivity than the deep learning model [11], which might be because machine learning models usually had a simpler structure and required less training data. These models were not tending to over-fitting, particularly when dealing with limited datasets or a small number of features [12]. Although deep learning models have proved strong capabilities in handling large-scale and complex datasets, they usually rely on a mass of high-quality training data. When the available data is limited or of poor quality, these models are prone to over-fitting, which may result in a reduction in the sensitivity of identifying positive cases during the prediction process. Moreover, disease symptoms can vary between regions and patient groups, i.e. the clinical data included in different studies are heterogeneous [13]. This inconsistency ends up decreasing how well the model works. This research also indicated that multi-center studies could help enhance the universality of research conclusions [14], which was likely because multi-center studies incorporated data from multiple regions or hospitals, enabling the collection of diverse clinical manifestations across different populations. The increased diversity and representativeness of the data allowed the model to learn a broader range of predictive features related to mortality during training and led to more accurate classification and prediction of fatal outcomes, resulting in higher specificity in multi-center studies compared to single-center studies [15]. This study demonstrated that the model's specificity was higher in the interior test set compared with the external one. This discrepancy likely came from internal test samples and training sets typically being drawn from the same population, making the internal datasets more adaptable [16]. In contrast, the external test set often comprises data collected from different institutions or regions, leading to variations in demographic characteristics such as age and gender when compared to the training set [17]. Consequently, the model' s performance on external data tended to be relatively less effective. These discrepancies could lead to reduced model performance when applied to external data. When AI models were used to predict mortality rates by integrating clinical data with medical imaging data, their specificity tended to be lower than the models that relied solely on clinical data [18], which might be attributed to image quality varies, affected by different imaging equipment or the technical skills of scanning medical personnel. Inconsistencies in image quality might negatively affect diagnostic accuracy and model performance. Models that entirely based on clinical data utilize structured and standardized variables such as age, gender, and laboratory test

data, which may be more helpful in predicting the happening of the risk of death [19].

This research still had limitations, which included that the research subjects were spread unevenly across regions. Sample sizes were inconsistent. The model's predictive universality hadn't been fully tested. In addition, the count of studies included in different groups in the subgroup analysis varied significantly, which limited the reliability of heterogeneity assessment among subgroups. Further, the study did not fully incorporate the assessment of the impact of different COVID-19 variants on death rate prediction, which might lead to a decline in the predictive efficacy of the model when new epidemic strains emerged. Furthermore, if the included studies had high heterogeneity in terms of design, data, methods, *etc*., it might affect the reliability of the Deeks' funnel plot conclusion, thereby leading to the conclusion of no publication bias. Although the constraints, this research offered useful perspectives to support the precise forecasting of death risk among individuals with COVID-19. The results could underpin the reasonable deployment of limited healthcare resources and aid in advancing personalized strategies for diagnosis and treatment.

Conclusion

This study integrated 34 studies and 36 models to conduct a meta-analysis evaluating the predictive accuracy of artificial intelligence technologies for COVID-19 mortality. The key indicators included sensitivity, specificity, combined AUC, likelihood ratio, and diagnostic odds ratio. The findings revealed that AI models, which included ML and DL techniques, held potential utility for predicting the outcomes of COVID-19 patients. However, due to the existence of uncertainties, more high-quality research results needed to be included for verification. Meanwhile, the application should complement clinical expertise and contextual considerations in practice to help make more reliable decisions. If the primary objective was to enhance sensitivity, machine learning methods were generally more advantageous. However, if the goal was for high specificity, structured data should be given priority. Nevertheless, irrespective of the chosen methodology, it was imperative to incorporate case data from diverse regions and populations, which would help mitigate biases arising from restricted datasets and ensure that the model remained both explainable and broadly applicable.

Acknowledgement

This work was supported by the High-Level Talent Research Start-Up Fund Project of Guizhou Medical University (Grant No. Xiao Bo He J [2025] 003).

