BMC Medical Informatics and Decision Making

SYSTEMATIC REVIEW

Open Access

Accuracy of artificial intelligence algorithms in predicting acute respiratory distress syndrome: a systematic review and meta-analysis

Yaxin Xiong[1], Yuan Gao[1], Yucheng Qi[1], Yingfei Zhi[1], Jia Xu[1], Kuo Wang[1], Qiuyue Yang[1], Changsong Wang[1,2], Mingyan Zhao[1,2*] and Xianglin Meng[1,2*]

Abstract

Background Acute respiratory distress syndrome (ARDS) is a serious threat to human life. Hence, early and accurate diagnosis and treatment are crucial for patient survival. This meta-analysis evaluates the accuracy of artificial intelligence in the early diagnosis of ARDS and provides guidance for future research and applications.

Methods A search on PubMed, Embase, Cochrane, Web of Science, CNKI, Wanfang, Chinese Biomedical Literature (CBM), and VIP databases was systematically conducted, from their establishment to November 2023, to obtain eligible studies for the analysis and evaluation of the predictive effect of AI on ARDS. The retrieved literature was screened according to inclusion and exclusion criteria, the quality of the included studies was assessed using QUADAS-2, and the included studies were statistically analyzed.

Results Among the 2, 996 studies, 33 were included in this meta-analysis, which showed that the pooled sensitivity of AI in predicting ARDS was 0.81 (0.76–0.85), the pooled specificity was 0.88 (0.84–0.91), and the area under the receiver operating characteristic curve (AUC) was 0.91 (0.88–0.93). The analyzed studies included 28 models, with a pooled sensitivity of 0.79 (0.76–0.82), a pooled specificity of 0.85 (0.83–0.88), and an AUC of 0.89 (0.86–0.91). In the subgroup analysis, the pooled AUC of the AI models ANN, CNN, LR, RF, SVM, and XGB were 0.86 (0.83–0.89), 0.91 (0.88–0.93), 0.86 (0.83–0.89), and 0.89 (0.86–0.91), 0.90 (0.87–0.92), 0.93 (0.90–0.95), respectively. In an additional subgroup analysis, we evaluated the predictive performance of the AI models trained using different predictors. This meta-analysis was registered in PROSPERO (CRD42023491546).

Conclusion AI has good sensitivity and specificity for predicting ARDS, indicating a high clinical application value. Algorithmic models such as CNN, SVM, and XGB have improved prediction performance. The subgroup analysis revealed that the model trained using images combined with other predictors had the best predictive performance.

Keywords Artificial intelligence, Acute respiratory distress symptoms, Prediction, Meta-analysis

*Correspondence:
Mingyan Zhao
mingyan197@gmail.com
Xianglin Meng
mengzi98@163.com
Full list of author information is available at the end of the article

Background

Acute respiratory distress syndrome (ARDS) is a severe condition, affecting more than 3 million patients worldwide every year. Further, the mortality rate of severe ARDS can reach up to 46.1%, indicating a severe threat to patients' lives [1]. The etiology of ARDS is complex and its disease progression is rapid, thus accurate early prediction, diagnosis, and individualized treatment plans are crucial for improving patient survival rates. However, the early identification of potentially high-risk ARDS requires the timely integration and analysis of basic patient information, disease characteristics, laboratory assays, imaging data, and respiratory mechanical characteristics for predictive analytics [2]. Developing methods or tools to accurately predict ARDS early is crucial for effectively treating patients with ARDS.

With the rapid development of artificial intelligence (AI) technology, its application in medicine has attracted widespread attention. AI technologies, such as machine learning and deep learning, show great potential for analyzing medical images, bioinformatics, and clinical decision support. Their ability to process data and recognize patterns makes them powerful tools for processing large-scale clinical data. These technologies have been applied to the early diagnosis of diseases, patient management, and prognosis assessment in clinical medicine [3, 4].

Patients with ARDS often exhibit complex and highly heterogeneous conditions, accompanied by an abundance of clinical, biomarker, and imaging data. This establishes a solid informational foundation for AI research in the field of ARDS. Studies have demonstrated that AI technology can effectively be utilized for early prediction of ARDS, by conducting in-depth analysis of clinical data, imaging materials, and monitoring information. This aids doctors in achieving more rapid and precise diagnoses and predictions, thereby facilitating timely interventions and improving patient outcomes [5]. However, no meta-analyses of the efficacy of AI in the prediction and diagnosis of ARDS have been conducted. A search on research databases shows that studies on AI in ARDS have been proliferating, especially since 2021. Hence, this study conducted a meta-analysis of AI prediction for ARDS to evaluate the accuracy of AI in predicting ARDS, providing useful guidance for future research and applications.

Methods

This meta-analysis was conducted according to the PRISMA guidelines [6]. This meta-analysis was registered in PROSPERO (CRD42023491546).

Literature search strategy and screening

We searched PubMed, Embase, Cochrane, Web of Science, CNKI, Wanfang, Chinese Biomedical Literature (CBM), and VIP databases from their earliest available records up to November 2023.The search terms were "acute respiratory distress syndrome," "artificial intelligence," "computer intelligence," "machine learning," "computer reasoning," "deep learning," and "random forest." (The specific search strategy is described in Supplementary Material 3)

Inclusion and exclusion criteria

Articles were screened against the inclusion criteria by two independent researchers, first by title or abstract. If this step did not provide clear results, the entire text was examined to determine whether the article satisfied the inclusion criteria. Disagreements between researchers were resolved by consulting a third researcher.

The studies included in this paper were retrospective, and satisfied the following inclusion criteria: (1) literature in different languages; (2) studies conducted in patients with ARDS or those likely to develop ARDS; (3) studies that provide direct or indirect data to calculate the true positive (TP), false positive (FP), true negative (TN), and false negative (FN) values of the study to construct a complete four-fold table; (4) studies that present clear descriptions of the AI models and predictors used; (5) studies that clearly illustrate the source of the dataset used.

The exclusion criteria were as follows: (1) studies with incomplete data in the literature and inability to obtain TP, FP, TN, and FN directly or indirectly; (2) reviews, conference reports, letters, and experiments with animals; and (3) Literature duplicating experimental data.

Data extraction and literature quality assessment

Information was extracted and crosschecked independently by two researchers. Disagreements were resolved through discussion or consultation with a third researcher. We extracted the results from the validation or test sets of the study, and when there were no clear grouping in the text, we analyzed them using the total sample size. Data extracted included authors, year of publication, study population, study area, study type, sample size, prevalence, cross-validation method, TP, FP, TN, FN, sensitivity, and specificity. Quality evaluation of the included studies was performed using QUADAS-2 [7].

Statistical analysis

Statistical analysis was performed for the included studies. Sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR), and diagnostic odds ratio (DOR) were summarized using a bivariate mixed-effects model (MIDAS), and a summary receiver operating characteristic (SROC) curve was plotted, with the area under the curve (AUC) value calculated. All results were expressed with 95% confidence intervals. Calculate the Spearman correlation coefficient to detect heterogeneity caused by threshold effects. The magnitude of heterogeneity was assessed using the I^2 statistic. Heterogeneity was considered high if I^2 was $>50\%$, and sources of heterogeneity were explored through meta-regression. The stability of the results of diagnostic studies is tested through sensitivity analysis. The Deeks test was employed to assess the publication bias of the included studies, which were deemed to have publication bias when $P<0.05$.

Results

Literature search results and characteristics of the included studies

A total of 2, 996 studies were searched through databases. After removing duplicates, the titles and abstracts of 1919 studies were reviewed, and the full text of 126 studies was screened. Ultimately, 33 studies were included in the pooled analysis [2, 8–39]. The specific literature screening process is shown in Fig. 1. The characteristics of the included studies are summarized in Table 1. The selected studies were conducted in eight countries and regions, with 21 single-center

