Review

Artificial intelligence in clinical care amidst COVID-19 pandemic: A systematic review

Eleni S. Adamidi *, Konstantinos Mitsis, Konstantina S. Nikita

Biomedical Simulations and Imaging Lab, School of Electrical and Computer Engineering, National Technical University of Athens, Greece

ARTICLE INFO

ABSTRACT

The worldwide health crisis caused by the SARS-Cov-2 virus has resulted in>3 million deaths so far. Improving early screening, diagnosis and prognosis of the disease are critical steps in assisting healthcare professionals to save lives during this pandemic. Since WHO declared the COVID-19 outbreak as a pandemic, several studies have been conducted using Artificial Intelligence techniques to optimize these steps on clinical settings in terms of quality, accuracy and most importantly time. The objective of this study is to conduct a systematic literature review on published and preprint reports of Artificial Intelligence models developed and validated for screening, diagnosis and prognosis of the coronavirus disease 2019. We included 101 studies, published from January 1st, 2020 to December 30th, 2020, that developed AI prediction models which can be applied in the clinical setting. We identified in total 14 models for screening, 38 diagnostic models for detecting COVID-19 and 50 prognostic models for predicting ICU need, ventilator need, mortality risk, severity assessment or hospital length stay. Moreover, 43 studies were based on medical imaging and 58 studies on the use of clinical parameters, laboratory results or demographic features. Several heterogeneous predictors derived from multimodal data were identified. Analysis of these multimodal data, captured from various sources, in terms of prominence for each category of the included studies, was performed. Finally, Risk of Bias (RoB) analysis was also conducted to examine the applicability of the included studies in the clinical setting and assist healthcare providers, guideline developers, and policymakers.

Contents

1. Introduction

The World Health Organization (WHO) declared on March 11th, 2020 the COVID-19 outbreak, emerged in December 2019 in Wuhan, China [1] resulting at the time of writing in more than 3 million deaths and 150 million cases worldwide. The most critical steps in assisting healthcare professionals to save lives during this pandemic are early screening, diagnosis and prognosis of the disease. Several studies have been conducted using Artificial Intelligence (AI) techniques to optimize these steps on clinical settings in terms of quality, accuracy and time. AI techniques, employing Deep Learning (DL) methods, have demonstrated great success in the medical imaging domain due to DL's advanced capability for feature extraction [2]. Apart from the medical imaging domain, AI techniques are widely used to screen, diagnose and predict prognosis of COVID-19 based on clinical, laboratory and demographic data.

Early clinical course of SARS-CoV2 infection can be difficult to distinguish from other undifferentiated medical presentations to hospital and SARS-CoV-2 PCR testing can take up to 48 h for operational reasons. Limitations of the gold-standard PCR test for COVID-19 have challenged healthcare systems across the world due to shortages of specialist equipment and operators, relatively low test sensitivity and prolonged turnaround times [3]. Hence, rapid identification of COVID-19 is important for delivering care, aiding proper triage among patients admitting to hospitals, accelerating proper treatment and minimizing the risk of infection during presentation and waiting hospital admission time. Several studies have been conducted to face the need of early screening by using AI methods [4,5].

Challenges on COVID-19 diagnosis are also present due to the difficulties of differentiating Chest X-Ray radiographs (CXRs) with COVID-19 pneumonia symptoms from those with common pneumonia and insufficient empirical understanding of the radiological morphology in CT scans of this new type of pneumonia among other. Moreover, CXR or CT-based diagnosis may need laboratory confirmation. Therefore, there is an imperative demand for accurate methods to assist clinical diagnosis of COVID-19. Multiple studies using AI techniques have been conducted towards this direction, to extract valuable features from CXRs or CTs [6,7], to use clinical data and laboratory exams [89] or to even combine both imaging quantitative features and clinical data to result in accurate diagnosis [10,11]. Finally, prognosis is an essential step towards assisting healthcare professionals to predict ICU need,
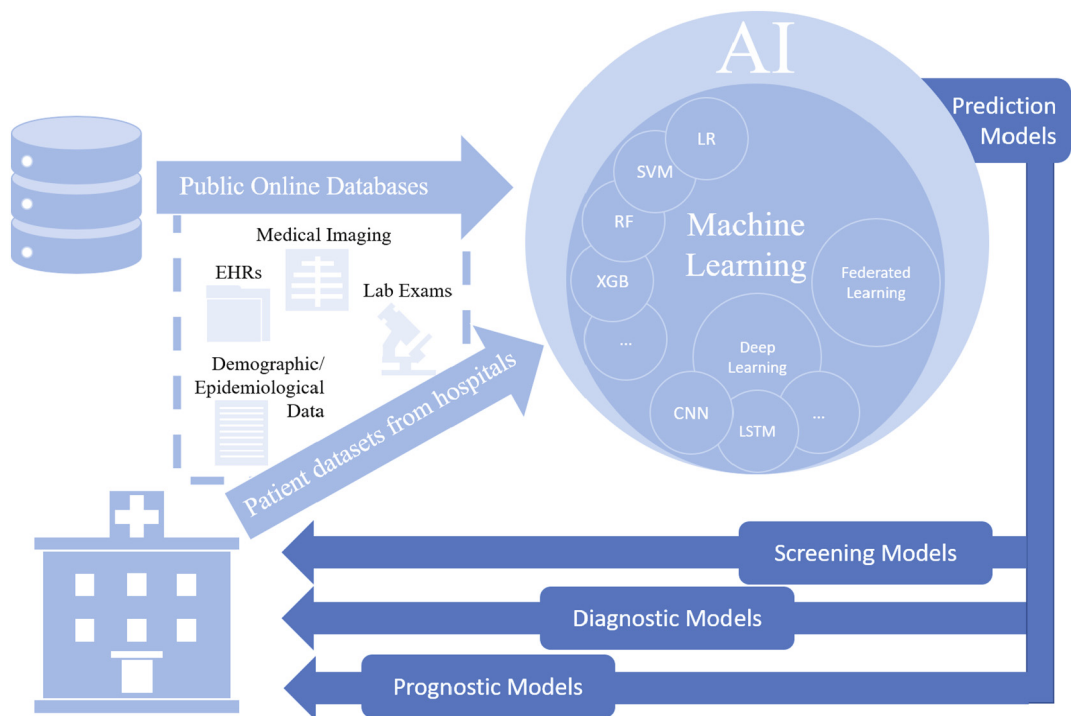

Fig. 1. AI-based clinical prediction models.

mechanical ventilator need, hospitalization time, mortality risk or severity assessment of the disease.

The objective of our study was to conduct a systematic literature review on published and preprint reports of Artificial Intelligence techniques developed and validated for screening, diagnosis, and prognosis of the coronavirus disease 2019. Studies that developed AI prediction models for screening, diagnosis or prognosis that can be applied in the clinical setting were included (see Fig. 1). Screening studies describe prediction models developed for early identification of COVID-19 infection, whereas diagnostic studies propose prediction models developed to establish a diagnosis of the disease. In these studies, several predictors were recognized, including clinical parameters (e.g., comorbidities, symptoms) laboratory results (e.g., hematological, biochemical tests), demographic features (e.g., age, sex, province, country, travel history) or imaging features extracted from CT scans or CXRs. Identification of the most prominent predictors was also part of our analysis. Furthermore, novel technologies incorporated in AI techniques were investigated to determine the current state of research in developing AI prediction models. Additionally, the advantages of using imaging, clinical and laboratory data or the combination of those were analyzed. To achieve this objective, each study was analyzed in terms of COVID-19 positive patients included in the primary datasets, AI methods employed, predictors identified, validation methods applied, and performance metrics used. Finally, a Risk of Bias (RoB) analysis was conducted to examine the applicability of the included studies in the clinical setting and support decisions made by healthcare providers, guideline developers, and policymakers.

The paper is organized as follows. In Section 2, we describe the used methods in the approach and protocols including the description of the AI algorithms performance metrics, and the inclusion–exclusion criteria of the reviewed studies. In Section 3, results are presented on the primary datasets, AI algorithms, validation methods, as well as prediction models developed for screening, diagnostic and prognostic purposes. In this section, we also provide results on the most prominent predictors for each category of the included prediction models. Moreover, the results of the RoB assessment, are provided in Section 3. In Section 4, we discuss the results and the limitations related to the applicability of the developed prediction models and we identify possible future directions aiming at enhancing the adoption of AI-based prediction models in clinical practice.

2. Methods

A. Review approach and protocols

In this systematic literature review, we followed the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) protocol to ensure transparent and complete reporting (see Fig. 2) [12]. This study focused on peer-reviewed publications, as well as preprints published in English, that applied AI techniques to develop prediction models for diagnosis or prognosis of COVID-19. A systematic literature search was conducted for collecting research articles available from January 2020 through December 2020, using the online databases PubMed, Nature, Science Direct, IEEE Xplore, Arxiv and medRxiv. By combining appropriate keywords with Boolean operators, the following expression was formed:

[("artificial intelligence" OR "AI") OR ("machine learning" OR "ML") OR ("deep learning" OR "DL")] AND ("hospital" OR "clinical" OR "healthcare system") AND ("triage" OR "early screening" OR "diagnosis" OR "mortality prediction" OR "severity assessment") AND ("covid-19" OR "sars-cov-2" OR "Coronavirus" OR "pandemic") AND ("prediction models")

Title and abstract screening, full text review, data extraction and Risk of Bias Analysis were conducted by two independent reviewers using Covidence [13], a software for systematic review management.
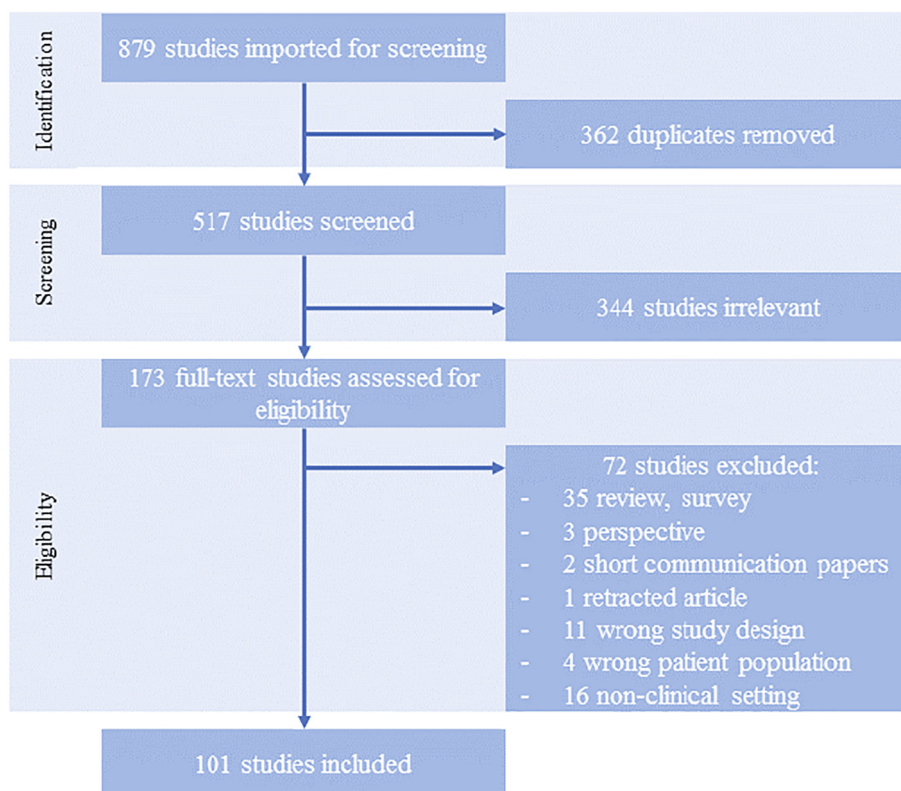

Fig. 2. PRISMA (preferred reporting items for systematic reviews and *meta*-analyses) flowchart.