References

1. Nigatu BZ, Dessie NT. 2025. Prevalence of comorbidities and their association with disease severity and mortality in COVID-19 patients: A systematic review and meta-analysis. J Multimorb Comorb. 15:26335565251371256.
2. Zheng Q, Zeng Z, Tang X, Ma L. 2024. Impact of an ICU bed capacity optimization method on the average length of stay and average cost of hospitalization following implementation of China's open policy with respect to COVID-19: A difference-in-differences analysis based on information management system data from a tertiary hospital in southwest China. BMJ Open. 14(4):e078069.
3. Misan N, Wilf-Miron R, Saban M. 2024. Comparing emergency nursing measures before and during COVID-19: A retrospective study of assessment, triage, and workflow. SAGE Open Nurs. 10:23779608241274766.
4. Shillan D, Sterne JAC, Champneys A, Gibbison B. 2019. Use of machine learning to analyze routinely collected intensive care unit data: A systematic review. Crit Care. 23(1):284.
5. Hamidi HH, Moradi M. 2025. Analysis of disease severity and mortality prediction using machine learning during COVID-19. Acta Psychol (Amst). 258:105136.
6. Li J, Li X, Hutchinson J, Asad M, Liu Y, Wang Y, *et al*. 2022. An ensemble prediction model for COVID-19 mortality risk. Biol Methods Protoc. 7(1):bpac029.
7. Liang C, Zhang X, Zhu Q, Guo J, Wu Y, Wang Z, *et al*. 2025. Lactate/albumin ratio predicts mortality in critically ill COVID-19 patients: A retrospective machine learning study. Sci Rep. 15(1):36422.
8. Lages Dos Santos A, Oliveira MCL, Colosimo EA, Mak RH, Pinhati CC, Gallante SC, *et al*. 2025. Comparative performance of

twelve machine learning models in predicting COVID-19 mortality risk in children: A population-based retrospective cohort study in Brazil. Peer J Comput Sci. 11:e2916.

9. Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, *et al*. 2011. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomized controlled trials. BMJ. 343:d4002.
10. Tang JW, Caniza MA, Dinn M, Dwyer DE, Heraud JM, Jennings LC, *et al*. 2022. An exploration of the political, social, economic and cultural factors affecting how different global regions initially reacted to the COVID-19 pandemic. Interface Focus. 12(2):20210079.
11. Chavoshi M, Zamani S, Mirshahvalad SA. 2024. Diagnostic performance of deep learning models versus radiologists in COVID-19 pneumonia: A systematic review and meta- analysis. Clin Imaging. 107:110092.
12. Daramola O, Kavu TD, Kotze MJ, Marnewick JL, Sarumi OA, Kabaso B, *et al*. 2025. Predictive modelling and identification of critical variables of mortality risk in COVID-19 patients. Sci Rep. 15(1):2184.
13. Bizuneh FK, Biwota GT, Tsheten T, Bizuneh TK. 2025. Incidence of recovery rate and predictors among hospitalized COVID-19 infected patients in Ethiopia: A systemic review and meta-analysis. BMC Public Health. 25(1):1644.
14. Sun R, Li S, Yang C, Huang G, Tang C, Li W, *et al*. 2025. Association between preoperative cognitive performance and postoperative delirium in older patients: Results from a multicenter, prospective cohort study, and a mendelian randomization study. MedComm. 6(8):e70302.
15. Ma X, Ng M, Xu S, Xu Z, Qiu H, Liu Y, *et al*. 2020. Development and validation of prognosis model of mortality risk in patients with COVID-19. Epidemiol Infect. 148:e168.
16. Yin H, Jiang Y, Xu Z, Huang W, Chen T, Lin G. 2022. Apparent diffusion coefficient-based convolutional neural network model can be better than sole diffusion-weighted magnetic resonance imaging to improve the differentiation of invasive breast cancer from breast ductal carcinoma *in situ*. Front Oncol. 11:805911.
17. Kablan R, Miller HA, Suliman S, Frieboes HB. 2023. Evaluation of stacked ensemble model performance to predict clinical outcomes: A COVID-19 study. Int J Med Inform. 175:105090.
18. Ahmed SR, Befano B, Egemen D, Rodriguez AC, Desai KT, Jeronimo J, *et al*. 2025. Generalizable deep neural networks for image quality classification of cervical images. Sci Rep. 15(1):6312.
19. Rosca D, Krishna V, Chetarajupalli C, Jianu AM, Deak IE, Virzob CRB, *et al*. 2024. Comparative analysis of qSOFA, PRIEST, PAINT, and ISARIC4C scores in predicting severe COVID-19 outcomes among patients aged over 75 years. Diseases. 12(12):304.