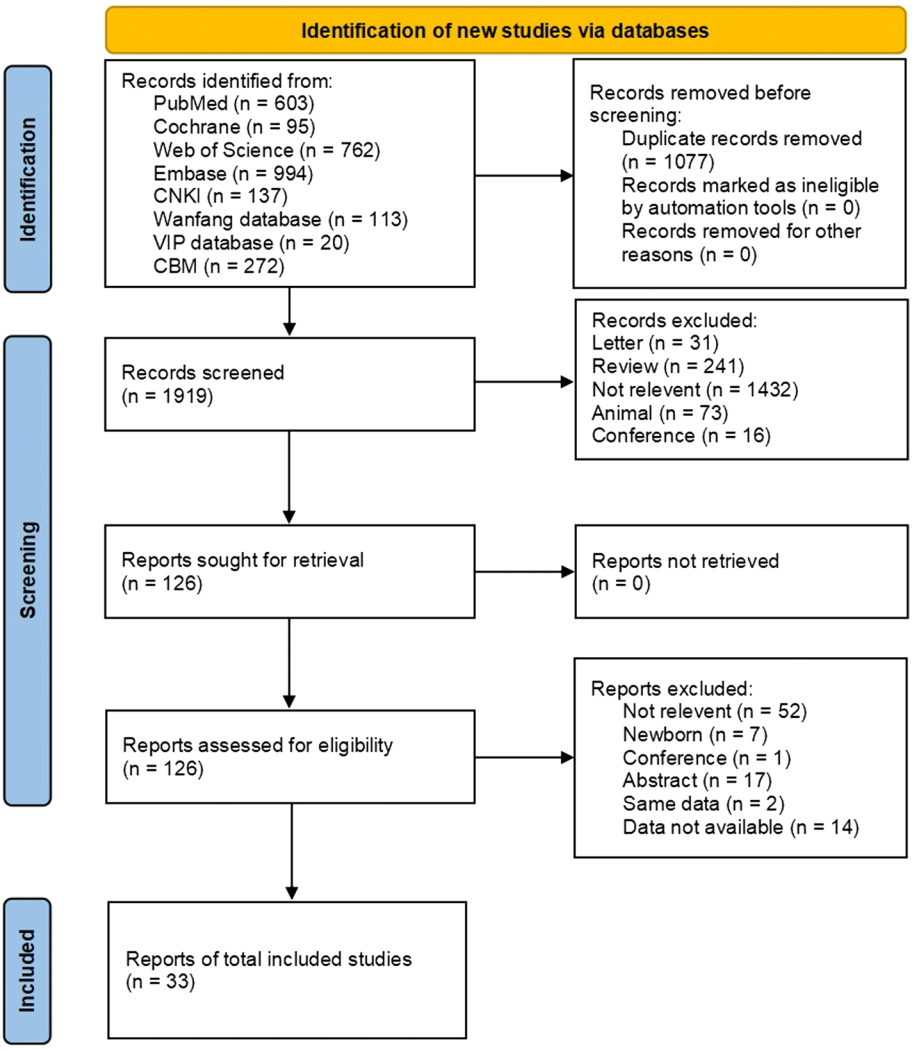

Fig. 1 Literature screening flow

Table 1 Characteristics of the included studies

Author, Year	Country	Selected time	Study type	Single/Multi-center	Database	Model	Patients (n)	ARDS Incidence (%)	Method of validation
Kai Chih Pai 2022 [19]	Taiwan of China	2018.10-2019.12	Retrospective	Single center	Taichung Veterans General Hospital	XGB,RF,LR,XGB+RF+LR,CNN,XGB+CNN,RF+CNN,LR+CNN,XGB+RF+LR+CNN	1577	24.29	five-fold cross-validation
Negar Farzaneh 2023 [22]	America	2017	Retrospective	Multi-center	CheXpert, MIMIC-CXR,University of Michigan	CNN	115	26.09	/
Yang Zhou 2023 [23]	China	2022.4-2022.6	Retrospective	Single center	Shanghai Renji Hospital	CNN,XGB,XGB+CNN	31	12.9	five-fold cross-validation
Michael W Sjoding 2021 [24]	America	2016.1.1-2017.6 30/2017.7.1-12.1/2015.11-2017.12.31	Retrospective	Multi-center	CheXpert, MIMIC-CXR, University of Michigan, Hospital of the University of Pennsylvania	CNN	1560	28.08	/
Hyun Woo Lee 2023 [25]	Korea	2020.2-2020.10	Retrospective	Multi-center	Boramae Medical Center, Korean Imaging Cohort of COVID-19(KICC-19)	faster-RCNN,LR1,LR2	467	9.85	/
Yuntian Chen 2020 [31]	China	2020.1.20-2020.3.31	Retrospective	Single center	Chengdu Public Health Center	LR	105	14.29	/
Mehak Arora 2023 [37]	America	2015.8-2019.5	Retrospective	Multi-center	MIMIC-CXR, Emory University Hospital	CNN	1639	61.56	ten-fold cross-validation
Suganya D 2023 [30]	India	/	Retrospective	Multi-center	publicly available datasets (https://github.com/mr7495/COVID-CTset),be created by fusing multiple images from different source	Mask R-CNN	12452	6.87	five-fold cross-validation
Nesrine Zaglam 2014 [10]	Canada	/	Retrospective	Single center	a previous clinical database	SVM	90	58.89	/
Sebastian Röhrich 2021 [15]	Austria	/	Retrospective	Single center	Med Univ Vienna	GBT	123	32.52	forty-fold cross-validation
Yang Fei 2018 [20]	China	2015.1-2018.4	Retrospective	Multi-center	Medical School of Nanjing University, Jinling Hospital, Nanjing General Hospital of Nanjing Military Region	ANNs,LR	32	21.88	five-fold cross-validation
Zara Izadi 2022 [26]	America	2020.3.24-2021.5.12	Retrospective	Multi-center	the COVID-19 Global Rheumatology Alliance	KNN,SVM,GLMNET,BAYESGLM,GAM,GBM,NN,multivariate LR	891	3.93	ten-fold cross-validation
Mengran Zhang 2023 [12]	China	2017.1-2022 .8	Retrospective	Single center	Xuanwu Hospital of Capital Medical University	SVM,EDTs,BC,LR	92	18.48	five-fold cross-validation
Pengcheng Yang 2020 [28]	China	2001-2012	Retrospective	Single center	MIMIC-III	SLP-FNN,L2-LR,AdaBoost,XGBoost	47352	76.83	ten-fold cross-validation
Menglian Zhou 2019 [16]	America	/	Retrospective	Single center	University of Michigan	ML+PCA+LDA	85	48.24	four-fold cross-validation

Table 1 (continued)

Author, Year	Country	Selected time	Study type	Single/Multi-center	Database	Model	Patients (n)	ARDS Incidence (%)	Method of validation
Wan Xu 2021 [33]	China	2020.1.22-2020.4.1	Retrospective	Multi-center	hospitals from 11 regions: NingBo, ZhouShan, HuBei, Lishui, Jiaxin, HangZhou, TaiZhou, DongYang, ShaoXing, WenZhou, HuZhou	DT,LR,RF,SVM,DNN	71	19.72	ten-fold cross-validation
Jipeng Mo 2023 [34]	China	2018.2-2021.2	Retrospective	Multi-center	Changzhou Second People's Hospital, Jiangsu Provincial People's Hospital, Nanjing Military Region General Hospital, Wuxi Fifth People's Hospital	ANN	288	37.15	/
Kang Zou 2023 [35]	China	2013.7-2022.4	Retrospective	Single center	Hospital of Southwest Medical University	ANN,LR	64	38.01	five-fold cross-validation
Wanyue Zhang 2023 [36]	China	2017.8-2022 .8	Retrospective	Single center	Zhongda hospital	LR,RF,SVM,DT,XGB	308	50.97	five-fold cross-validation
Gregory B Rehm 2021 [38]	America	2015-2019	Retrospective	Single center	UC Davis Medical Center	RF,NN,Adaboost ,LR,Naïve Bayes,SVM	100	50	five-fold cross-validation
Majid Afshar 2018 [9]	America	2011.1.1-2016.12.31	Retrospective	Single center	a tertiary academic center	SVM	106	25.47	ten-fold cross-validation
Amir Gandomi 2022 [8]	America	2001-2012	Retrospective	Single center	MIMIC-III	linear SVM	300	23.67	five-fold cross-validation
Zhiwei Yang 2022 [21]	China	2017.4.10-2021.10 .31	Retrospective	Single center	Hebei General Hospital	multivariate LR	516	42.05	/
Sidney Le 2020 [18]	England	2001-2012	Retrospective	Single center	MIMIC-III	XGBoost	9251	3.2	ten-fold cross-validation
Marshall C.E. 2023 [14]	America	2016.10-2018.9	Retrospective	Single center	Emory University Healthcare ICUs	XGB	2078	6.16	/
Anoop Mayampurath 2020 [2]	America	/	Retrospective	Multi-center	Loyola University Medical Center,University of Chicago Medicine	SVM	235	46.81	/
Yu Wang 2023 [11]	China	2021.4-2021.7/2021.8-2022.7/2023.1-2023.3	Retrospective	Single center	Union Hospital of Tongji Medical College Hospital	SLR	150	66.67	ten-fold cross-validation
Xiaoqiang Wang 2023 [29]	China	2019.1.1 - 2020.12.31	Retrospective	Single center	Eastern Hepatobiliary Surgery Hospital,Shanghai,China	LR,LASSO regression	542	5.72	ten-fold cross-validation
Elyas Sabeti 2021 [13]	America	/	Retrospective	Single center	University of Michigan Hospital	SVM,SVM+	9362	7.71	five-fold cross-validation
Pengcheng Yang 2019 [32]	China	2001-2012	Retrospective	Single center	MIMIC-III	ANNs,LR,AdaBoost,Bagging	7529	22.18	ten-fold cross-validation

Table 1 (continued)

Author, Year	Country	Selected time	Study type	Single/Multi-center	Database	Model	Patients（n）	ARDS Incidence (%)	Method of validation
Carson Lam 2021 [17]	America	2019.5.1-2021.5.1	Retrospective	Multi-center	7 US hospitals	RNN	3457	97.86	/
Jipeng Mo 2022 [27]	China	2020.2-2021.2	Retrospective	Multi-center	Changzhou Second People's Hospital,Jiangsu Provincial People's Hospital,Nanjing Military Region General Hospital,Wuxi Fifth People's Hospital	ANNs	124	17.74	/
Weiwei Lu 2023 [39]	China	2015-2020	Retrospective	Single center	Xinhua Hospital affiliated to Shanghai Jiaotong University School of Medicine	XGBoost,SVM,ANNs,LR,RF	308	7.79	ten-fold cross-validation

XGB Xtreme Gradient Boosting, *RF* Random forest, *LR* Logistic regression, *CNN* Convolutional neural network, *SVM* Support vector machine, *GBT* Gradient boosted trees, *KNN* k-nearest neighbors, *GLMNET* The lasso and elastic-net regularized generalized linear models, *BAYESGLM* Bayesian generalized linear models, *GAM* Generalized additive models, *GBM* Gradient boosting machines, *NN* Neural networks, *EDTs* Ensembles of Decision Trees, *BC* Bayesian Classifier, *SLP-FNN* Single hidden layer feedforward neural network, *L2-LR* L2 regularized logistic regression, *AdaBoost* Adaptive boosting, *ML* Machine learning, *PCA* Principal Component Analysis, *LDA* Linear Discriminant Analysis, *DT* Decision tree, *DNN* Deep Neural Network, *ANN* Artificial neural network, *NLP* Natural Language Processing, *RNN* Recurrent Neural Network, *SLR* Stepwise logistic regression algorithm