Data were extracted with Covidence software using a customized extraction form. The extraction form included the following fields for each included study: Covidence Study ID, Lead Author, Title, Database source, Country (the country of dataset origin), Hospital Name, No of hospitals, Start date, End date, Outcome, No of days for mortality prediction, Type of AI model, AI Methods used, Type of input data, Source of input data, Sample Size of input data, Predictors, Study design, Number of participants for model development (with outcomes), Total Number of COVID-19 positive patients, Population description, Validation method, Number of participants for model validation (with outcomes), Performance (Area under the curve (AUC%), Accuracy (Acc%), Sensitivity (SEN %), Specificity (SPE%), Positive Predictive Values/ Negative Predictive Values (PPV/NPV) (%), (95% CI)), Code availability, Limitations, Ethical Considerations, Risk of Bias for participants/ predictors/ outcome/ analysis/, overall risk of bias.

The performance of each AI model was reported in terms of metrics defined using the number of True Positives (TP), True Negatives (TN), False Positive (FP) and False Negatives (FN) [14], as follows:

1) AUC is the area under the Receiver Operating Characteristic (ROC) curve, which plots the true positive rate against the false positive rate. This metric is a standard method for evaluating medical tests and risk models [15].

2) Accuracy is the percentage of cases correctly identified calculated by:

$$Acc = \frac{TP + TN}{TP + FN + FP + TN}$$

3) Sensitivity is the rate of true positives. It measures the proportion of true positives that the model predicts accurately as positive [16], expressed by:

$$SEN = \frac{TP}{TP + FN}$$

4) Specificity is the rate of true negatives. It measures the proportion of true negatives that the model accurately predicts as negative [16], calculated by:

$$SPE = \frac{TN}{FP + TN}$$

5) The positive predictive value (PPV), can be expressed as the ratio of the true positives to the sum of the true positives and false positives and NPV is defined as the ratio of the true negatives to the predicted negatives [17].

In this systematic review, we used PROBAST (Prediction model Risk Of Bias ASsessment Tool) [18], to assess the risk of bias and applicability of the included studies with a focused and transparent approach. PROBAST protocol is organized into the following four domains: participants, predictors, outcome, and analysis. These domains contain a total of 20 signaling questions to facilitate structured judgment of ROB, which was defined to occur when shortcomings in study design, conduct, or analysis led to systematically distorted estimates of model predictive performance.

B. Inclusion - Exclusion criteria

Studies that reported the use of AI techniques, including but not limited to techniques from the AI subfield of Machine Learning (ML) and techniques from the ML subfield of DL for developing pre-
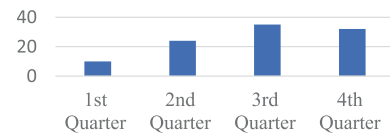
Fig. 3. Number of studies per year 2020 quarter.

diction models for Triage, Diagnosis and Prognosis (such as disease progression, mortality prediction, severity assessment) were included in this systematic review. Cohorts, retrospective cohorts, randomized controlled trials, diagnostic test accuracy, single-centered or multicentered retrospective studies were selected for further analysis.

Restrictions were applied concerning the setting of the studies. If the outcome of the studies could not be applied on a clinical setting, these studies were excluded. Concerning the type of participants, studies that did not include COVID-19 patient data, were excluded. Additionally, studies affecting mental health were excluded. Finally, studies that did not use exclusively AI, ML, or DL to develop these type of prediction models were also excluded.

3. Results

In this review, 879 titles were screened and 101 studies presenting 101 AI-based models for screening, diagnosis and prognosis of COVID-19, were included for full-text review. A significant increase in the number of studies published in the 3rd and 4th quarter of the year was observed (see Fig. 3). We identified in total 14 models for screening (5 based on medical imaging), 38 diagnostic models for detecting COVID-19 (31 based on medical imaging) and 50 prognostic models (7 based on medical imaging) for predicting ICU need, ventilator need, mortality risk, severity assessment or hospital length stay (see Fig. 4). The results are presented in Tables 1-8, including Lead Author of each included study, Country (in case datasets from specific hospitals were used), outcome of the study, number of COVID-19 Positive Patients (CPP) included in the development of the model, AI methods, validation methods and performance of developed prediction models. Missing values of CPP were mainly found in imaging studies that did not specify the number of CT scans or CXR images corresponding to each COVID-19 positive patient. There are in total 11 studies [9,15,19,20,21,22,23,24,25,26,27] with unclear reporting of the number of CPP included.

A. Primary multimodal datasets

Several datasets were identified consisting of multimodal data (e.g., demographic, clinical, imaging) among the included studies. In total 70 studies used COVID-19 positive patient data derived from hospitals from various countries (41 in China, 12 in United States, 6 in Italy, 4 in Brazil, 3 in South Korea, 3 in UK and one study for each of the following countries: France, Spain, Germany, Singapore, Greece, Denmark, Mexico and Israel).

Apart from patient datasets derived from hospitals in the above countries, various publicly available online databases of COVID-19 and pneumonia CXRs ad CT scans were used in the included studies such as:

1) Italian Society of Medical and Interventional Radiology (SIRM) COVID-19 database [28]:

SIRM COVID-19 database reports 384 COVID-19 positive radiographic images (CXR and CT) with varying resolution. Out of 384
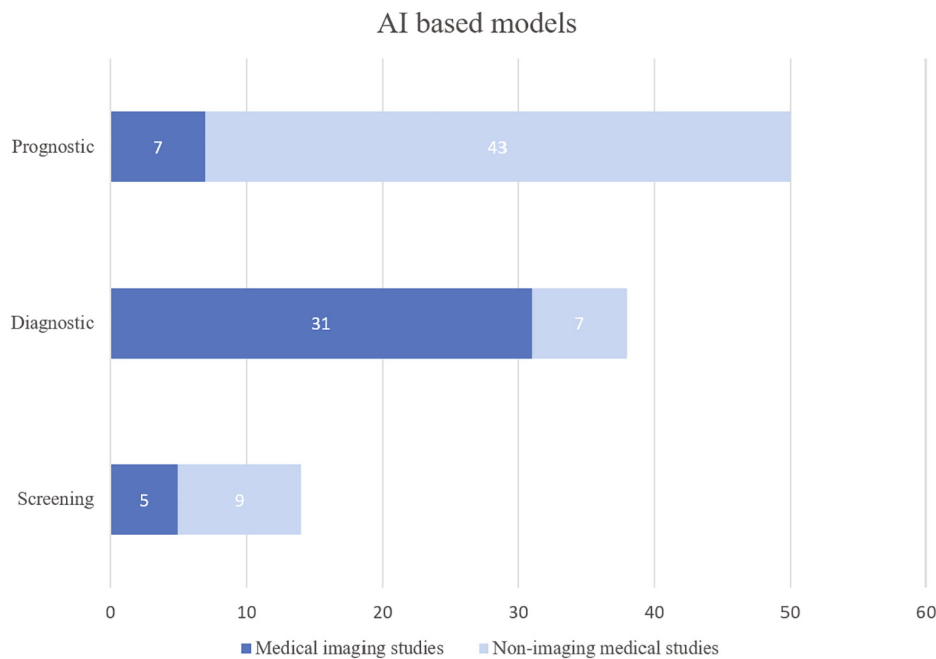
AI based models

Fig. 4. Included AI based prediction models.

radiographic images, 94 images are chest X-ray images and 290 images are lung CT images. This database is updated in a random manner and until 9th December 2020, 115 confirmed COVID-19 cases were reported in this database. This dataset was used by three included studies [29,30,31].

2) *Novel Corona Virus 2019 Dataset* [32]:

This is a public database in GitHub by collecting 319 radiographic images of COVID-19, Middle East respiratory syndrome (MERS), Severe acute respiratory syndrome (SARS) and ARDS from the published articles and online resources. In this database, they have collected 250 COVID-19 positive chest X-ray images and 25 COVID-19 positive lung CT images with varying image resolutions. This dataset was used by twelve studies [5,30,33,19,20,34,35,36,37,21,22,23].

3) *J. Kaggle chest X-ray database* [38]:

This is a very popular database, which has 5,863 images chest X-ray images of normal, viral and bacterial pneumonia with resolution varying from 400p to 2000p. Two included studies [30,22] used this dataset.

4) *K. COVID-19 Radiography Database* [39]:

This database was created for three different types of images classified as chest x-ray images belonging to patients infected with COVID-19, chest x-ray images of cases with viral pneumonia and Chest x-ray images of healthy persons. There are currently 1200 COVID-19 positive images, 1341 normal images, and 1345 viral pneumonia images. This database was used by [40] and [24].

5) *L. COVID-19 cases open database* [41]:

This database contains temporal acquisitions for 42 patients with up to 5 X-ray images per patient, with ground truth annotation as a therapeutic issue for each patient: death or survival. The ground truth annotation can aid in developing prognostic

models like the one presented by Fakhfakh et al. [42] to classify multi-temporal chest X-ray images and predict the evolution of the observed lung pathology based on the combination of convolutional and recurrent neural networks.

Moreover, regarding the type of input data employed in the developed models, several studies used demographic, clinical and imaging data or even a combination of those multimodal data types. Fang *et al.* [43] study, reported to be among the first attempts to fuse clinical data and sequential CT scans to improve the performance of predicting COVID-19 malignant progression in an end to end manner. Liu et al. [44] also used a combination of quantitative CT features of pneumonia lesions with traditional clinical biomarkers to predict the progression to severe illness in the early stages of COVID-19. Additionally, Chassagnon et al. [45] reported an AI solution for performing automatic screening and prognosis based on imaging, clinical, comorbidities and biological data.

B. AI algorithms

The most frequently used AI algorithms for classification purposes in all studies were Random Forests (RF), Linear Regression (LR), Support Vector Machines (SVM), Convolutional Neural Networks (CNN), Decision Trees (DT) and XGBoost (XGB) which are in general six widely used classification methods. Random Forest is a tree-based learning algorithm that utilizes decision trees rising from the training subset which are selected randomly to solve a classification problem [16]. A logistic regression model predicts the probability of a categorical dependent variable occurring [16]. The SVM model seeks to find the hyperplane that has a maximal distance between two classes [16]. CNN is one of the most widely used deep neural networks with multiple layers, including convolutional layer, non-linearity layer, pooling layer, and fully connected layer. CNN has an excellent performance in machine learning problems, especially in imaging studies [46]. DTs are one of the most popular approaches for representing classifiers, expressed as a recursive partition of the instance space [47]. XGB generates a series of decision trees in sequential order; each decision tree is fitted to the residual between the prediction of the pre-

Table 1
Results for screening models.