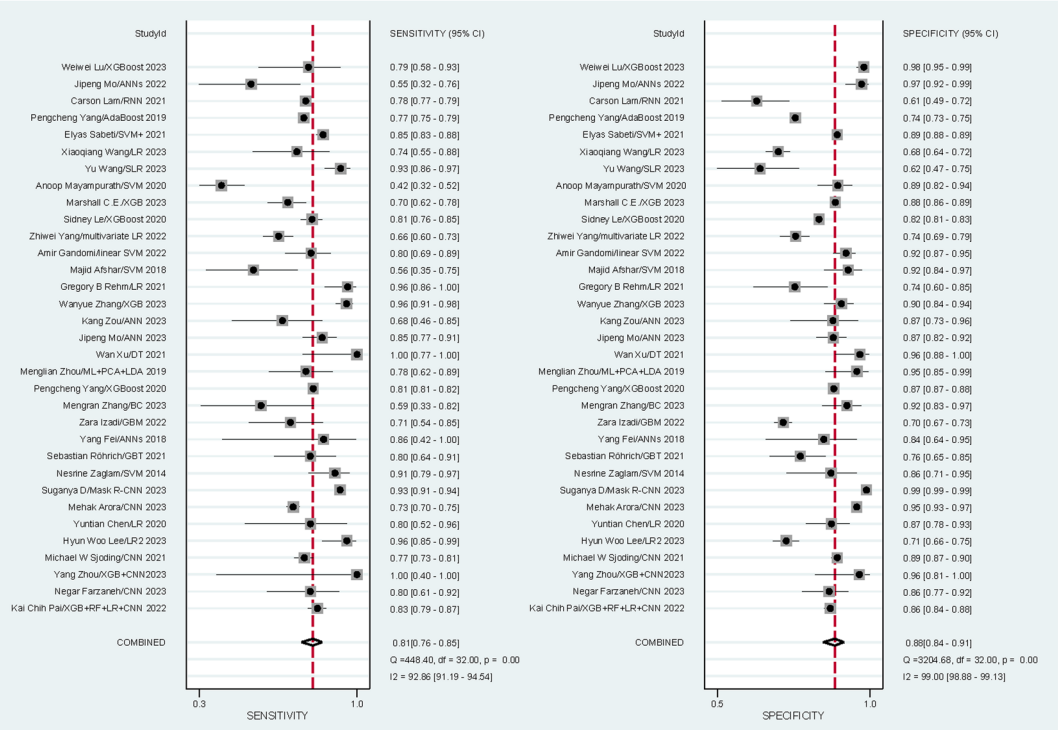

Fig. 2 Forest plots of the pooled sensitivity and specificity for best models

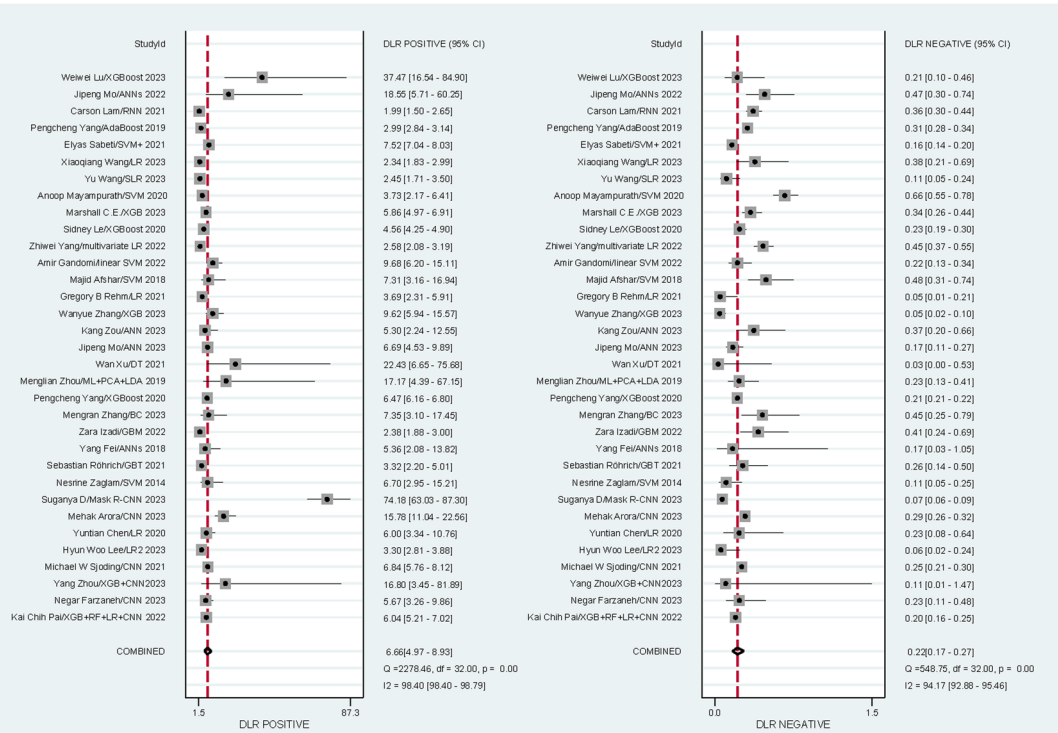

Fig. 3 Forest plots of the pooled PLR and NLR for best models

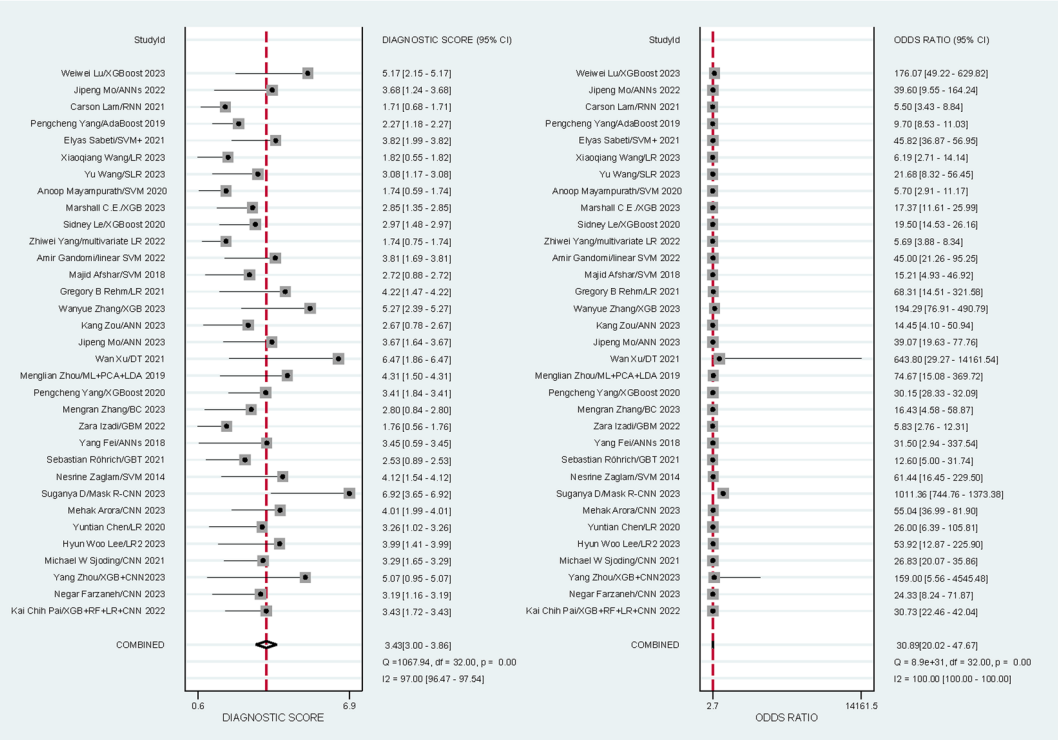

Fig. 4 Forest plots of the pooled DOR for best models

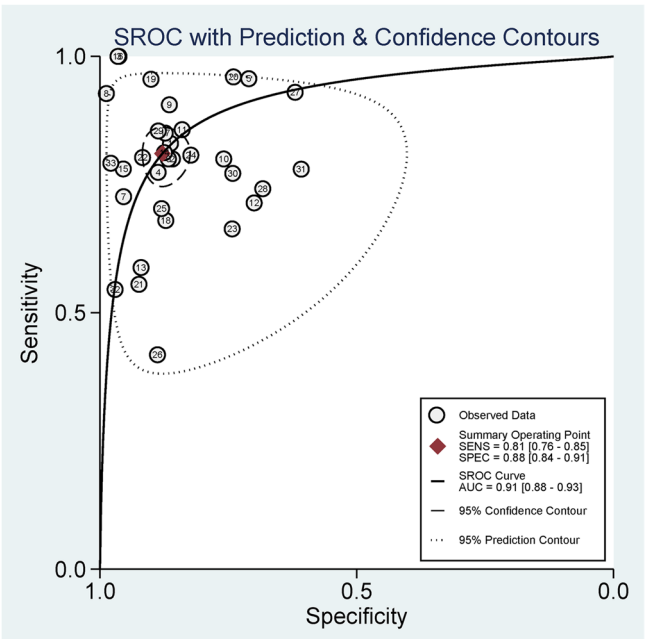

Fig. 5 SROC of best models for predicting ARDS

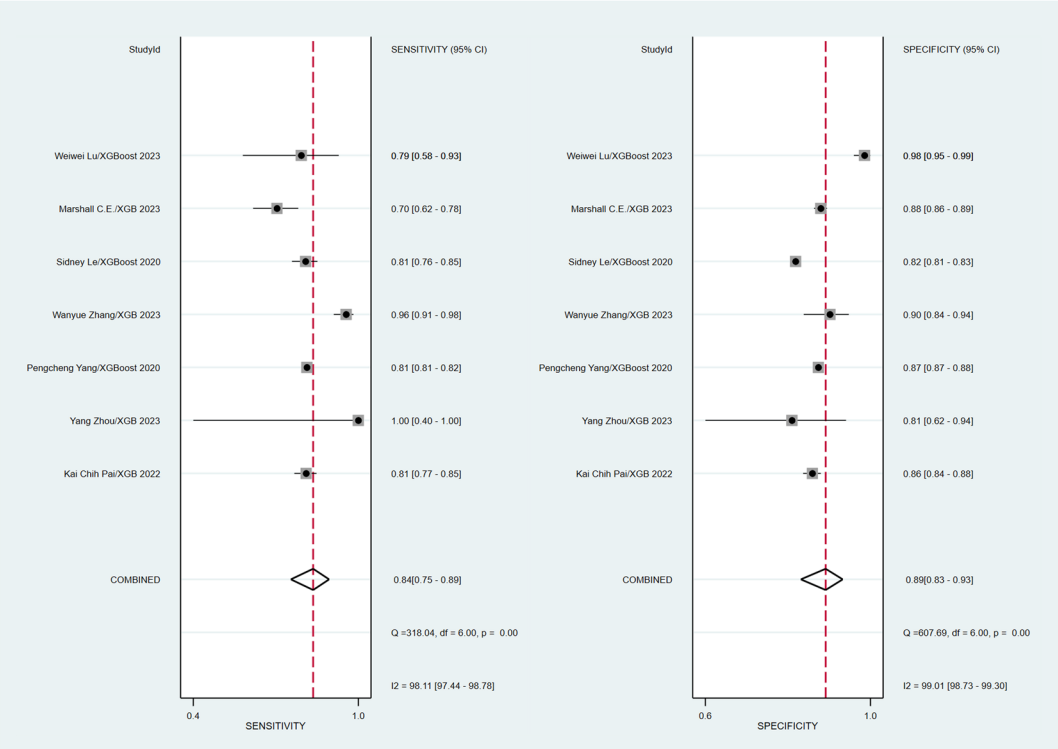

Fig. 6 Forest plots of the pooled sensitivity and specificity for XGB models

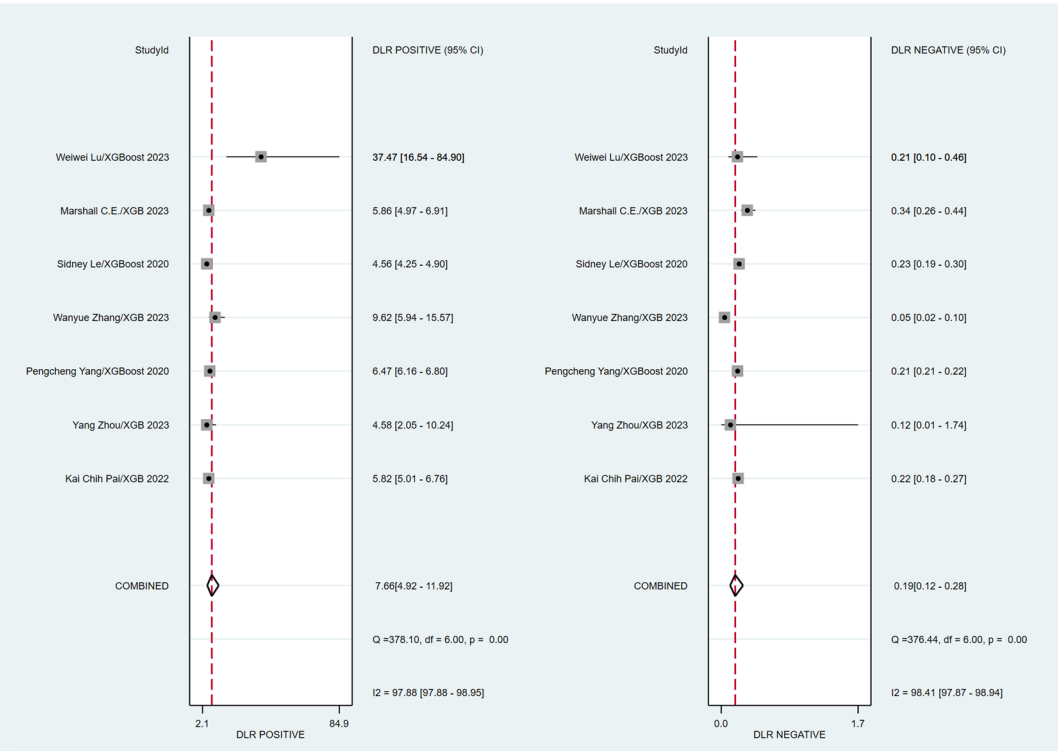

Fig. 7 Forest plots of the pooled PLR and NLR for XGB models

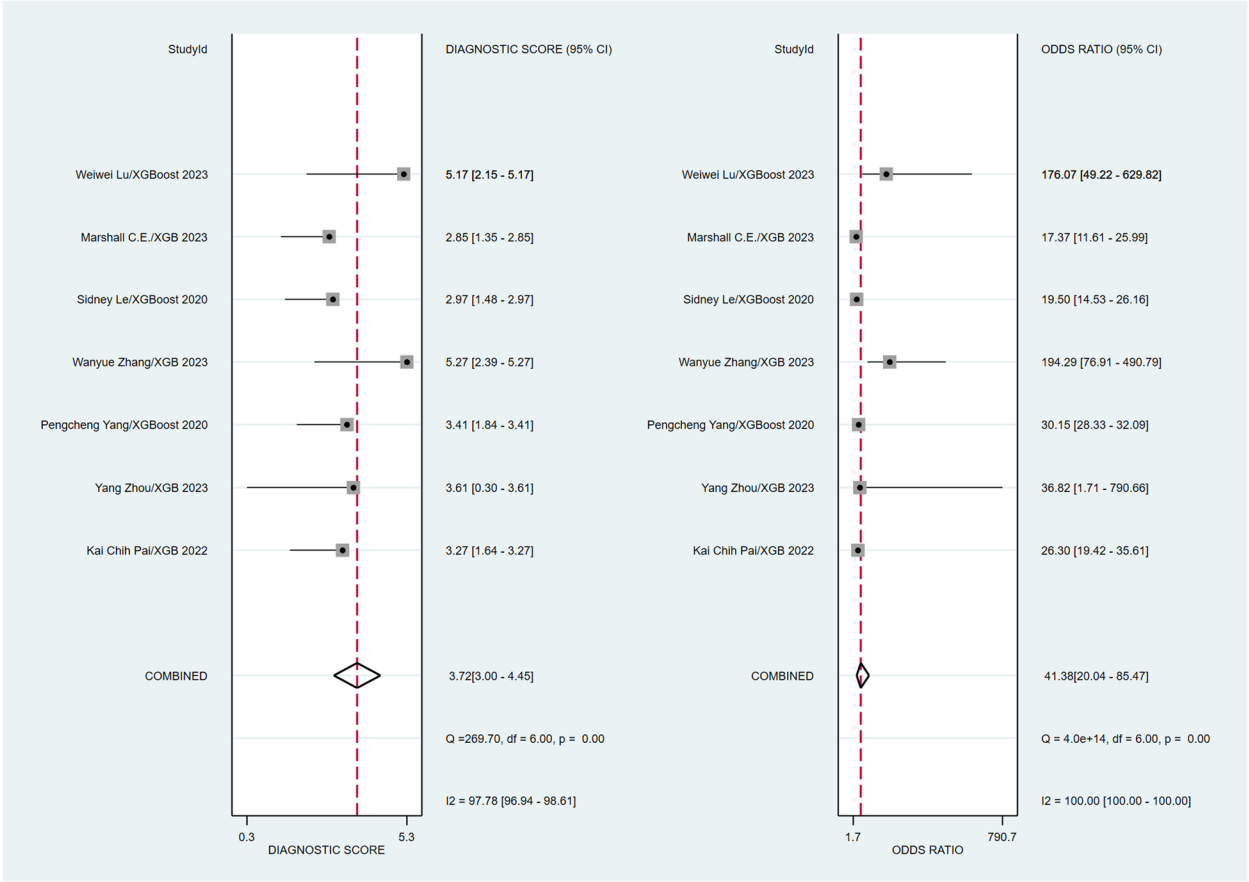

Fig. 8 Forest plots of the pooled DOR for XGB models

and 12 multicenter studies. All 33 studies were retrospective.

Literature quality evaluation

According to the QUADAS-2 tool, the overall risk of bias in patient selection was "high" for 1 study and "unclear" for 6 studies. The risk of bias related to the index test and the flow and timing was "low" for all studies. The risk of bias for the reference standard test was "unclear" for 6 studies. In terms of overall applicability, the risk of patient selection bias was "unclear" for 5 studies, and the risk of bias for the reference standard test was "unclear" for 6 studies. (Supplementary Material 2 Figures S1, S2)

Results of the meta-analysis

Best models

The performance of the best AI model in predicting ARDS was evaluated in the 33 studies. The pooled sensitivity was 0.81 (0.76–0.85), the pooled specificity was 0.88 (0.84–0.91), the pooled PLR was 6.66 (4.97–8.93), the pooled NLR was 0.22 (0.17–0.27), the pooled diagnostic odds ratio (DOR) was 31 (20–48), and the overall pooled AUC was 0.91 (0.88–0.93). (Figures 2, 3, 4 and 5) After excluding seven studies with patient selection bias risk and six studies with reference standard test bias risk, it was found that their impact on the final results was very limited, with AUC values of 0.91 (0.88–0.93) and 0.91 (0.89–0.94), respectively. (Supplementary Material 1 Table S2) Five studies included external test data, and 12 studies were multicenter studies. Their pooled AUC values were 0.91 (0.88–0.93) and 0.92 (0.89–0.94), respectively, further validating the reliability of our research findings. (Supplementary Material 1 Table S3)

All models

A total of 28 models predicted ARDS. The pooled sensitivity was 0.79 (0.76–0.82), the pooled specificity was 0.85 (0.83–0.88), the pooled PLR was 5.37 (4.55–6.35), the pooled NLR was 0.25 (0.21–0.28), the DOR was 22 (17–28), and the overall pooled AUC was 0.89 (0.86–0.91). (Supplementary Material 2 Figures S3-S6)

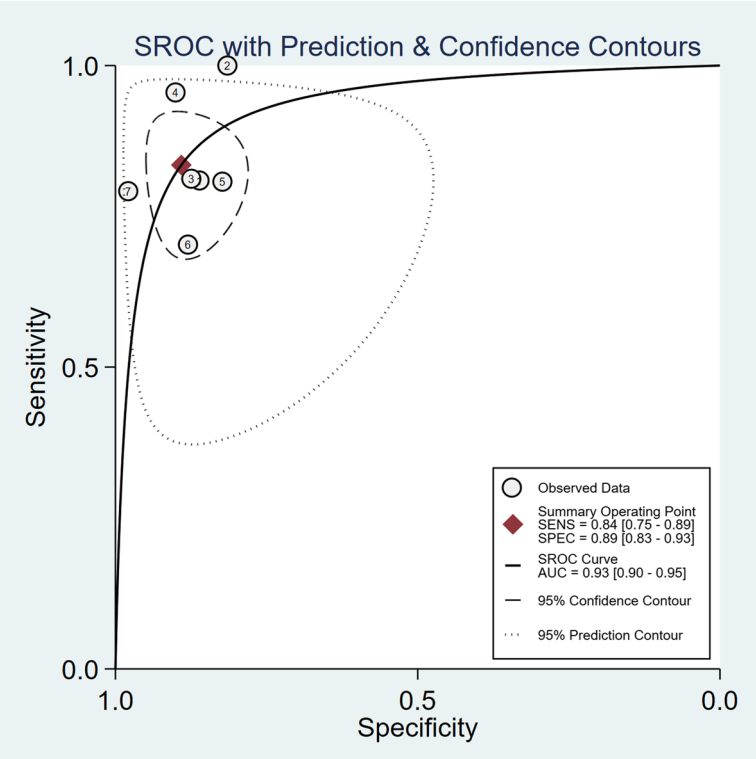