Study, Country, Outcome	No. of CPP*	AI methods	Predictors	Val. methods	Performance (AUC, Accuracy (Acc%), Sensitivity (SEN%), Specificity (SPE%), PPV/NPV (%), (95% CI))	Risk of Bias**: Participants/ Predictors/ Outcome/Analysis/ Overall
Yang et al. [4], USA, Early and rapid identification of high-risk SARS-CoV-2 infected patients	1,898	LR, DT, RF, GBDT	age, gender, race and 27 routine laboratory tests	5-FCV	AUC 0.854 (95% CI: 0.829–0.878)	L U H H H
Li et al. [63], China, Screening based on ocular surface features	104	DL	Imaging features	5-FCV	AUC 0.999 (95%CI, 1670.997–1.000, SEN 98.2, SPE 97.8	U U U H H
AS Soltan et al. [3], UK, Early detection, Screening	437	multivariate LR, RF, XGBoost	Presentation laboratory tests and vital signs	TTS, 10-FCV	ED model: AUC 0.939, SEN 77.4, SPE 95.7Admissions model: AUC 0.940, SEN 77.4, SPE 94.8Both models achieve high NPP (>99)	H H H H H
Nan et al. [57], China, Early screening	293	DL, LR, SVM, DT, RF	4 epidemiological features, 6 clinical manifestations (muscle soreness, dyspnea, fatigue, lymphocyte count, WBC, imaging features)	TTS	AUC 0.971, Acc 90, SPE 0.95 (LR optimal screening model)	H U H H H
Soares et al. [58], Brazil, Screening of suspect COVID-19 patients	81	ML, SVM, SMOTE Boost, ensembling, k-NN	Hemogram: (Red blood cells, MCV, MCHC, MCH, RDW, Leukocytes, Basophils, Monocytes, Lymphocytes, Platelets, Mean platelet volume, Creatinine, Potassium, Sodium, CRP, Age	unspecified	AUC 86.78 (95%CI: 85.65–87.90), SEN 70.25 (95% CI: 66.57–73.12), SPE 85.98 (95%CI: 84.94–86.84), NPV 94.92 (95%CI: 94.37–95.37), PPV 44.96 (95% CI: 43.15–46.87)	L U H H H
Feng et al. [59], China, Early identification of suspected COVID-19 pneumonia on admission	32	ML, LR (LASSO), DT, Adaboost	lymphopenia, elevated CRP and elevated IL-6 on admission	10-FCV	AUC 0.841, SPE 72.7	H H H H H
Wu et al. [60], China, Early detection	27	RF	11 key blood indices: TP, GLU, Ca, CK-MB, Mg, BA, TBIL, CREA, LDH, K, PDW	10-FVC, Ext. Val.	Acc 95.95, SEN 95.12, SPE 96.97	L L L H H
Banerjee et al. [61], Brazil, Initial screening	81	RF, ANN	platelets, leukocytes, eosinophils, basophils, lymphocytes, monocytes.	10-FCV	AUC 0.95	H H H H H
Peng et al. [62], China, Quick and accurate diagnosis	32	SRLSR, non-dominated radial slots-based algorithm, ARMED, GFS, RFE	18 diagnostic factors: WBC, eosinophil count, eosinophil ratio, 2019 new Coronavirus RNA (2019n-CoV), Amyloid-A, Neutrophil ratio, basophil ratio, platelet, thrombocytocrit, monocyte count, procalcitonin, neutrophil count, lymphocyte ratio, lymphocyte count, monocyte ratio, MCHC, Urine SG	not performed	not performed	L L U H H

*CPP = COVID-19 Positive Patients, Abbreviations of medical terms included in this Table are provided in the Appendix.
**L: Low, H: High, U: Unclear

Table 2
Results for screening imaging models.

Study, Country, Outcome	No. of CPP*	AI methods	Predictors	Val. methods	Performance (AUC, Accuracy (Acc%), Sensitivity (SEN %), Specificity (SPE%), PPV/NPV (%), (95% CI))	Risk of Bias**: Participants/ Predictors/ Outcome/Analysis/ Overall
Abdani et al. [30], Fast screening	219	DL, CNN	Imaging features	5-FCV	Acc 94	H U H H H
Ahammed et al. [5], Early detection	285	ML, DL, CNN, SVM, RF, k-NN, LR, GNB, BNB, DT, XGB, MLP, NC, perceptron.	Imaging features	10-FCV	AUC 95.52, Acc 94.03, SEN 94.03, SPE 97.01	H H H H H
Barstugan et al. [31], Early detection	53	ML, SVM	Imaging features	10-FCV	Acc 99.68, SEN 93, SPE 100	U U U H H
Wu et al. [55], China, Fast and accurately identification	368	DL	Imaging features	TTS	AUC 0.905, Acc 83.3, SEN 82.3	L U U H H
Wang et al. [56], China, Triage	1647	DL	Imaging features	Ext. val.	AUC 0.953 (95% CI 0.949–0.959), SEN 92.3 (95% CI 91.4–93.2), SPE 85.1 (84.2–86.0), PPV 79 (77.7–80.3), NPV 94.8 (94.1–95.4)	L U U H H

*CPP = COVID-19 Positive Patients, Abbreviations of medical terms included in this Table are provided in the Appendix.
**L: Low, H: High, U: Unclear

Table 3
Results for diagnostic models.

Study, Country, Outcome	No. of CPP*	AI methods	Predictors	Val. methods	Performance (AUC, Accuracy (Acc%), Sensitivity (SEN%), Specificity (SPE%), PPV/NPV (%), (95% CI))	Risk of Bias**: Participants/ Predictors/ Outcome/Analysis/ Overall
Diagnostic						
Cabitza et al. [64], Italy, Fast identification	845	ML	LDH, AST, CRP, calcium, WBC, age	Int.-ext. val.	AUC 0.83–0.90	L H L L H
Batista et al. [8], Brazil, Diagnosis	102	ML, NN, RF, GB trees, LR, SVM	lymphocytes, leukocytes, eosinophils	10-FVC	AUC 0.85, SEN 68, SPE 85, PPV 78, NPV 77	H L H H H
Cai et al. [10], China, predict RT-PCR negativity during clinical treatment	81	DL	9 CT quantitative features and radiomic features	TTS	AUC 0.811–0.812, SEN 76.5, SPE 62.5	H H H H H
Mei et al. [65], China, Diagnosis	419	DCNN	Imaging features, age, exposure to SARS-CoV-2, fever, cough, cough with sputum, WBC	TTS	AUC 0.92, SEN 84.3	H H H L H
Ren et al. [11], China, Diagnosis	58	AI	unclear	unspecified	AUC 0.740, SEN 91.2, SPE 58.8	L U U H H

*CPP = COVID-19 Positive Patients, Abbreviations of medical terms included in this Table are provided in the Appendix.
**L: Low, H: High, U: Unclear

vious decision tree and the target value, and this is repeated until a predetermined number of trees or a convergence criterion is reached [48]. AI techniques based on the above algorithms or other algorithmic approaches for classification purposes that were identified among the reviewed studies included GDCNN (Genetic Deep Learning Convolutional Neural Network) [74], CRT (Classification and Regression Decision Tree) [116], ET (Extra Trees) [15], GBDT (Gradient Boost Decision Tree) [4,71], GBM light (Gradient Boosting Machine light) [100,115], Adaboost (Adaptive Boosting) [27], Boost Ensembling [58,45,59], k-NN (K-Nearest Neighbor) [5,9,58,70,88,106,112], NB (Naïve Bayes) [9], BNB (Bernoulli Naïve Bayes) [5], GNB (Gaussian Naïve Bayes) [5,27], Inception Resnet (Inception Residual Neural Network) [117], LDA (linear discriminant analysis) [43], RBF (Radial Basis Function) [45], LSTM (Long-Short Term Memory) [43].

Moreover, the use of AI techniques for data preprocessing purposes was reported in the reviewed studies. In particular, AI techniques, including ARMED (Attribute Reduction with Multi-objective Decomposition Ensemble optimizer) [62], GFS (Gradient boosted feature selection) [62], MRMR (Maximum Relevance Minimum Redundancy) [84], and RFE (Recursive Feature Elimination)

[62] were used for feature selection. Regression models, including L1LR (L1 Regularized Logistic Regression) [17], LASSO (Least Absolute Shrinkage and Selection Operator) [50,59,84,89,112], CoxPH (Cox Proportional Hazards) [100], and SRLSR (Sparse Rescaled Linear Square Regression) [62] were also employed for feature selection. The use of Nadam optimizer (NesterovAccelerated Adaptive Moment optimizer) for model optimization was reported in two studies [33,117], while application of SMOTE (Synthetic Minority Oversampling TEchnique) for data augmentation was reported in one study [58]. Finally, NLP (Natural Language Processing) was employed in three studies [7,48,86], for data mining purposes.

Regarding novel technologies used in the included studies, explainable AI (XAI) methods, and Federated Learning (FL) were investigated. XAI refers to the ability to explain to a domain expert the reasoning that enables the algorithm to produce its results and is deemed increasingly important in health AI applications [49]. FL, used in [24] and [50], is a nascent field for data-private multi-institutional collaborations, where model-learning leverages all available data without sharing data between hospitals, by distributing the model-training to the data-owners and aggregating their results [51].

Table 4

Results for diagnostic imaging models – part 1.

Study, Country, Outcome	No. of CPP*	AI methods	Predictors	Val. methods	Performance (AUC, Accuracy (Acc%), Sensitivity (SEN%), Specificity (SPE%), PPV/NPV (%), (95% CI))	Risk of Bias**: Participants/ Predictors/ Outcome/Analysis/ Overall				
Chen et al. [92], China, Diagnosis	51	DL	Imaging features	TTS	Acc 95.24, SEN 100, SPE 93.55, PPV 84.62, NPV 100	H	U	L	H	H
Rahimzadeh et al. [33], Diagnosis	118	DNN, Nadam optimizer	Imaging features	TTS	Acc 99.50	L	U	L	H	H
Roy et al. [72], Italy, Diagnosis	17	DL	Imaging biomarkers	5-FCV	F1-score 65.9	H	U	H	H	H
Zhou et al. [73], China, Diagnosis	35	DL	Imaging features	Ext. val.	AUC > 0.93	H	U	U	H	H
Ter-Sarkisov et al. [93], China, Diagnosis	150	DL, CNN	Imaging features	TTS	Acc 91.66, SEN 90.80	H	U	U	H	H
Qjidaa et al. [19], Early detection	unclear	DL, CNN	Imaging features	Int.- ext. val.	Acc 92.5, SEN 92	U	U	U	H	H
Babukarthik et al. [74], Early diagnosis	102	GDCNN	Imaging features	unclear	Acc 98.84, SEN 100, SPE 97.0	H	H	H	H	H
Minaee et al. [20], Diagnosis	unclear	CNN	Raw images without feature extraction	TTS	SEN 98, SPE 92	U	U	U	H	H
Yan et al. [94], China, Diagnosis	206	CNN	Imaging features	TTS	SEN 99.5 (95%CI: 99.3–99.7), SPE 95.6 (95%CI: 94.9–96.2)	L	H	H	H	H
Lokwani et al. [95], India, Diagnosis	55	NN	Imaging features	TTS	SEN 96.4 (95% CI: 88–100), SPE 88.4 (95% CI: 82–94)	U	U	H	H	H
Jin et al. [96], China, Screening (early detection)	751	DL, DNN	Imaging features	TTS	AUC 0.97, SEN 90.19, SPE 95.76	H	U	U	H	H
Ko et al. [29], South Korea, Diagnosis	20	2D DL	Imaging features	TTS, ext.val.	Acc 99.87, SEN 99.58, SPE 100.00	U	U	U	H	H
Ezzat et al. [34], Diagnostic imaging	99	Hybrid CNN	Not applicable	TTS	Acc 98	H	U	H	L	H
Ouchicha et al. [40], Diagnosis	43	DCNN	Imaging features	5-FCV	Acc 97.20	U	U	L	H	H
Xiong et al. [75], China, Diagnosis	521	DL, CNN	Imaging features	TTS, ext. val.	AUC 0.95, Acc 96 (95% CI: 90–98), SEN 95 (95% CI: 83–100), SPE 96 (95% CI: 88–99)	H	U	H	H	H
Li et al. [2], China, Diagnosis	468	DL, CNN	Imaging features	TTS	AUC 0.96, SEN 90 (95% CI: 83–94), SPE 96 (95% CI: 93–98)	L	U	H	H	H
Mahmud et al. [25], China, Diagnosis	unclear	DCNN	Imaging features	5- FCV	Acc 97.4	U	U	U	H	H
Li et al. [97], China, Diagnosis	305	NN	Imaging features	TTS	Precision 93%	U	U	L	H	H
Sun et al. [98], China, Diagnosis	1495	LR, SVM, RF, NN	30 Imaging features: Volume features, Infected lesion number, Histogram distribution, Surface area, Radiomics features	5-FCV	ACC 91.79, SEN 93.05, SPE 89.95	L	U	H	H	H