Fig. 9 SROC of XGB models for predicting ARDS

Different types of AI models

In this meta-analysis, we analyzed the performance of XGB, RF, LR, CNN, SVM, and ANN models in predicting ARDS. Their pooled AUCs were 0.93 (0.90–0.95), 0.89 (0.86–0.91), 0.86 (0.83–0.89), 0.91 (0.88–0.93), 0.90 (0.87–0.92), and 0.86 (0.83–0.89), respectively. The XGB was the best predictive model, with the pooled sensitivity of 0.84 (0.75–0.89), the pooled specificity of 0.89 (0.83–0.93), the pooled PLR of 7.66 (4.92–11.92), the pooled NLR of 0.19 (0.12–0.28), the pooled DOR of 41 (20–85) (Table 2; Figs. 6, 7, 8 and 9). (Forest plots and SROC curves for the other analyzed models are shown in Supplementary Material 2 Figures S7-S26.)

Models with different predictors

The effectiveness of different predictors as variables to train models for predicting ARDS was analyzed by training the models with images, images combined with other predictors, mechanical ventilation parameters, laboratory assays, and other predictors. Their pooled AUC were 0.90 (0.88–0.93), 0.92 (0.89–0.94), 0.87 (0.83–0.89), 0.91 (0.88–0.93), and 0.78 (0.74–0.81) respectively. In particular, the model trained with images combined with other predictors exhibited the best prediction with the pooled sensitivity of 0.85 (0.80–0.89), the pooled specificity of 0.86 (0.82–0.89), the pooled PLR of 6.10 (4.98–7.46), the pooled NLR of 0.18 (0.14–0.22), and the pooled DOR of 35 (29–41) (Table 3; Figs. 10, 11, 12 and 13). (Forest

Table 2 Different types of AI models

Model	Study	Sensitivity	Sen-I^2	Specificity	Spe-I^2	AUC
XGB	7	0.84 (0.75–0.89)	98.11 (97.44–98.78)	0.89 (0.83–0.93)	99.01 (98.73–99.30)	0.93 (0.90–0.95)
RF	5	0.80 (0.73–0.86)	75.58 (53.66–97.50)	0.90 (0.80–0.96)	94.03 (90.34–97.72)	0.89 (0.86–0.91)
LR	17	0.80 (0.71–0.86)	97.91 (97.44–98.39)	0.80 (0.74–0.84)	96.98 (96.22–97.75)	0.86 (0.83–0.89)
CNN	7	0.80 (0.72–0.86)	95.30 (93.07–97.52)	0.91 (0.83–0.96)	99.47 (99.35–99.59)	0.91 (0.88–0.93)
SVM	12	0.72 (0.54–0.85)	94.43 (92.36–96.49)	0.89 (0.82–0.94)	98.80 (98.53–99.07)	0.90 (0.87–0.92)
ANN	5	0.74 (0.64–0.83)	95.72 (93.32–98.12)	0.88 (0.77–0.94)	98.92 (98.53–99.31)	0.86 (0.83–0.89)

XGB Xtreme Gradient Boosting, *RF* Random forest, *LR* Logistic regression, *CNN* Convolutional neural network, *SVM* Support vector machine, *ANN* Artificial neural network

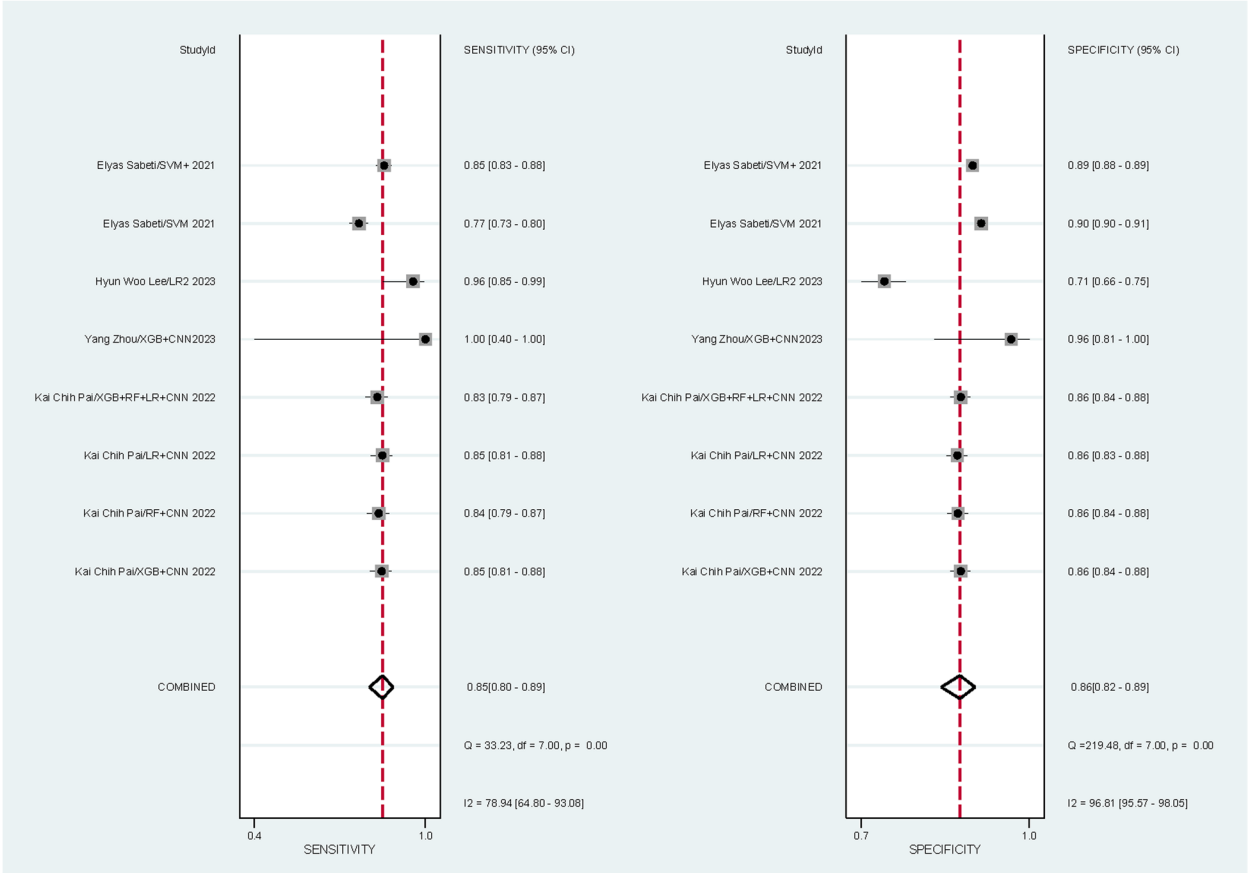

Fig. 10 The pooled sensitivity and specificity for models trained with images combined with other predictors

plots and SROC curves for training models for different predictors are shown in Supplementary Material 2 Figures S27-S42) (Details of the predictors are shown in Supplementary Material 1 Table S1)

Model trained with specific factors

We conducted a thorough analysis of models trained using specific predictive variables. Specifically, CNNs were primarily trained using Images data, achieving an AUC of 0.91 (0.88–0.93); LR focused on Mechanical ventilation parameters, yielding an AUC of 0.80 (0.76–0.83); and Laboratory assays were utilized to train LR, ANN, and SVM, with their respective AUCs being 0.88 (0.85–0.90), 0.92 (0.89–0.94), and 0.91 (0.89–0.94). (Supplementary Material 1 Table S4)

Publication bias detection

The publication bias test using Deeks showed $P=0.95$ ($P>0.05$). The Deeks funnel plot (Fig. 14) revealed that the angle between the regression line and the DOR axis was close to 90, and no significant asymmetry was present, suggesting a low likelihood of publication bias.

Heterogeneity analysis

Threshold effect analysis

The threshold effect analysis showed a Spearman correlation coefficient of 0.090 and a *P*-value of 0.619. There was no threshold effect between the studies in this inclusion.

Meta-regression analysis

In the combined analysis of the optimal models from 33 studies, significant heterogeneity was observed in both sensitivity and specificity (with I^2 values of 92.86% and 99.00%, respectively). To explore the underlying causes of this heterogeneity, we conducted a meta-regression analysis, examining factors such as study region, definition of ARDS, type of research center (multi-center vs. single-center), number of patients, incidence of ARDS, type of AI model, and type of predictor variable. As illustrated in fig. 15 all factors except for the type of research

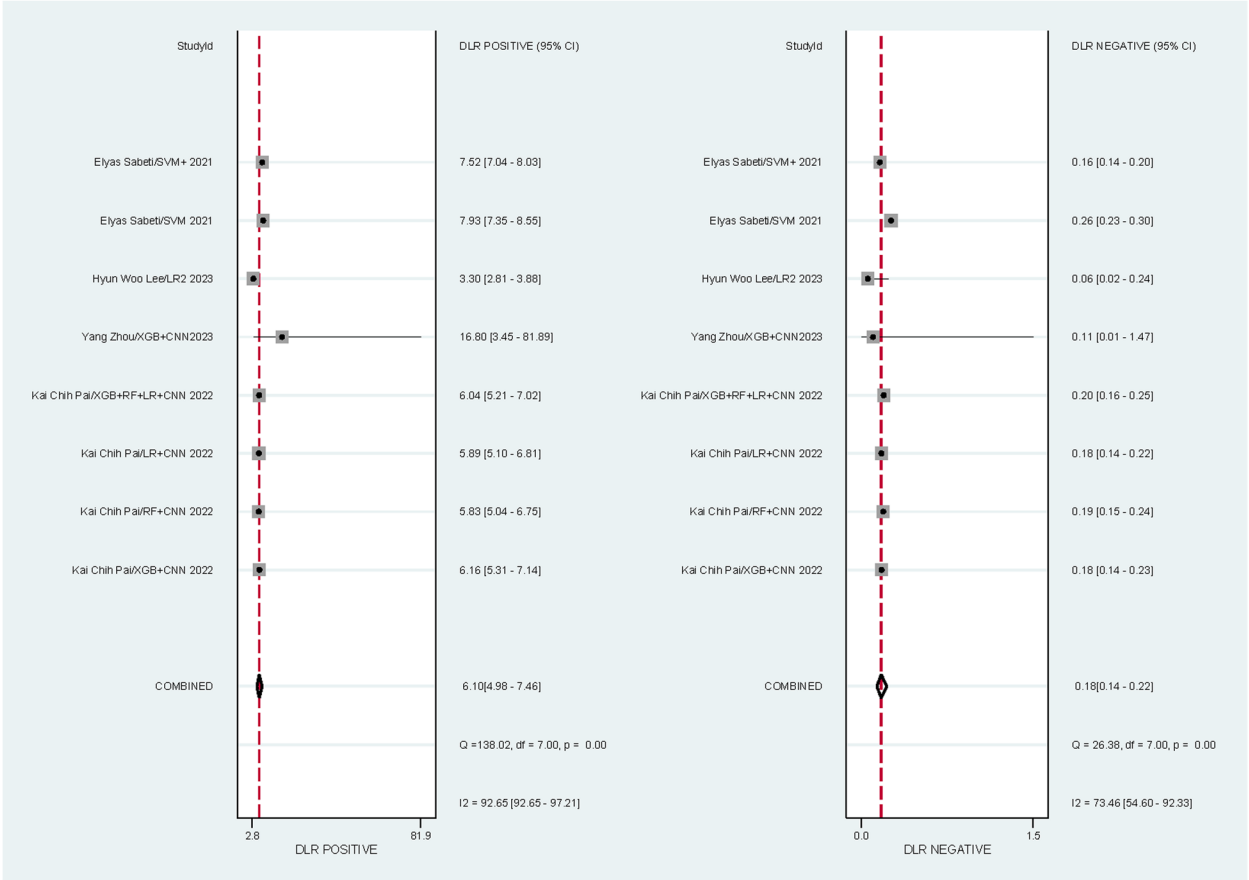

Fig. 11 The pooled PLR and NLR for models trained with images combined with other predictors

center were considered as sources of heterogeneity (with *P*-values < 0.05).

Sensitivity analysis

We conducted a comprehensive sensitivity analysis on the included studies, with particular attention to the impact of excluding each of the studies by Suganya D/2023, Wanyue Zhang/2023, Anoop Mayampurath/2020, and Yu Wang/2023. After recalculating the I^2 values for the remaining studies, we found that heterogeneity still persisted (I^2 value > 50%). Using MIDAS to integrate the effect sizes, the combined AUC value did not exhibit significant fluctuation compared to the total combined AUC value of 0.91 (0.88–0.93). (Fig. 16; Table 4) This suggests that these four studies have a limited impact on the accuracy of AI in predicting ARDS, and the research findings are stable and reliable. Additionally, we performed sensitivity analyses on various subgroups and observed that after excluding the study by Wanyue Zhang/XGB 2023, the combined AUC value for the XGB group decreased to 0.82 (with a total combined AUC value of 0.93). Similarly, after removing the study by Suganya D/Mask R-CNN 2023, the combined AUC values for the CNN group and the image group decreased to 0.80 (with a total combined AUC value of 0.91) and 0.82 (with a total combined AUC value of 0.90), respectively. (Supplementary Material 1 Table S5, Supplementary Material 2 Figs. 43, 44, 45, 46, 47, 48, 49, 50 and 51).

Discussion

AI, with several powerful algorithms, has significantly progressed the fields of image recognition, analysis of big data, natural language processing, and decision-making assistance, substantially developing various medical fields. In recent years, AI has been gradually applied to predict, recognize, and diagnose multiple diseases. Bacci et al. systematically evaluated the use of AI models in AKI prediction. The authors reported that the AUC results of the included studies could reach up to 0.70 [40]. Moreover, Silva et al. also indicated that AI performed well to detect prostate cancer with an optimal sensitivity of 1.0 (0.93–1.0) and a specificity of 0.78 (0.64–0.89) [41]. Several studies have conducted systematic evaluations on the application of AI in ARDS, yet these studies have solely

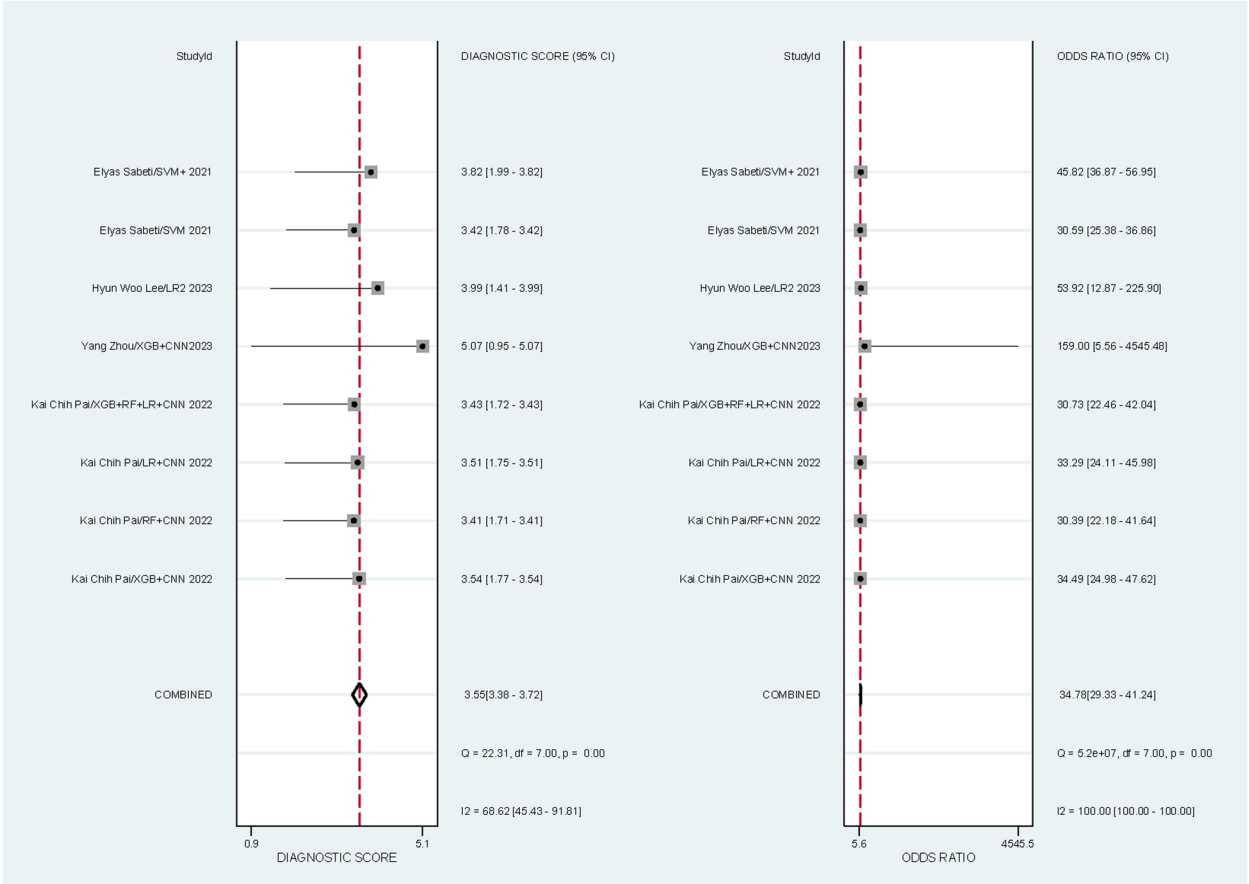

Fig. 12 The pooled DOR for models trained with images combined with other predictors

focused on analyzing AI models that interpret imaging data [42–45]. In contrast, the research by Muhammed Rashid et al. provides a comprehensive examination of the multifaceted applications of AI in ARDS, highlighting its promising role in diagnosis, risk stratification, severity prediction, management, mortality prediction, and decision-making analysis [46]. Similarly, T.K. Tran et al. have analyzed the utilization of machine learning in ARDS research, encompassing a broad range of research areas spanning seven distinct categories, notably including diagnosis [47]. However, none of these studies conducted a meta-analysis to delve deeper into the research questions. This study addresses this gap by assessing the accuracy of AI in predicting ARDS, as well as the actual effectiveness and advantages of AI in managing ARDS.