Results for diagnostic imaging models – part 2 Study, Country, Outcome	No. of CPP*	AI methods	Predictors	Val. methods	Performance (AUC, Accuracy (Acc%), Sensitivity (SEN%), Specificity (SPE%), PPV/NPV (%), (95% CI))	Risk of Bias**: Participants/ Predictors/ Outcome/ Analysis/ Overall				
Ozturk et al. [35], Diagnosis (automatic COVID-19 detection)	125	DNN	Imaging features (1. Ground-glass opacities (GGO) (bilateral, multifocal, subpleural, peripheral, posterior, medial and basal) 2. A crazy paving appearance (GGOs and inter-/intra-lobular septal thickening) 3. Air space consolidation. 4. Broncho vascular thickening (in the lesion). 5. Traction bronchiectasis	5- FCV	Acc 98.08, SEN 95.13, SPE 95.03, F1-score 96.51	H		U	H	H	H
Zhang et al. [7], United States, Diagnosis	2060	DNN (CV19-Net), NLP	Imaging features	TTS	AUC 0.92 (95% CI: 0.91–0.93), SEN 88% (95% CI: 87–89), SPE 79% (95% CI: 77–80)	H		U	H	H	H
Borkowski et al. [37], Diagnosis	103	Automated ML platform Microsoft CustomVision	Imaging features	Ext. val. with US Dep. of Vet. Affairs (VA) PACS	Acc 97, SEN 100, SPE 95, PPV 91, NPV 100	H		H	H	H	H
Xu et al. [6], China, Diagnosis	432	DL, CNN, FL	Imaging features	5-FCV	SEN 77.2, SPE 91.9	L		U	U	H	H

Table 4 (*continued*)

Study, Country, Outcome	No. of CPP*	AI methods	Predictors	Val. methods	Performance (AUC, Accuracy (Acc%), Sensitivity (SEN%), Specificity (SPE%), PPV/NPV (%), (95% CI))	Risk of Bias**: Participants/ Predictors/ Outcome/ Analysis/ Overall				
Wang et al [99], Diagnosis	266	DCNN	Imaging features	Int. val.	Acc 93.3, SEN 91.0, PPV 98.9	L	U	U	H	H
Elaziz et al. [21], Diagnosis	unclear	CNN	Imaging features	Ext. val.	Acc 98.09	H	U	U	L	H
Gozes et al. [26], United States, China, Diagnostic imaging	unclear	DCNN	Imaging features	TTS	AUC 0.996, (95%CI: 0.989–1.00), SEN 98.2, SPE 92.2	U	U	H	H	H
Salman et al. [22], Diagnosis	unclear	DL, CNN	Imaging features	TTS	Acc 100, SEN 100, SPE 100, PPV 100, NPV 100	U	U	U	H	H
Liu et al. [24], Diagnosis	unclear	FL (Federated Learning)	Imaging features	Comparison with and without FL 10-	ResNet18 (highest accuracy): Acc 91,26, SEN 96.15, SPE 91.26	U	U	L	L	H
Castiglioni et al. [76], Italy, Diagnosis	250	CNN	unclear							
FCV	AUC 0.89 (95%CI 0.86–0.91), SEN 78 (95% CI 74–81), SPE 82 (95% CI: 78–85)	H	U	H	H	H				
Padma et al. [23], Diagnosis	unclear	ML, CNN	Imaging features	TTS	Acc 99	U	U	U	H	H
Zhang et al. [100], China, Diagnosis	752	Light GBM, CoxPH	Imaging features	5-FCV	AUC 0.9797 (95% CI: 0.966–0.9904), Acc 92.49, SEN 94.93, SPE 91.13	U	U	U	H	H
Wang et al. [77], China, Diagnosis	79	DL, CNN	Imaging features	Int. – ext.						
val.	Acc 89.5, SEN 87, SPE 88	H	U	U	H	H				

*CPP = COVID-19 Positive Patients, Abbreviations of medical terms included in this Table are provided in the Appendix.

**L: Low, H: High, U: Unclear

C. Validation methods

The most frequent validation methods used in the included studies were training test split (TTS) (34 studies), 5-fold cross-validation (FCV) (23 studies) and 10-fold cross-validation (19 studies). Other studies performed internal and external validation methods. Internal validation techniques, are advocated when no other data than the study sample are being used, to estimate the potential of overfitting in the performance of the developed model [52]. External validation is used to adjust or update the model in other data than the study sample [52]. TTS validation method splits the data into training and testing datasets based on a predefined threshold. FCV is the traditional method for estimating the future error rate of a prediction rule constructed from a training set of data [53]. To estimate the potential for overfitting and optimism in a prediction's model performance, internal validation techniques are advocated, meaning that no other data than the study sample are being used [52]. External validation uses new participant level data, external to those used for model development, to examine whether the model's predictions are reliable and ade-

quately accurate, in individuals from potential population for clinical use [54]. All validation methods used in the included studies, are reported in Tables 1-7.

D. Models developed for screening purposes

We identified 14 models for screening COVID-19 (see Tables 1 and 2). Five screening models were based in medical imaging [5,30,31,55,56] using CXRs and CT scans from public databases or hospitals in China. Wang *et al.* [56] were the first to develop and validate a deep learning algorithm on the basis of chest CT scans of 1647 COVID-19 positive patients acquired from fever clinics of five hospitals in Wuhan, China, for rapid triaging, achieving AUC 0.953 (95% CI 0.949–0.959), SEN 0.923 (95% CI 0.914–0.932), SPE 0.851 (0.842–0.860), PPV 0.790 (0.777–0.803) and NPV 0.948 (0.941–0.954). The rest of the screening studies [4,3,57,58,59,60,61,62] used as input data -among other- demographic data, comorbidities, epidemiological history of exposure to COVID-19, vital signs, blood test values, clinical symptoms, infection-related biomarkers, and days from illness onset to first

Table 5
Results for prognostic models – part 1.

Study, Country, Outcome	No. of CPP*	AI methods	Predictors	Val. methods	Performance (AUC, Accuracy (Acc%), Sensitivity (SEN%), Specificity (SPE%), PPV/NPV (%), (95% CI))	Risk of Bias**: Participants/ Predictors/ Outcome/Analysis/ Overall				
Muhammad et al. [9], South Korea, Recovery prediction, disease progression	unclear	DT, SVM, NB, LR, RF, K-NN	unclear	5-FCV	Acc 99.85 (Decision Tree)	H	U	U	U	H
Cheng et al. [78], United States, Severity Assessment, (risk prioritization tool that predicts ICU transfer within 24 h)	1987	RF	respiratory failure, shock, inflammation, renal failure	TTS, 10-FCV	AUC 79.9 (95% CI: 75.2–84.6), Acc 76.2 (95% CI: 74.6–77.7), SEN 72.8 (95% CI: 63.2–81.1), SPE 76.3% (95% CI: 74.7–77.9)	H	H	H	H	H
Kim et al. [79], South Korea, ICU need prediction	4787	55 ML models developed, (XGBoost model revealed the highest discrimination perf.)	age, sex, smoking history, body temperature, underlying comorbidities, activities of daily living (ADL), symptoms	TTS	AUC 0.897, (95% CI 0.877–0.917)	H	U	H	U	H
Yadaw et al. [101], United States, Mortality prediction	4802	ML, RF, LR, SVM, XGBoost	age, minimum oxygen saturation over the course of their medical encounter, type of patient encounter (inpatient vs outpatient and telehealth visits)	TTS	AUC 91	L	H	H	H	H
Klann et al. [102], USA, France, Italy, Germany, Singapore, Severity assessment	4227	ML	PaCO2, PaO2, ARDS, sedatives, d-dimer, immature granulocytes, albumin, chlorhexidine, glycopyrrolate, palliative care encounter	5-FCV, TTS	AUC 0.956 (95% CI: 0.952, 0.959)	U	U	U	H	H
Navlakha et al. [103], United States, Severity assessment in cancer patients (predicting severity occurring after 3 days)	354	ML, RF, DT	40 out of 267 clinical variables (3 most important individual lab variables: platelets, ferritin, and AST (aspartate aminotransferase)	10-FCV	AUC 70–85	L	H	H	H	H
Shashikumar et al. [104], United States, Mechanical ventilation need prediction (24 h in advance)	777	DL	vital signs, laboratory values, sequential-organ failure assessment (SOFA) scores, Charlson comorbidity index scores (CCI) index, demographics, length of stay, outcomes	Ext. val., 10-FCV	AUC 0.918	L	H	H	L	H
Bertsimas et al. [80], Greece, Italy, Spain, United States, Mortality risk	3,927	XGBoost	Increased age, decreased oxygen saturation (<93%), elevated levels of CRP (>130 mg/L), blood urea nitrogen, blood creatinine	Cross-validation	AUC 0.90 (95% CI: 0.87–0.94)	H	H	U	L	H

Results for prognostic models – part 2

Study, Country, Outcome	No. of CPP*	AI methods	Predictors	Val. methods	Performance (AUC, Accuracy (Acc%), Sensitivity (SEN%), Specificity (SPE%), PPV/NPV (%), (95% CI))	Risk of Bias**: Participants/ Predictors/ Outcome/Analysis/ Overall				
Youssef et al. [105], UK, Severity assessment (ventilation, ICU need prediction), early warning system	472	GBT, LR, RF	Vital signs: Heart Rate, Oxygen Saturation, Respiratory Rate, Systolic Blood Pressure Temperature, AVPUVenous blood tests: Albumin, ALK. Phosphatase, ALT, APTT, Basophils, Bilirubin Creatinine, CRP, Eosinophils, Haemocrit, Haemoglobin, INR Lymphocytes, Mean Cell Vol., Monocytes, Neutrophils Platelets, Potassium, Prothrombin Time, Sodium, Urea, WCs, eGFR	5-FCV	AUC 0.94 (best performance, GBT)	U	U	H	H	H
Vaid et al. [50], United States, Mortality Prediction	4029	FL, LR, federated LASSO, federated MLP	gender, age, ethnicity, race, past medical history (such as asthma), lab tests (white blood cell counts)	10-FCV	AUC 0.693–0.805	L	L	H	H	H
Karthikeyan et al. [106], China, Mortality prediction	370	ML, XGBoost, k-NN, LR, RF, DT	neutrophils, lymphocytes, LDH, hs-CRP, age	5-FCV	Acc 96 (16 days in advance)	H	U	U	U	H

Table 5 (continued)