In this meta-analysis, the best models exhibited a pooled sensitivity of 0.81 (0.76–0.85), a pooled specificity of 0.88 (0.84–0.91), and a pooled AUC of 0.91 (0.88–0.93). This suggests that AI can recognize sick and non-sick people with credible results. The results of subgroup analysis conducted on various artificial intelligence models indicate that XGB demonstrates optimal performance, with an AUC value of 0.93 (0.90–0.95), which aligns with the general trends observed in machine learning. As an ensemble tree model, XGB excels in handling data with complex interactions and high dimensionality. Compared to other machine learning models, XGB not only leads in prediction accuracy but also exhibits superior training efficiency and scalability [48]. A study suggests that AI models with high AUC values should incorporate additional predictive indicators [48]. However, Izadi et al. suggested that a large number of predictors might not improve the predictive performance of a model. In particular, selecting appropriate predictors may be crucial, which is more practical in clinical practice [26]. In this meta-analysis, the model trained with images combined with other predictors had an AUC of 0.92 (0.89–0.94), showing the best prediction performance. The AUCs of the models trained for images and laboratory assays were 0.90 (0.88–0.93) and 0.91 (0.88–0.93), respectively. This suggests that imaging and laboratory assays may be more suitable for training ARDS

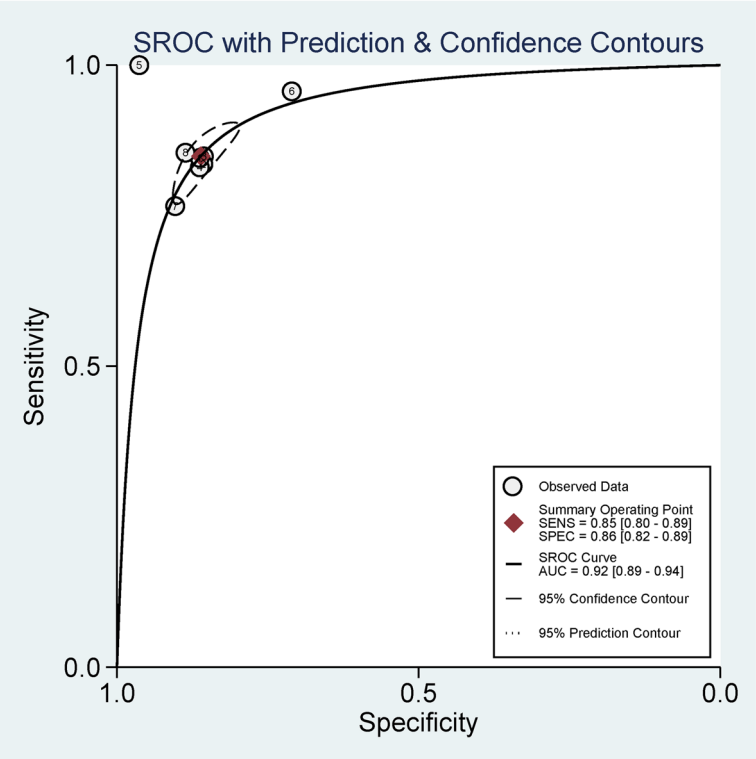

Fig. 13 SROC of models trained with images combined with other predictors for predicting ARDS

prediction models. The two subgroups analyzed in this study provide useful references for selecting appropriate models and predictors in future research.

We further observed that CNNs, a deep learning model specialized for processing image and video data, yielded promising prediction outcomes. Models trained based on laboratory examinations encompassed LR, ANN, and SVM. Among them, ANNs demonstrated the superior predictive performance, attributed to their high accuracy, parallel distributed processing capabilities, and robustness against noise. In our dataset, LR models were more frequently utilized in conjunction with mechanical ventilation parameters for training; however, their predictive performance was not satisfactory, with an AUC of 0.80 (0.76–0.83). This may be attributed to the fact that LR models typically only consider pairwise interactions while neglecting other variables and are susceptible to disturbances from nonlinear relationships among predictor variables [35, 49]. In conclusion, selecting an appropriate model according to the specific task and data characteristics is crucial for achieving satisfactory prediction results and performance.

The implementation of artificial intelligence algorithms for real-time prediction aids in timely stratified care for high-risk patients with ARDS, optimizes resource allocation in ICUs, and enhances treatment efficiency. Our comprehensive analysis, integrating numerous related studies, reveals that the application of artificial intelligence in predicting ARDS is accurate and reliable. However, significant challenges persist in data sharing and regulation during the implementation of AI algorithms. To address these issues, the establishment of unified data standards and sharing mechanisms, along with the enhancement of data regulation, is imperative.

Table 3 Models with different predictors

Predictors	Study	Sensitivity	Sen-I^2	Specificity	Spe-I^2	AUC
Images	10	0.81 (0.75–0.86)	93.27 (90.36–96.18)	0.90 (0.82–0.94)	99.25 (99.09–99.41)	0.90 (0.88–0.93)
Images combined with other predictors	8	0.85 (0.80–0.89)	78.94 (64.80–93.08)	0.86 (0.82–0.89)	96.81 (95.57–98.05)	0.92 (0.89–0.94)
Mechanical ventilation parameters	18	0.82 (0.78–0.86)	98.63 (98.37–98.89)	0.78 (0.75–0.81)	98.99 (98.82–99.16)	0.87 (0.83–0.89)
Laboratory assays	31	0.77 (0.70–0.83)	90.26 (87.67–92.86)	0.90 (0.87–0.93)	94.20 (92.88–95.52)	0.91 (0.88–0.93)
Other predictors	14	0.68 (0.62–0.74)	81.65 (72.82–90.48)	0.77 (0.70–0.83)	98.36 (97.98–98.74)	0.78 (0.74–0.81)

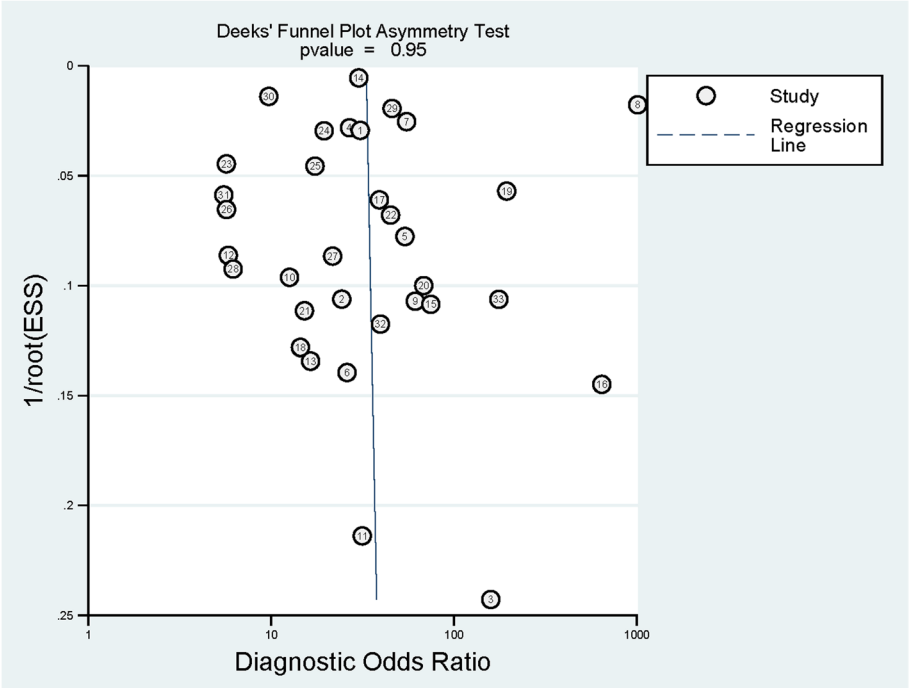

Fig. 14 Funnel plot of studies included in the meta-analysis

Protecting patient privacy is a fundamental principle that must be adhered to in the clinical application of AI. The studies included in this meta-analysis have adopted rigorous data encryption and anonymization measures to ensure the security of patient data, and future research should also focus on this to ensure the reasonable, safe, and sustainable application of AI technology.

This study acknowledges its limitations as well. Firstly, all 33 studies included were retrospective, leading to a high degree of heterogeneity among them. Meta-regression analysis identified various sources of heterogeneity, including study region, definition of ARDS, type of research center, patient number, incidence of ARDS, type of AI model used, and type of predictor variables utilized. High heterogeneity was also observed within subgroups. Different models vary in algorithmic principles, data processing methods, and predictive capabilities, which may serve as sources of heterogeneity. Furthermore, different studies employed various predictors for model training, reflecting different aspects of patients' pathophysiological processes with different sensitivities and specificities, thereby influencing the study results. Sensitivity analysis indicated a lack of robustness in the study outcomes for the XGB group. Within this group, the study by Wanyue Zhang/XGB 2023 reported good prediction results, potentially due to the selection of more appropriate predictors. Upon excluding this study, the combined effect size of the remaining studies significantly decreased. The study by Suganya D/Mask R-CNN 2023 had a large sample size and excellent predictive performance; after excluding this study, the combined effect sizes of both the CNN group and the imaging group declined. Future studies should include more high-quality research for further analysis. Secondly, during the literature retrieval process, we only searched English and Chinese databases, and all articles eventually included in the meta-analysis were published in either English or Chinese. This outcome may inadvertently exclude important studies in other languages, introducing language bias and affecting the comprehensiveness and accuracy of the results. While we did conduct a comprehensive search across multiple authoritative databases and adhered to a systematic literature review method, even the most thorough search may fail to capture all relevant studies, particularly those that are unpublished or difficult to access, which may result in selection bias in the literature. Lastly, the included studies lacked data on model calibration metrics (such as the Brier score), which hindered our accurate assessment of the model's reliability. The interpretability of a model is crucial for understanding its practicality.

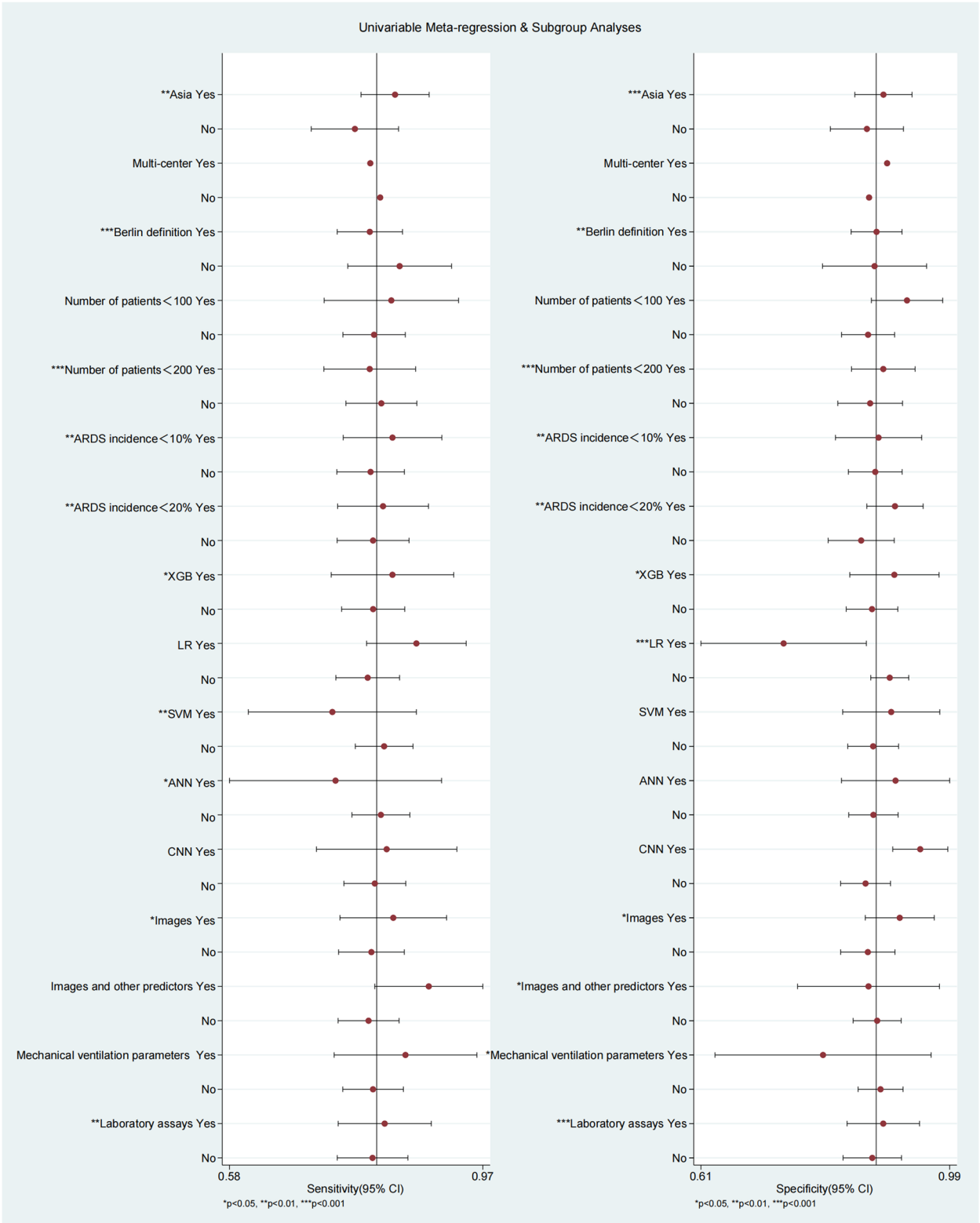

Fig. 15 Meta-regression of heterogeneous sources

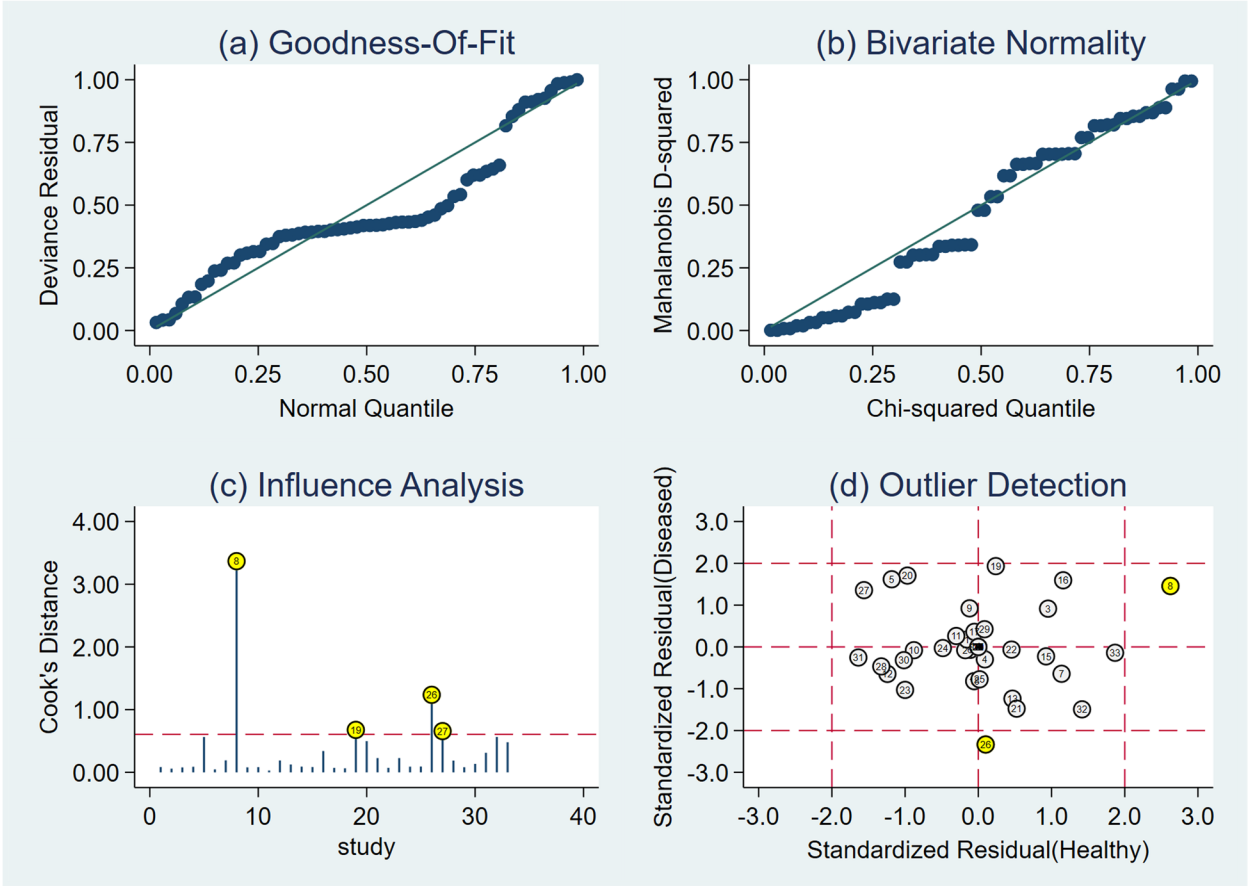

Fig. 16 The analysis of sensitivity

Table 4 The results of excluding each of the four studies one by one

Excluded studies	Sensitivity	Sen-I^2	Specificity	Sep-I^2	AUC
Suganya D 2023 [30]	0.80 (0.75–0.84)	90.32	0.87 (0.83–0.90)	97.89	0.90 (0.87–0.93)
Wanyue Zhang 2023 [36]	0.80 (0.75–0.84)	92.21	0.88 (0.84–0.91)	99.01	0.90 (0.87–0.93)
Anoop Mayampurath 2020 [2]	0.82 (0.78–0.85)	91.94	0.88 (0.84–0.91)	99.05	0.91 (0.88–0.93)
Yu Wang 2023 [11]	0.80 (0.76–0.85)	92.92	0.88 (0.84–0.91)	99.03	0.91 (0.88–0.93)
All four studies were excluded	0.79 (0.75–0.82)	84.69	0.87 (0.83–0.90)	98.03	0.89 (0.85–0.91)

In the studies included in this paper, although various variables such as imaging, vital signs, and laboratory indicators were used to train the model, the lack of data that can quantify variable contributions, such as SHAP values, prevented us from conducting an in-depth integrated analysis of these variables

Conclusion

AI models show good sensitivity and specificity for predicting ARDS, promising future clinical applications. Among them, CNN, SVM, and XGB models exhibited the best prediction performance. The subgroup analyses revealed that models trained with images combined with other predictors showed the best predictive performance. A future work could focus on selecting the model and predictors according to the specific task and data characteristics.

Abbreviations

ARDS	Acute respiratory distress syndrome
AI	Artificial intelligence
XGB	Xtreme Gradient Boosting
RF	Random forest
LR	Logistic regression

CNN	Convolutional neural network
SVM	Support vector machine
GBT	Gradient boosted trees
KNN	k-nearest neighbors
GLMNET	The lasso and elastic-net regularized generalized linear models
BAYESGLM	Bayesian generalized linear models
GAM	Generalized additive models
GBM	Gradient boosting machines
NN	Neural networks
EDTs	Ensembles of Decision Trees
BC	Bayesian Classifier
SLP-FNN	Single hidden layer feedforward neural network
L2-LR	L2 regularized logistic regression
AdaBoost	Adaptive boosting
ML	Machine learning
PCA	Principal Component Analysis
LDA	Linear Discriminant Analysis
DT	Decision tree
DNN	Deep Neural Network
ANN	Artificial neural network
SLR	Stepwise logistic regression algorithm
RNN	Recurrent Neural Network
NLR	Negative likelihood ratio
PLR	Positive likelihood ratio
DOR	Diagnostic odds ratio
SEN	Sensitivity
SPE	Specificity
TP	True positive
FP	False positive
TN	True negative
FN	False negative
AUC	Area under the curve
SROC	Summary receiver operating characteristic

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12911-025-02869-0.

Supplementary Material 1.

Supplementary Material 2.

Supplementary Material 3.

Supplementary Material 4.

Supplementary Material 5.

Supplementary Material 6.

Acknowledgements
Heilongjiang Province Key R&D Program (JD22C005), the National Natural Scientific Foundation of China [82172164], and Heilongjiang Province Key R&D Program (GA21C011).

Authors' contributions
Y. X. wrote the manuscript, Y. G., Y. Q. and Y. Z. performed the literature review. J. X. , K. W. and Q. Y. performed the statistical analysis. C. W., M. Z, X. M revised the text. All authors read and approved the final manuscript.

Funding
Heilongjiang Province Key R&D Program (JD22C005), the National Natural Scientific Foundation of China [82172164], and Heilongjiang Province Key R&D Program (GA21C011).

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Critical Care Medicine, First Affiliated Hospital of Harbin Medical University, Heilongjiang, China. [2]Heilongjiang Provincial Key Laboratory of Critical Care Medicine, Heilongjiang, China.

Received: 15 January 2024 Accepted: 14 January 2025
Published online: 28 January 2025

References

1. Bellani G, Laffey JG, Pham T, Fan E, Brochard L, Esteban A, Gattinoni L, van Haren F, Larsson A, McAuley DF, et al. Epidemiology, patterns of care, and mortality for patients with acute respiratory distress syndrome in intensive care units in 50 countries. JAMA. 2016;315(8):788–800.
2. Mayampurath A, Churpek MM, Su X, Shah S, Munroe E, Patel B, Dligach D, Afshar M. External Validation of an Acute Respiratory Distress Syndrome Prediction Model Using Radiology Reports. Crit Care Med. 2020;48(9):e791–98.
3. Schwalbe N, Wahl B. Artificial intelligence and the future of global health. Lancet. 2020;395(10236):1579–86.
4. Ngiam KY, Khor W. Big data and machine learning algorithms for healthcare delivery. Lancet Oncol. 2019;20(5):e262–73.
5. Rashid M, Ramakrishnan M, Chandran VP, Nandish S, Nair S, Shanbhag V, Thunga G. Artificial intelligence in acute respiratory distress syndrome: a systematic review. Artif Intell Med 2022:102361.
6. Lee SW, Koo MJ. PRISMA 2020 statement and guidelines for systematic review and meta-analysis articles, and their underlying mathematics: Life Cycle Committee Recommendations. Life Cycle. 2022;2:e9.
7. Whiting PF, Rutjes AW, Westwood ME, Westwood ME, Mallett S, Mallett S, Deeks JJ, Reitsma JB, Reitsma JB, Leeflang MMG, Sterne JAC, Bossuyt PMM. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Annals Int Med. 2011;155(8):529–36.
8. Gandomi A, Wu P, Clement DR, Xing J, Aviv R, Federbush M, et al. ARDSFlag: An NLP/Machine Learning Algorithm to Visualize and Detect High-Probability ARDS Admissions Independent of Provider Recognition and Billing Codes. 2022. Preprint at https://doi.org/10.1101/2022.09.27.22280416.
9. Afshar M, Joyce C, Oakey A, Formanek P, Yang P, Churpek MM, Cooper RS, Zelisko S, Price R, Dligach D. A computable phenotype for acute respiratory distress syndrome using natural language processing and machine learning. AMIA Annual Symp Proc AMIA Symp. 2018;2018:157–65.
10. Zaglam N, Jouvet P, Flechelles O, Emeriaud G, Cheriet F. Computer-aided diagnosis system for the acute respiratory distress syndrome from chest radiographs. Comput Biol Med. 2014;52:41–8.
11. Wang Y, Chen L, Yao C, Wang T, Wu J, Shang Y, Li B, Xia H, Huang S, Wang F, et al. Early plasma proteomic biomarkers and prediction model of acute respiratory distress syndrome after cardiopulmonary bypass: a prospective nested cohort study. Int J Surg (London England). 2023;109(9):2561–73.
12. Zhang M, Pang M. Early prediction of acute respiratory distress syndrome complicated by acute pancreatitis based on four machine learning models. Clinics 2023, 78;100215.
13. Sabeti E, Drews J, Reamaroon N, Warner E, Sjoding MW, Gryak J, Najarian K. Learning using partially available Privileged information and label uncertainty: application in detection of Acute Respiratory Distress Syndrome. IEEE J Biomedical Health Inf. 2021;25(3):784–96.