Results for prognostic models – part 2

Study, Country, Outcome	No. of CPP	AI methods	Predictors	Val. methods	Performance (AUC, Accuracy (Acc%), Sensitivity (SEN%), Specificity (SPE%), PPV/NPV (%), (95% CI))	Risk of Bias**: Participants/ Predictors/ Outcome/Analysis/ Overall				
Jimenez-Solem et al. [81], Denmark, UK, Severity assessment (ICU need, mechanical ventilation need), disease progression, mortality prediction	3944	ML, RF	age, gender, BMI, hypertension, diabetes	Ext. val. (UK biobank), 5-FCV	AUC 0.678–0.818 (hospital admission), AUC 0.587–0.736 (ICU admission), AUC 0.625–0.674 (mortality prediction)	H	H	H	U	H
Casiraghi et al. [107], Italy, Risk prediction	301	ML, RF	saturation values, lab-oratory values (lymphocyte counts, CRP, WBC counts, Haemoglobin), variables related to comorbidities (number of comorbidities, presence of cardiovascular pathologies and/or arterial hypertension), radiological values computed through CovidNet, and presence of symptoms (vomiting/nausea or dyspnea or respiratory failure).	10-FCV	AUC 0.81–0.76, Acc 74–68, SEN 72–66, SPE 76–71, F1 score 62–55	L	U	U	H	H
Burdick et al. [82], United States, Ventilation need prediction within 24 h and triage	197	ML, XGBoost	diastolic blood pressure (DBP), systolic blood pressure (SBP), heart rate (HR), temperature, respiratory rate (RR), oxygen saturation (SpO2), white blood cell (WBC), platelet count, lactate, blood urea nitrogen (BUN), creatinine, and bilirubin	Ext. val.	AUC 0.866, SEN 90, SPE 58.3	H	U	H	H	H

Results for prognostic models – part 3

Study, Country, Outcome	No. of CPP*	AI methods	Predictors	Val. methods	Performance (AUC, Accuracy (Acc%), Sensitivity (SEN%), Specificity (SPE%), PPV/NPV (%), (95% CI))	Risk of Bias**: Participants/ Predictors/ Outcome/Analysis/ Overall				
Cai et al. [83], China, Severity assessment, ICU need and length of stay prediction, O₂ inhalation duration prediction, sputum NAT-positive prediction and patient prognosis	99	ML, RF	CT quantification	10-FCV	AUC 0.945 (ICU treatment), AUC 0.960 (prognosis/ partial recovery vs pro-longed recovery)	H	U	U	H	H
Wu et al. [84], China, Mortality prediction	58	ML, LR, mRMR, LASSO LR	7 continuous laboratory variables: blood routine test, serum biochemical (including glucose, renal and liver function, creatine kinase, lactate dehydrogenase, and electrolytes), coagulation profile, cytokine test, markers of myocardial injury, infection-related makers, other enzymes	5-FCV	SEN 98 (95% CI: 93–100), SPE 91 (95% CI: 84–99)	H	L	L	H	H
Iwendi et al. [27], Severity and outcome prediction	unclear	ML, DT, SVM, GNB, Boosted RF, AdaBoost	COVID-19 patient's geographical, travel, health, and demographic data	unspecified	Acc 94	H	U	H	H	H
Gerevini et al. [15], Italy, Mortality prediction	unclear	DT, RF, ET (extra trees)	Age, sex, C-Reactive Protein (PCR), Lactate dehydrogenase (LDH), Ferritin (Male)Ferritin (Female), Troponin-T, White blood cell (WBC), D-dimer, Fibrinogen, Lymphocyte (over 18 years old patients), Neutrophils/Lymphocytes, Chest XRay-Score (RX)	Cross validation	AUC 90.2 for the 10th day	H	U	U	L	H
Yan et al. [108], China, Severity assessment	404	ML, XGBoost	3 biomarkers that predict the survival of individual patients: LDH, lymphocyte, high-sensitivity C-reactive protein (hs-CRP).	cross-validation	Acc 90	L	L	U	H	H
Gao et al. [71], China, Mortality prediction	2160	LR, SVM, GBDT, NN	8 features had a positive association with mortality (high risk: consciousness, male sex, sputum, blood urea nitrogen [BUN], respiratory rate [RR], D- dimer, number of comorbidities, and age) and 6 features were negatively correlated with mortality (low risk: platelet count[PLT], fever, albumin [ALB], SpO2, lymphocyte, and chronic kidney disease [CKD]).	Int. val. cohort	AUC 0.9621 (95% CI: 0.9464–0.9778)	H	H	L	L	H

Table 5 (continued)

Results for prognostic models – part 3										
Ma et al. [85], China, Mortality prediction	305	ML, RF, XG Boost	LDH, CRP, age		4-FCV	AUC 0.951	H	L	U	H	H

Results for prognostic models – part 4										
Study, Country, Outcome	No. of CPP*	AI methods	Predictors	Val. methods	Performance (AUC, Accuracy (Acc%), Sensitivity (SEN%), Specificity (SPE%), PPV/NPV (%), (95% CI))		Risk of Bias**: Participants/ Predictors/ Outcome/Analysis/ Overall			
Schwab et al. [109], Brazil, Prognosis for Hospitalization need, ICU need	558	LR, NN, RF, SVM, XGBoost	Predicting SARS-CoV-2 test results (first 3): presence of arterial lactic acid measurement, age, urine-leukocyte countPredicting Hospitalization: lactic dehydrogenase, gamma-glutamyl transferase, INR presencePredicting ICU need: pCO2, potassium presence, ionized calcium presence	TTS	AUC 0.92, (95% CI 0.81–0.98), sensitivity 75 (95% CI 67–81), specificity 49 (95% CI 46–51) for COVID-19 positive patients that require hospitalizationAUC 0.98 (95% CI 0.95–1.00) for COVID-19 positive patients that require critical care		L H H H H
Izquierdo et al. [86], Spain, ICU need prediction	10,504	ML, NLP	age, fever, tachypnea	TTS	AUC 0.76, Acc 0.68		H U U U H
Hao et al. [48], United States, Early prediction of level-of-care requirements (hospitalization/ ICU/ mechanical ventilation need prediction,	2566	NLP, RF, XGBoost, SVM, LR	Vital signs, age, BMI, dyspnea, comorbidities (Most important for hospitalization), Opacities on chest imaging, age, admission vital signs and symptoms, male gender, admission laboratory results, diabetes (most important risk factors for ICU admission and mechanical ventilation)	TTS	Acc 88 (hospitalization need), Acc 87 (ICU care), Acc 86 (mechanical ventilation need)		U L L L U
Wu et al. [87], China, Severity assessment, triage	299	LR	1) Clinical features: age, hospital employment, body temperature and the time of onset to admission 2) Laboratory features: Lymphocyte (proportion), neutrophil, (proportion), CRP, lactate dehydrogenase (LDH), creatine kinase, urea and calcium3) CT semantic features: lesion-Most prominent predictors: age, lymphocyte (proportion), CRP, LDH, creatine kinase, urea and calcium	5-FCV	AUC 0.84–0.93, Acc 74.4–87.5, SEN 75.0–96.9, SPE 55.0–88.0		H H H H H
Li et al. [16], Fatality prediction	30,406	DL, LR, RF, SVM	symptoms and comorbidities	TTS	Acc> 90, SPE> 90, SEN< 40		H H H H H
Nemati et al. [110], Predict patient length of stay in the hospital	1182	SVM, GB, DT	age, gender		Acc 71.47		U H U U H
Zhu et al. [111], China, Mortality prediction	181	DL, DNN	top predictors: D-dimer, O2 Index, neutrophil: lymphocyte ratio, C-reactive protein, lactate dehydrogenase	5-FCV	AUC 0.968 (95%CI: 0.87–1.0)		L H U L H
Vaid et al. [112], United States, Mortality and critical events prediction	4098	XGBoost, (LR, LASSO, k-NN for validation)	Within 7 days of admission most prominent predictors are:1) For critical event prediction: presence of acute kidney injury, high and low levels of LDH, respiratory rate, glucose2) For mortality prediction: age, anion gap, C-reactive protein, LDH	10-FCV, Ext. val.	1) Critical event prediction: AUC 0.80 (3 days), 0.79 (5 days), 0.80 (7 days), 0.81 (10 days)2) Mortality prediction: AUC 0.88 (3 days), 0.86 (5 days), 0.86 (7 days), 0.84 (10 days)		L H H H H

Results for prognostic models – part 5										
Study, Country, Outcome	No. of CPP*	AI methods	Predictors	Val. methods	Performance (AUC, Accuracy (Acc%), Sensitivity (SEN%), Specificity (SPE%), PPV/NPV (%), (95% CI))		Risk of Bias**: Participants/ Predictors/ Outcome/Analysis/ Overall			
Wollenstein-Betech et al. [113], Mexico, predict :(1) hospitalization, (2) mortality (3) ICU need (4) ventilator need	20,737	SVM, LR, RF, XGBoost	age, gender, diabetes, COPD, asthma, immunosuppression, hypertension, obesity, pregnancy (hospitalization need prediction), chronic renal failure, tobacco use, other disease, SARS-CoV-2 test results	10-FCV	AUC 0.63, Acc 79 (mortality prediction performance), AUC 0.74 (hospitalization prediction performance), AUC 0.55, Acc 89 (ICU need prediction performance), AUC 0.58, Acc 90 (ventilator need prediction performance)		L H H H H
Wang et al. [17], China, Mortality prediction of hospitalized COVID-19 patients	375	L1LR, L1SVM	LDH, lymphocytes percentage, hs-CRP, Albumin	TTS	Acc> 94, F1-score 97		H L U H H

Table 5 (continued)

Results for prognostic models – part 5

Study, Country, Outcome	No. of CPP*	AI methods	Predictors	Val. methods	Performance (AUC, Accuracy (Acc %), Sensitivity (SEN%), Specificity (SPE%), PPV/NPV (%), (95% CI))	Risk of Bias**
Jiang et al. [88], China, Severity assessment, Outcome Prediction	53	ML, LR, k-NN, DT, RF, SVM	Most predictive features: alanine aminotransferase (ALT), myalgias, hemoglobin (red blood cells) Less predictive features: lymphocyte count, white blood count, temperature, cycle threshold, creatinine, gender, CRP, age, fever, CK, LDH, Glu, AST, K+, N+	10-FCV	Acc 70–80	H U H H H
Sun et al. [114], China, Severity assessment	336	SVM	36 clinical indicators (mainly thyroxine, immune related cells and products including CD3, CD4, CD19, CRP, high-sensitive CRP, leukomonocytes and neutrophils), age, GSH, total protein	TTS	AUC 0.999, SEN 77.5, SPE 78.4	L H L H H
Razavian e al. [115], United States, Outcome prediction	3,317	LR, RF, Light GBM	significant oxygen support (including nasal cannula at flow rates>6 L/min, face mask or high-flow device, or ventilator), ICU admission, death (or discharge to hospice), return to the hospital after discharge within 96h of prediction	TTS	AUC 90.8, (95% CI: 90.7–91.0)	L U H L H
Liang et al. [89], China, Early triage of critically ill (Critical illness was defined as a composite event of admission to an ICU or requiring invasive ventilation, or death.)	1590	DL, LASSO algorithm	10 clinical features: X-ray abnormalities, age, dyspnea, COPD (chronic obstructive pulmonary disease), number of comorbidities, cancer history, neutrophil/lymphocytes ratio, lactate dehydrogenase, direct bilirubin, creatine kinase	Three int. ext. val.:1) Hospital in Wuhan2) Centers in Hubei3) Hospital in Guangdong	AUC 0.911 (95% CI, 0.875–0.945)	H H U L H
Gong et al. [90], China, Early identification of cases at high risk of progression to severe COVID-19	72	LR, DT, RF, SVM	Age, DBIL, RDW, BUN, CRP, LDH, and ALB	Ext. val.	AUC 0.912, (95% CI: 0.846–0.978), SEN 85.7, SPE 87.6	H H H H H