14. Marshall CE, Narendrula S, Wang J, De Souza Vale JG, Jeong H, Krishnan P, et al. A Machine Learning Algorithm to Predict Hypoxic Respiratory Failure and risk of Acute Respiratory Distress Syndrome (ARDS) by Utilizing Features Derived from Electrocardiogram (ECG) and Routinely Clinical Data. 2022. Preprint at https://doi.org/10.1101/2022.11.14.22282274.
15. Rohrich S, Hofmanninger J, Negrin L, Langs G, Prosch H. Radiomics score predicts acute respiratory distress syndrome based on the initial CT scan after trauma. Eur Radiol. 2021;31(8):5443–53.
16. Zhou M, Sharma R, Zhu H, Li Z, Li J, Wang S, Bisco E, Massey J, Pennington A, Sjoding M, et al. Rapid breath analysis for acute respiratory distress syndrome diagnostics using a portable two-dimensional gas chromatography device. Anal Bioanal Chem. 2019;411(24):6435–47.
17. Lam C, Tso CF, Green-Saxena A, Pellegrini E, Iqbal Z, Evans D, Hoffman J, Calvert J, Mao Q, Das R. Semisupervised Deep Learning techniques for Predicting Acute Respiratory Distress Syndrome from Time-Series Clinical Data: Model Development and Validation Study. JMIR Formative Res. 2021;5(9):e28028–28028.
18. Le S, Pellegrini E, Green-Saxena A, Summers C, Hoffman J, Calvert J, Das R. Supervised machine learning for the early prediction of acute respiratory distress syndrome (ARDS). J Crit Care. 2020;60:96–102.
19. Pai K-C, Chao W-C, Huang Y-L, Sheu R-K, Chen L-C, Wang M-S, Lin S-H, Yu Y-Y, Wu C-L, Chan M-C. Artificial intelligence-aided diagnosis model for acute respiratory distress syndrome combining clinical data and chest radiographs. Digit Health 2022, 8;20552076221120317.
20. Fei Y, Gao K, Li WQ. Artificial neural network algorithm model as powerful tool to predict acute lung injury following to severe acute pancreatitis. Pancreatology. 2018;18(8):892–9.
21. Yang Z. Clinical Treatment Analysis and Risk Prediction Model of Acute Respiratory Distress Syndrome [Dissertation]. Master of Medicine: North China University of Science and Technology; 2022. Article at https://d.wanfangdata.com.cn/thesis/ChJUaGVzaXNOZXdTMjAyMzA5MDESCUQwMjk4MTY5ORolYjR4MTd5M2U%3D.
22. Farzaneh N, Ansari S, Lee E, Ward KR, Sjoding MW. Collaborative strategies for deploying artificial intelligence to complement physician diagnoses of acute respiratory distress syndrome. NPJ Digit Med 2023, 6(1):62.
23. Zhou Y, Feng J, Mei S, Tang R, Xing S, Qin S, Zhang Z, Xu Q, Gao Y, He Z. A deep learning model for predicting COVID-19 ARDS in critically ill patients. Front Med 2023, 10;1221711.
24. Sjoding MW, Taylor D, Motyka J, Lee E, Co I, Claar D, McSparron JI, Ansari S, Kerlin MP, Reilly JP, et al. Deep learning to detect acute respiratory distress syndrome on chest radiographs: a retrospective study with external validation. Lancet Digit Health. 2021;3(6):E340–8.
25. Lee HW, Yang HJ, Kim H, Kim U-H, Kim DH, Yoon SH, Ham S-Y, Nam BD, Chae KJ, Lee D et al. Deep learning with chest radiographs for making prognoses in patients with COVID-19: Retrospective Cohort Study. J Med Internet Res 2023, 25;e42717.
26. Izadi Z, Gianfrancesco MA, Aguirre A, Strangfeld A, Mateus EF, Hyrich KL, Gossec L, Carmona L, Lawson-Tovey S, Kearsley-Fleet L, et al. Development of a prediction model for COVID-19 Acute respiratory distress syndrome in patients with rheumatic diseases: results from the global Rheumatology Alliance Registry. ACR open Rheumatol. 2022;4(10):872–82.
27. Mo J, Jia Z, Tang Y, Yang M, Qin H. Establishing prediction model of community-acquired pneumonia complicated with acute respiratory distress syndrome based on artificial neural network. Chin Crit Care Med. 2022;34(4):367–72.
28. Yang P, Wu T, Yu M, Chen F, Wang C, Yuan J, Xu J, Zhang G. A new method for identifying the acute respiratory distress syndrome disease based on noninvasive physiological parameters. PLoS ONE 2020, 15(2);e0226962.
29. Wang X, Zhang H, Zong R, Yu W, Wu F, Li Y. Novel models for early prediction and prevention of acute respiratory distress syndrome in patients following hepatectomy: a clinical translational study based on 1,032 patients. Front Med 2023, 9;1025764.
30. Suganya D, Kalpana R. Prognosticating various acute covid lung disorders from COVID-19 patient using chest CT images. Eng Appl Artif Intell 2023, 119;105820.
31. Chen Y, Wang Y, Zhang Y, Zhang N, Zhao S, Zeng H, Deng W, Huang Z, Liu S, Song B. A quantitative and Radiomics approach to monitoring ARDS in COVID-19 patients based on chest CT: a cohort. Int J Med Sci. 2020;17(12):1773–82.
32. Yang P, Chen F, Zhang G, Yu M, LU M, Wang C, Wang C, Wu T. Research on algorithms for identifying the severity of acute respiratory distress syndrome patients based on noninvasive parameters. J Biomed Eng 2019, 36(3);435.
33. Xu W, Sun N-N, Gao H-N, Chen Z-Y, Yang Y, Ju B, Tang L-L. Risk factors analysis of COVID-19 patients with ARDS and prediction based on machine learning. Sci Rep 2021;11(1);2933.
34. Mo J, Ling S, Yang M, Qin H. Risk of acute respiratory distress syndrome in community-acquired pneumonia patients: use of an artificial neural network model. Emerg Med Int. 2023;2023:2631779.
35. Zou K, Ren W, Huang S, Jiang J, Xu H, Zeng X, Zhang H, Peng Y, Lü M, Tang X. The role of artificial neural networks in prediction of severe acute pancreatitis associated acute respiratory distress syndrome: a retrospective study. Medicine. 2023;102(29):e34399.
36. Zhang W, Chang Y, Ding Y, Zhu Y, Zhao Y, Shi R. To establish an early prediction model for Acute respiratory distress syndrome in severe Acute Pancreatitis using machine learning algorithm. J Clin Med 2023, 12(5);1718.
37. Arora M, Davis CM, Gowda NR, Foster DG, Mondal A, Coopersmith CM, Kamaleswaran R. Uncertainty-aware convolutional neural network for identifying bilateral opacities on chest X-rays: a Tool to Aid diagnosis of Acute Respiratory Distress Syndrome. Bioengineering-Basel 2023, 10(8);946.
38. Rehm GB, Cortes-Puch I, Kuhn BT, et al. Use of Machine Learning to Screen for Acute Respiratory Distress Syndrome Using Raw Ventilator Waveform Data. Critical care explorations. 2021;3(1):e0313–e.
39. Lu W, Chen X, Liu W, Cai W, Zhu S, Wang Y, et al. XGBoost model predicts acute lung injury after acute pancreatitis. Signa Vitae. 2023;19(5):206–12.
40. Bacci MR, Minczuk Cvb Fau -, Fonseca FLA, Fonseca FLA. A systematic review of artificial intelligence algorithms for predicting acute kidney injury. Eur Rev Med Pharmacol Sci. 2023;27(20):9872–9.
41. da Silva LM, Pereira EM, Salles PG, Godrich R, Ceballos R, Kunz JD, Casson A, Viret J, Chandarlapaty S, Ferreira CG, et al. Independent real-world application of a clinical-grade automated prostate cancer detection system. J Pathol. 2021;254(2):147–58.
42. Albahri OS, Zaidan AA, Albahri AS, Zaidan BB, Abdulkareem KH, Al-Qaysi ZT, Alamoodi AH, Aleesa AM, Chyad MA, Alesa RM et al. Systematic review of artificial intelligence techniques in the detection and classification of COVID-19 medical images in terms of evaluation and benchmarking: taxonomy analysis, challenges, future solutions and methodological aspects. J Infect Public Health 2020;13(10):1381–96.
43. Bao C, Liu X, Zhang H, Li Y, Liu J. Coronavirus Disease 2019 (COVID-19) CT findings: a systematic review and Meta-analysis. J Am Coll Radiol 2020;17(6):701–9.
44. Montazeri MA-O, ZahediNasab RA-O, Farahani AA-O, Mohseni HA-O, Ghasemian FA-O. Machine learning models for image-based diagnosis and prognosis of COVID-19: systematic review. JMIR Med Inf 2021;9(4):e25181.
45. Suri JS, Agarwal S, Agarwal S, Gupta S, Puvvula A, Viskovic K, Suri N, Alizad A, El-Baz A, Saba L, Fatemi M, Naidu DS, et al. Systematic Review of Artificial Intelligence in Acute Respiratory Distress Syndrome for COVID-19 Lung Patients: A Biomedical Imaging Perspective. J Biomed Health Inform. 2021;25(11):4128–39.
46. Rashid M, Ramakrishnan M, Chandran VP, Nandish S, Nair S, Shanbhag V, Thunga G. Artificial intelligence in acute respiratory distress syndrome: a systematic review. Artif Intell Med 2022;131:102361.
47. Tran TK, Tran MC, Joseph A, Phan PA, Grau V, Farmery AD. A systematic review of machine learning models for management, prediction and classification of ARDS. Respir Res 2024;4(1):232.
48. Xin Y, Li H, Zhou Y, Yang Q, Mu W, Xiao H, Zhuo Z, Liu H, Wang H, Qu X, et al. The accuracy of artificial intelligence in predicting COVID-19 patient mortality: a systematic review and meta-analysis. BMC Med Inf Decis Mak. 2023;23(1):155.
49. Sutradhar R, Barbera L. Comparing an Artificial neural network to Logistic Regression for Predicting ED visit risk among patients with Cancer: a Population-based Cohort Study. Pain Symptom Manage. 2020;60(1):1–9.

Publisher's note