Results for prognostic models – part 6

Study, Country, Outcome	No. of CPP*	AI methods	Predictors	Val. methods	Performance (AUC, Accuracy (Acc %), Sensitivity (SEN%), Specificity (SPE%), PPV/NPV (%), (95% CI))	Risk of Bias**: Participants/ Predictors/ Outcome/Analysis/ Overall
Pourhomayoun et al. [70], 76 countries, Mortality prediction	117,000	SVM, ANN, RF DT, LR, k-NN	demographic features (age, sex, province, country, age, travel history), comorbidities(diabetes, cardiovascular disease), patient symptoms (chest pain, chills, colds, conjunctivitis, cough, diarrhea, discomfort, dizziness, dry cough, dyspnea, emesis, expectoration, eye irritation, fatigue, gasp, headache, lesions on chest radiographs, little sputum, malaise, muscle pain, myalgia, obnubilation, pneumonia, myelofibrosis, respiratory symptoms, rhinorrhea, somnolence, sputum, transient fatigue, weakness)	10-FCV	Acc 93.75, SEN 90, SPE 97	U L U H H
Abdulaal et al. [91], UK, Mortality risk (point-of-admission scoring system)	398	ANN	demographics, comorbidities, smoking history, presenting symptoms	Cross validation	AUC 90.12, Acc 86.25, SEN 87.50 (95% CI: 61.65–98.45), SPE 85.94 (95% CI: 74.98–93.36), PPV 60.87 (95% CI: 45.23–74.56), NPV 96.49 (95% CI: 88.23–99.02)	H L H H H
Assaf et al. [116], Israel, Severity assessment risk prediction model	162	ML, NN, RF, CRT	clinical, hematological and biochemical parameters at admission	10-FCV	Acc 92.0, SEN 88.0, SPE 92.7	L H H H H
Fang et al. [43], China, Early warning to predict malignant progression	104	linear discriminant analysis (LDA), SVM, multilayer perceptron (MLP), LSTM	Troponin, Brain natriuretic peptide, White cell count, Aspartate aminotransferase, Creatinine, and Hypersensitive C-reactive protein.	5-FCV	AUC 0.920, (95% CI: 0.861, 0.979)	H H H H H

Table 5 (continued)

Results for prognostic models – part 6

Study	No.	AI method	Predictors	Val.	Performance	Risk of Bias
Trivedi et al. [117], Mortality prediction	739	DL, InceptionResnetV2 model, Nadam optimizer	Imaging features	TTS	Acc 95.3, AUC 96.0, SEN 94.0, SPE, 94.0, F1-score 93.3	L U U H H
Ryan et al. [118], United States, Mortality prediction for COVID-19 pneumonia, and mechanically ventilated patients	114	ML, XGBoost	Age, Heart Rate, Respiratory Rate, Peripheral Oxygen Saturation (SpO2), Temperature, Systolic Blood Pressure, Diastolic Blood Pressure, WBCs, Platelets, Lactate, Creatinine, and Bilirubin, over an interval of 3h and their corresponding differentials in that interval	5-FCV	1) For mortality prediction on mechanically ventilated patients: AUC 0.82 (12h), 0.81 (24h), 0.77 (48h), 0.75 (72h)2) For mortality prediction on pneumonia patientsAUC 0.87 (12h), 0.78 (24h), 0.77 (48h), 0.73 (72h)	L H H H H

*CPP = COVID-19 Positive Patients, Abbreviations of medical terms included in this Table are provided in the Appendix.
**L: Low, H: High, U: Unclear

Table 6
Results for Prognostic Imaging models.

Study, Country, Outcome	No. of CPP*	AI methods	Predictors	Val. methods	Performance (AUC, Accuracy (Acc %), Sensitivity (SEN%), Specificity (SPE%), PPV/NPV (%), (95% CI))	Risk of Bias**: Participants/ Predictors/ Outcome/Analysis/ Overall
Fakhfakh et al. [42], Prognosis	42	RNN, CNN	Unclear	unspecified	Acc 92	H U U L H
Zhu et al. [66], China, Disease progression prediction	408	SVM, LR	Imaging features	5-FCV	Acc 85.91	L U U H H
Qi et al. [67], China, Hospital stay prediction (Short-term (<10 days), long-term (>10 days))	31	LR, RF	Imaging features (CT radiomics)	5-FCV	AUC 0.97, SEN 100, SPE 89, (95%CI 0.83–1.0)	U L L H H
Xiao et al. [68], China, Severity assessment, disease progression	408	DL, CNN, ResNet34 (RNN)	Imaging features	5-FCV	AUC 0.987 (95% CI: 0.968–1.00), Acc 97.4	L U U H H
Cohen et al. [36], Severity assessment for COVID19 Pneumonia	80	NN	CXR features	not performed		U U U H H
Salvatore et al. [69], Italy, Prognosis prediction (discharging at home, hospitalization in stable conditions, hospitalization in critical conditions, death)	98	LR	Imaging features	not performed	Acc 81, SEN 88, SPE 78	H U U H H
Liu et al. [44], China, Severity assessment	134	CNN	Imaging features (APACHE-II, NLR, d-dimer level)	Ext. val.	AUC 0.93, (95% CI: 0.87–0.99)	L U U U U

*CPP = COVID-19 Positive Patients, Abbreviations of medical terms included in this Table are provided in the Appendix.
**L: Low, H: High, U: Unclear

Table 7
Results for diagnostic and prognostic Imaging models.

Study, Country, Outcome	No. of CPP*	AI methods	Predictors	Val. methods	Performance (AUC, Accuracy (Acc%), Sensitivity (SEN%), Specificity (SPE%), PPV/NPV (%), (95% CI))	Risk of Bias**: Participants/ Predictors/ Outcome/ Analysis/Overall
Chassagnon et al. [45], France, Quantification, Staging and Prognosis of COVID-19 Pneumonia	693	DL, 2D-3D CNN, RBF SVM, Linear SVM, AdaBoost, RF, DT, XGBoost	15 radiomics features: imaging from the disease regions (5features), lung regions (5features) and heart features (5features), biological and clinical data (6features: age, sex, high blood pressure (HBP), diabetes, lymphocyte count, CRP level), image indexes (2features: disease extent and fat ratio).	TTS	Acc 70, SEN 64, SPE 77 (Holistic Multi-Omics Profiling & Staging), Acc 71, SEN 74, SPE 82 (AI prognosis model performance)	L U L L U

*CPP = COVID-19 Positive Patients, Abbreviations of medical terms included in this Table are provided in the Appendix.
**L: Low, H: High, U: Unclear

admission. One study only [63], used ocular surface photographs (eye-region images) as input data demonstrating that asymptomatic and mild COVID-19 patients have distinguished ocular features from others.

In screening studies, the most prominent predictors were age, platelets, leukocytes, monocytes, eosinophils, lymphocytes, CRP, WBC from routine blood tests and imaging features from CXR and CT images. Only one study [57] reported as predictors 4 epidemiological features (relationship with a cluster outbreak, travel or residence history over the past 14 days in Wuhan, exposure to patients with fever or respiratory symptoms over the past 14 days who had a travel or residence history in Wuhan, exposure to

patients with fever or respiratory symptoms over the past 14 days who had a travel or residence history in other areas with persistent local transmission, or community with definite cases) and 6 clinical manifestations (muscle soreness, dyspnea, fatigue, lymphocyte count, white blood cell count, imaging changes of Chest X-ray or CT).

E. Diagnostic prediction models

Thirty-eight diagnostic prediction models for detecting COVID-19 were identified out of which 32 were based on medical imaging using CXRs or CT scans as input data. Results are presented in Tables 3 and 4. The rest of the diagnostic studies [64,8,9,10,65,11,45] used as input, among other data, age, gender, demographics, symptoms, routine blood exam results, and clinical characteristics. The largest dataset of COVID-19 positive patients (CPP) used in the included diagnostic imaging studies was 2060 patients (5806 CXRs; mean age 62 ± 16, 1059 men) [7]. In this retrospective study, a deep neural network, CV19-Net, was trained, validated, and tested on CXRs, to differentiate COVID-19 related pneumonia from other types of pneumonia, achieving AUC 0.92 (95% confidence interval [CI]: 0.91, 0.93), SEN 88% (95% CI: 87%, 89%) and SPE 79% (95% CI: 77%, 80%). In the non-imaging diagnostic studies, the largest number of CPPs was 845 patients admitted at an Italian hospital from February to May 2020 [64]. Routine blood tests of 1,624 patients were exploited in this study to develop Machine Learning models to diagnose COVID-19 patients, achieving an AUC ranged from 0.83 to 0.90. The most prominent predictors were age, WBC, LDH, AST, CRP and calcium [64].

Most frequently reported predictors in the included studies for identification or diagnosis of COVID-19 cases were age, lymphocytes, WBC and quantitative and radiomic features derived from CXR and CT images.

F. Prognostic prediction models

We identified 50 prognostic models (7 based on medical imaging) [36,42,66,67,68,69,44] for predicting hospitalization need (8 studies), ICU need (10 studies), ventilator need (8 studies), mortality risk (17 studies), severity assessment (16 studies), recovery prediction or disease progression (9 studies) or hospital length stay (3 studies). The results are presented in Table 5 and Table 6. Table 7 presents the results for one diagnostic and prognostic model. The first study to jointly predict the disease progression and the conversion time, which could help clinicians to deal with the potential severe cases in time or even save the patients' lives, used Support Vector Machines (SVM) and Linear Regression (LR) methods on 408 chest CT scans from COVID-19 positive patients [66]. The largest dataset of COVID-19 positive patients included in the prognostic studies was 117,000 patients worldwide. In this study, Support Vector Machine, Artificial Neural Network, Random Forest, Decision Tree, Logistic Regression, and K-Nearest Neighbor were used to predict the mortality rate based on the patients physiological conditions, symptoms, and demographic information [70].

Most of the prognostic prediction models did not report how many days in advance they can produce predictions. Only fourteen studies reported predictions with a time range varying from 12 h to 20 days before the outcome. Highest time frame prediction prior to the outcome was reported by Gao et al. [71] who presented a mortality risk prediction model for COVID-19 (MRPMC) that uses patients clinical data on admission to stratify them by mortality risk, which enables prediction of physiological deterioration and death up to 20 days in advance.

Most frequently reported predictors for prognosis of COVID-19 cases were age, CRP, lymphocyte, LDH and imaging features derived from CXR and CT images.

G. Risk of Bias assessment

The results of the RoB analysis (H: High, L: Low, U: Unclear) are provided for each included study in Tables 1-8. In total 98 prediction models were at high overall risk of bias and three studies [44,45,48] were at unclear risk of bias, indicating that further research needs to be conducted to fully apply them in the clinical practice. In the participants domain, 30 studies had high risk of bias and 21 unclear. Sources of bias in the participants domain varied from small or incomplete datasets to exclusion criteria indicating the need of further data collection to test the generalizability of the developed AI models to other patient populations [2,7,15,16,17,35,37,43,48,57,59,65,69,71–91]. Thirty studies had high RoB in the "predictors" domain related to different ways of definitions and assessment for all participants or predictors availability.

Finally, RoB analysis was high for 99 studies and unclear for two studies [44,48]. Most frequent reasons for assessing high analysis using the PROBAST protocol were number of participants, missing data on predictors and outcomes and exclusion criteria for participants reported as limitations in the included studies.

4. Discussion

In the present systematic review, we included 101 studies that developed or validated screening, diagnostic and prognostic prediction models that can be applied in clinical practice and were published from January 1st, 2020 to December 30th, 2020. Even though most of the studies reported high performance algorithms, results of the RoB analysis conducted in the present review indicate that application in clinical practice may be problematic. Limitations related to the applicability of the developed prediction models were reported by several studies. The most prominent limitation reported was the use of a single data source (one hospital from one geographical area) for the algorithm's training [4,10,11,15,17,30,56,58,59,62,63,64,69,44,76,83,84,91, 92,78,103,107,106,108,109,112,114,115,116,118]. Generalizability of the trained models can be enhanced by adding multiple data sources in future studies.

Concerning the results of the present review, a clear distinction between prediction models that relied on imaging features and models that relied on clinical or laboratory data, proved to be evident amongst the included studies. In particular, medical imaging studies were most prominent for diagnostic purposes (31 out of 37) and least prominent for prognostic purposes (7 out of 50). Models developed for screening purposes were rather few (14) compared to the other two categories and distinction between the use of medical images and clinical or laboratory data was not clear (5 and 9 respectively). Analysis regarding the handling of unbalanced data and application of appropriate performance metrics across the developed models was not feasible due to the fact that different validation methods were applied. Student's *t*-test was deployed to investigate differences in dataset sizes among included studies, in terms of total number of participants and COVID-19 Positive Patients. Such comparison between datasets employed in screening, diagnostic and prognostic models did not display statistically significant differences (p > 0.05). Average number of total participants included in datasets of non-imaging medical studies proved to be significantly higher than in datasets of medical imaging studies (11,033.45 versus 1,528.27, p = 0.04). This was not identified for the number of COVID-19 Positive Patients between these two categories (p > 0.05). Regarding the application of performance criteria for evaluation purposes in the included studies, some dissimilarities were observed between models developed for different purposes. In particular, AUC score was

the most prevalent metric in models developed for prognostic (69.38%) and screening (71.14%) purposes, but not for diagnostic (35.12%) purposes. The most widely used metric in models for diagnostic purposes was SEN (70.27%), followed by SPE (59.45%) and Acc (54.05%). Dissimilarities among performance criteria were also evident between models in medical imaging studies, with Acc (63.63%) and SEN (63.63%) being the most prevalent, and models in non-imaging medical studies, with AUC score (76.78%) being the most common. Comparison between different types of prediction models that employ heterogeneous predictors is exceedingly difficult, taking into consideration that performance criteria applied were not similar. Additionally, since most prediction models were trained on a specific localized dataset, the evaluation of AI techniques used and the importance of predictors cannot be discerned through *meta*-analysis of the results presented in each study. This heavily underlines the need for global collaboration and data sharing, to enable the development of validated benchmarks for the evaluation of newly introduced AI techniques. This would expedite the application of new prediction models in clinical practice, especially in times of extreme urgency such as the COVID-19 pandemic.

Concerning the type of multimodal input data used in the developed models, three studies [43,45,44] demonstrated the advantages of using the combination of clinical features and image features (CT scans, CXR images), indicating that both CT scans and clinical data are of paramount importance to the diagnosis and prognosis of COVID-19. Moreover, developing AI diagnostic and prognostic models in an end-to-end manner enables the use of raw data without the need for manual design of feature patterns or interference of clinicians. Therefore, future AI studies can explore more methods of fusing clinical and image features as well as developing end-to-end models for use in the clinical practice.

Regarding novel technologies, we investigated the use of explainable AI and Federated Learning in the developed prediction models. Evaluation of the AI methods presented in this review study in terms of explainability proved to be difficult, due to the lack of uniform adoption of interpretability assessment criteria across the research community [119]. A multidisciplinary approach of combining medical expertise and data science engineering in future studies might be necessary to overcome this difficulty [119]. The integration of explainability modalities in the developed models can enhance human understanding on the reasoning process, maximize transparency and embellish trust towards the models' use in clinical practice [120]. Therefore, XAI techniques could prove to be critical in tackling volatile crises like the COVID-19 pandemic and as such, it is of paramount importance that XAI should be taken into consideration in future works. Moreover, clinical adoption of FL is expected to lead to models trained on datasets of unprecedented size, thus having a catalytic impact towards precision and personalized medicine [51]. FL shows promise in handling the new coronavirus electronic health record data to develop robust predictive models without compromising patient privacy [50] and can inspire more research on future COVID-19 applications, while boosting global data sharing.

One of the limitations of the present review is the rapid increase of new COVID-19 related AI models in the literature, challenging the completion of a full list of available studies. Another limitation is the fact that there were 31 included pre-prints which may differ from the final versions once accepted for official publication. A follow-up of the included studies at the time of writing indicated that 7 of the included pre-prints had already been peer-reviewed [34,59,45,70,76,81,102].

Based on the results of the Risk of Bias analysis, further research needs to be conducted to decrease the sources of bias in the included studies. Future studies can investigate the role of the most prominent predictors such as age, platelets, leukocytes, monocytes, eosinophils, lymphocytes, CRP, LDH, WBC from routine blood tests, imaging features from CXR and CT images. Most prominent AI methods (RF, LR, SVM, CNN, DT and XGBoost) along with the aforementioned predictors can be used as a leading approach in developing and validating future screening, diagnostic and prognostic prediction models that can be applied in clinical practice.

5. Conclusion

Artificial Intelligence methods are critical tools for utilizing the rapidly growing body of COVID-19 positive patient datasets, with a vast contribution in the fight against this pandemic. These multimodal datasets may include collected vitals, laboratory tests, comorbidities, CT scans or CXRs. In the present systematic review, we discussed the applicability and provided an overview of the AI-based prediction models developed by the rapidly growing literature, which can be used for screening, diagnosis, or prognosis of COVID-19 in the clinical setting. Limitations and considerations regarding the design and development of said prediction models were identified and future directions were proposed. Moreover, novel technologies such as explainable AI and Federated Learning could prove to be critical in tackling volatile crises like the COVID-19 pandemic. Increased collaboration in the development of the AI prediction models can enhance their applicability in the clinical practice and assist healthcare providers and developers in the fight against this pandemic and other public health crises.

CRediT authorship contribution statement

Eleni S. Adamidi: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Visualization, Writing - original draft, Writing - review & editing. **Konstantinos Mitsis:** Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Visualization, Writing - original draft. **Konstantina S. Nikita:** Conceptualization, Methodology, Supervision, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Zhu Na, Zhang D, Wang W, Li X, Yang Bo, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. N. Engl. J. Med. 2020;382 (8):727–33. https://doi.org/10.1056/NEJMoa2001017.

[2] Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, et al. Using Artificial Intelligence to Detect COVID-19 and Community-acquired Pneumonia Based on Pulmonary CT: Evaluation of the Diagnostic Accuracy. Radiology 2020;296(2):E65–71. https://doi.org/10.1148/radiol.2020200905.

[3] Soltan AAS et al. Rapid triage for COVID-19 using routine clinical data for patients attending hospital: development and prospective validation of an artificial intelligence screening test. Lancet Digit. Heal. Dec. 2020;3(2): e78–87. https://doi.org/10.1101/2020.07.07.20148361.

[4] H. S. Yang et al., "Routine Laboratory Blood Tests Predict SARS-CoV-2 Infection Using Machine Learning," Clin. Chem., vol. 66, no. 11, pp. 1396–1404, Nov. 2020, doi: 10.1093/clinchem/hvaa200.

[5] Ahammed K, Satu M, Abedin MZ, Rahaman M, Islam SMS. "Early Detection of Coronavirus Cases Using Chest X-ray Images. Employing Machine Learning and Deep Learning Approaches" 2020. https://doi.org/10.1101/2020.06.07.20124594.

[6] Y. Xu et al., "A collaborative online AI engine for CT-based COVID-19 diagnosis.," medRxiv Prepr. Serv. Heal. Sci., May 2020, doi: 10.1101/2020.05.10.20096073.

[7] R. Zhang et al., "Diagnosis of COVID-19 Pneumonia Using Chest Radiography: Value of Artificial Intelligence.," Radiology, p. 202944, Sep. 2020, doi: 10.1148/radiol.2020202944.

[8] A. F. de M. Batista, J. L. Miraglia, T. H. R. Donato, and A. D. P. Chiavegatto Filho, "COVID-19 diagnosis prediction in emergency care patients: a machine learning approach," 2020, doi: 10.1101/2020.04.04.20052092.

[9] L. J. Muhammad, M. M. Islam, S. S. Usman, and S. I. Ayon, "Predictive Data Mining Models for Novel Coronavirus (COVID-19) Infected Patients' Recovery.," SN Comput. Sci., vol. 1, no. 4, p. 206, Jul. 2020, doi: 10.1007/s42979-020-00216-w.

[10] Cai Q, Du S-Y, Gao Si, Huang G-L, Zhang Z, Li S, et al. A model based on CT radiomic features for predicting RT-PCR becoming negative in coronavirus disease 2019 (COVID-19) patients. BMC Med. Imaging 2020;20(1). https://doi.org/10.1186/s12880-020-00521-z.

[11] Ren HW et al. Analysis of clinical features and imaging signs of COVID-19 with the assistance of artificial intelligence. Eur. Rev. Med. Pharmacol. Sci. 2020;24(15):8210–8. https://doi.org/10.26355/eurrev_202008_22510.

[12] Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. PLoS Med. Jul. 2009;6(7):e1000100. https://doi.org/10.1371/journal.pmed.1000100.

[13] A. V. H. I. Melbourne, "Covidence Better systematic review management." https://www.covidence.org/.

[14] Grunkemeier GL, Jin R. Receiver operating characteristic curve analysis of clinical risk models. Ann. Thorac. Surg. 2001;72(2):323–6. https://doi.org/10.1016/S0003-4975(01)02870-3.

[15] A. E. Gerevini, R. Maroldi, M. Olivato, L. Putelli, and I. Serina, "Prognosis Prediction in Covid-19 Patients from Lab Tests and X-ray Data through Randomized Decision Trees," Oct. 2020, Accessed: Nov. 09, 2020. [Online]. Available: http://arxiv.org/abs/2010.04420.

[16] Li Y, Horowitz MA, Liu J, Chew A, Lan H, Liu Q, et al. Individual-Level Fatality Prediction of COVID-19 Patients Using AI Methods. Front. Public Heal. 2020;8. https://doi.org/10.3389/fpubh.2020.587937.

[17] Wang T, Paschalidis A, Liu Q, Liu Y, Yuan Y, Paschalidis IC. Predictive Models of Mortality for Hospitalized Patients With COVID-19: Retrospective Cohort Study. JMIR Med. Informatics 2020;8(10):. https://doi.org/10.2196/21788e21788.

[18] 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(1):51. https://doi.org/10.7326/M18-1376.

[19] M. Qjidaa et al., "Development of a clinical decision support system for the early detection of COVID-19 using deep learning based on chest radiographic images," 2020 Int. Conf. Intell. Syst. Comput. Vision, ISCV 2020, pp. 1–6, 2020, doi: 10.1109/ISCV49265.2020.9204282.

[20] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, and G. Jamalipour Soufi, "Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning," Med. Image Anal., vol. 65, 2020, doi: 10.1016/j.media.2020.101794.

[21] M. A. Elaziz, K. M. H. Id, A. Salah, M. M. Darwish, S. Lu, and A. T. Sahlol, "New machine learning method for image- based diagnosis of COVID-19," 2020, doi: 10.1371/journal.pone.0235187.

[22] F. M. Salman, S. S. Abu-Nasser, E. Alajrami, B. S. Abu-Nasser, and B. A. M. Ashqar, "COVID-19 Detection using Artificial Intelligence," Int. J. Acad. Eng. Res., vol. 4, no. 3, pp. 18–25, 2020, [Online]. Available: www.ijeais.org/ijaer.

[23] T. Padma and C. Usha Kumari, "Deep Learning Based Chest X-Ray Image as a Diagnostic Tool for COVID-19," Proc. - Int. Conf. Smart Electron. Commun. ICOSEC 2020, no. Icosec, pp. 589–592, 2020, doi: 10.1109/ICOSEC49089.2020.9215257.

[24] Liu B, Yan B, Zhou Y, Yang Y, Zhang Y. Experiments of Federated Learning for COVID-19 Chest X-ray Images. arXiv 2020.

[25] T. Mahmud, M. A. Rahman, and S. A. Fattah, "CovXNet: A multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization," Comput. Biol. Med., vol. 122, no. May, p. 103869, 2020, doi: 10.1016/j.compbiomed.2020.103869.

[26] O. Gozes, M. Frid, H. Greenspan, and D. Patrick, "Rapid AI Development Cycle for the Coronavirus (COVID-19) Pandemic : Initial Results for Automated Detection & Patient Monitoring using Deep Learning CT Image Analysis Article Type : Authors : Summary Statement : Key Results : List of abbreviati," arXiv:2003.05037, 2020, [Online]. Available: https://arxiv.org/ftp/arxiv/papers/2003/2003.05037.pdf.

[27] Iwendi C, Bashir AK, Peshkar A, Sujatha R, Chatterjee JM, Pasupuleti S, et al. COVID-19 patient health prediction using boosted random forest algorithm. Front. Public Heal. Jul. 2020;8. https://doi.org/10.3389/fpubh.2020.00357.

[28] S. I. di R. M. e Interventistica, "Covid-19: Casistica radiologica italiana," 2020. https://www.sirm.org/category/senza-categoria/covid-19/.

[29] H. Ko et al., "COVID-19 pneumonia diagnosis using a simple 2d deep learning framework with a single chest CT image: Model development and validation," J. Med. Internet Res., vol. 22, no. 6, Jun. 2020, doi: 10.2196/19569.

[30] S. R. Abdani, M. A. Zulkifley, and N. Hani Zulkifley, "A Lightweight Deep Learning Model for COVID-19 Detection," 2020 IEEE Symp. Ind. Electron. Appl. ISIEA 2020, 2020, doi: 10.1109/ISIEA49364.2020.9188133

[31] Barstugan M, Ozkaya U, Ozturk S. Coronavirus (COVID-19) Classification using CT Images by Machine Learning Methods. arXiv 2020;5:1–10.

[32] J. C. Monteral, "COVID-Chestxray Database," 2020. https://github.com/ieee8023/covid-chestxray-dataset.

[33] Rahimzadeh M, Attar A. A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2. Informatics Med. Unlocked 2020;19:100360. https://doi.org/10.1016/j.imu.2020.100360.

[34] D. Ezzat, A. ell Hassanien, and H. A. Ella, "GSA-DenseNet121-COVID-19: a Hybrid Deep Learning Architecture for the Diagnosis of COVID-19 Disease based on Gravitational Search Optimization Algorithm," arXiv, pp. 1–29, 2020.

[35] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. Rajendra Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," Comput. Biol. Med., vol. 121, Jun. 2020, doi: 10.1016/j.compbiomed.2020.103792.

[36] J. P. Cohen et al., "Predicting COVID-19 Pneumonia Severity on Chest X-ray With Deep Learning," Cureus, vol. 8, Jul. 2020, doi: 10.7759/cureus.9448.

[37] Borkowski A. Using Artificial Intelligence for COVID-19 Chest X-ray Diagnosis. Fed. Pract. 2020;vol. 19, no. 37(9):398–404. https://doi.org/10.12788/fp.0045.

[38] P. Mooney, "Chest X-Ray Images (Pneumonia)." https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia (accessed Dec. 15, 2020).

[39] T. Rahman, "COVID-19 Radiography Database," Kaggle, 2020. https://www.kaggle.com/tawsifurrahman/covid19-radiography-database (accessed Dec. 15, 2020).

[40] Ouchicha C, Ammor O, Meknassi M. "CVDNet: A novel deep learning architecture for detection of coronavirus (Covid-19) from chest x-ray images", Chaos. Solitons and Fractals 2020;140:110245. https://doi.org/10.1016/j.chaos.2020.110245.

[41] Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020;395 (10223):497–506. https://doi.org/10.1016/S0140-6736(20)30183-5.

[42] M. Fakhfakh, B. Bouaziz, F. Gargouri, and L. Chaari, "ProgNet: Covid-19 prognosis using recurrent and convolutional neural networks," pp. 1–7, 2020, doi: 10.1101/2020.05.06.20092874.

[43] C. Fang, S. Bai, Q. Chen, Y. Zhou, L. Xia, and L. Qin, "Deep learning for predicting COVID-19 malignant progression," 2020.

[44] Liu F, Zhang Q, Huang C, Shi C, Wang L, Shi N, et al. CT quantification of pneumonia lesions in early days predicts progression to severe illness in a cohort of COVID-19 patients. Theranostics 2020;10(12):5613–22. https://doi.org/10.7150/thno.45985.

[45] G. Chassagnon et al., "Holistic AI-Driven Quantification , Staging and Prognosis of COVID-19 Pneumonia," pp. 1–26, 2020.

[46] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," Proc. 2017 Int. Conf. Eng. Technol. ICET 2017, vol. 2018-Janua, pp. 1–6, 2018, doi: 10.1109/ICEngTechnol.2017.8308186.

[47] Rokach L, Maimon O. In: Data Mining and Knowledge Discovery Handbook. New York: Springer-Verlag; 2005. p. 165–92. https://doi.org/10.1007/0-387-25465-X_9.

[48] B. Hao et al., "Early prediction of level-of-care requirements in patients with COVID-19," Elife, vol. 9, Oct. 2020, doi: 10.7554/eLife.60519.

[49] Holzinger A, Carrington A, Müller H. Measuring the Quality of Explanations: The System Causability Scale (SCS): Comparing Human and Machine Explanations. KI - Kunstl. Intelligenz 2020;34(2):193–8. https://doi.org/10.1007/s13218-020-00636-z.

[50] Vaid A, Jaladanki SK, Xu J, Teng S, Kumar A, Lee S, et al. Federated Learning of Electronic Health Records to Improve Mortality Prediction in Hospitalized Patients With COVID-19: Machine Learning Approach. JMIR Med. Informatics Jan. 2021;9(1):e24207. https://doi.org/10.2196/24207.

[51] Sheller MJ, Edwards B, Reina GA, Martin J, Pati S, Kotrotsou A, et al. Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. Sci. Rep. 2020;10(1). https://doi.org/10.1038/s41598-020-69250-1.

[52] Moons KGM, Kengne AP, Woodward M, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. Heart 2012;98 (9):683–90. https://doi.org/10.1136/heartjnl-2011-301246.

[53] B. Efron and R. Tibshirani, "Improvements on Cross-Validation: The .632+ Bootstrap Method," J. Am. Stat. Assoc., vol. 92, no. 438, p. 548, Jun. 1997, doi: 10.2307/2965703.

[54] R. D. Riley et al., "External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: Opportunities and challenges," BMJ, vol. 353, pp. 27–30, 2016, doi: 10.1136/bmj.i3140.

[55] Wu X, Hui H, Niu M, Li L, Wang L, He B, et al. Deep learning-based multi-view fusion model for screening 2019 novel coronavirus pneumonia: A multicentre study. Eur. J. Radiol. 2020;128:109041. https://doi.org/10.1016/j.ejrad.2020.109041.

[56] Wang M, Xia C, Huang L, Xu S, Qin C, Liu J, et al. Deep learning-based triage and analysis of lesion burden for COVID-19: a retrospective study with external validation. Lancet Digit. Heal. 2020;2(10):e506–15. https://doi.org/10.1016/S2589-7500(20)30199-0.

[57] S. N. Nan et al., "A prediction model based on machine learning for diagnosing the early COVID-19 patients," pp. 1–12, 2020, doi: 10.1101/2020.06.03.20120881.

[58] F. Soares, "A novel specific artificial intelligence-based method to identify COVID-19 cases using simple blood exams," 2020, doi: 10.1101/2020.04.10.20061036.

[59] Feng C et al. A Novel Triage Tool of Artificial Intelligence Assisted Diagnosis Aid System for Suspected COVID-19 Pneumonia in Fever Clinics. SSRN Electron. J. 2020. https://doi.org/10.1101/2020.03.19.20039099.

[60] J. Wu et al., "Rapid and accurate identification of COVID-19 infection through machine learning based on clinical available blood test results," 2020, doi: 10.1101/2020.04.02.20051136.

[61] Banerjee A, Ray S, Vorselaars B, Kitson J, Mamalakis M, Weeks S, et al. Use of Machine Learning and Artificial Intelligence to predict SARS-CoV-2 infection from Full Blood Counts in a population. Int. Immunopharmacol. 2020;86:106705. https://doi.org/10.1016/j.intimp.2020.106705.

[62] Peng M et al. Artificial Intelligence Application in COVID-19 Diagnosis and Prediction. *SSRN Electron. J.*, no. April, Mar. 2020. https://doi.org/10.2139/ssrn.3541119.

[63] Li, Feng, "A machine-learning method based on ocular surface features for COVID-19 screening," pp. 1–16, 2020, doi: https://doi.org/10.1101/2020.09.03.20184226.

[64] F. Cabitza et al., "Development, evaluation, and validation of machine learning models for COVID-19 detection based on routine blood tests," Clin. Chem. Lab. Med., pp. 1–54, Oct. 2020, doi: 10.1515/cclm-2020-1294.

[65] Mei X, Lee H-C, Diao K-yue, Huang M, Lin B, Liu C, et al. Artificial intelligence–enabled rapid diagnosis of patients with COVID-19. Nat. Med. 2020;26 (8):1224–8. https://doi.org/10.1038/s41591-020-0931-3.

[66] Zhu X, Song B, Shi F, Chen Y, Hu R, Gan J, et al. Joint prediction and time estimation of COVID-19 developing severe symptoms using chest CT scan. Med. Image Anal. 2021;67:101824. https://doi.org/10.1016/j.media.2020.101824.

[67] X. Qi et al., "Machine learning-based CT radiomics model for predicting hospital stay in patients with pneumonia associated with SARS-CoV-2 infection: A multicenter study," 2020, doi: 10.1101/2020.02.29.20029603.

[68] Xiao L-shan, Li P, Sun F, Zhang Y, Xu C, Zhu H, et al. Development and Validation of a Deep Learning-Based Model Using Computed Tomography Imaging for Predicting Disease Severity of Coronavirus Disease 2019. Front. Bioeng. Biotechnol. 2020;8. https://doi.org/10.3389/fbioe.2020.00898.

[69] C. Salvatore et al., "Clinical and laboratory data, radiological structured report findings and quantitative evaluation of lung involvement on baseline chest CT in COVID-19 patients to predict prognosis," Radiol. Medica, no. 0123456789, 2020, doi: 10.1007/s11547-020-01293-w.

[70] M. Pourhomayoun and M. Shakibi, "Predicting Mortality Risk in Patients with COVID-19 Using Artificial Intelligence to Help Medical Decision-Making," medRxiv, p. 2020.03.30.20047308, Apr. 2020, doi: 10.1101/2020.03.30.20047308.

[71] Y. Gao et al., "Machine learning based early warning system enables accurate mortality risk prediction for COVID-19," Nat. Commun., vol. 11, no. 1, Dec. 2020, doi: 10.1038/s41467-020-18684-2.

[72] Roy S, Menapace W, Oei S, Luijten B, Fini E, Saltori C, et al. Deep Learning for Classification and Localization of COVID-19 Markers in Point-of-Care Lung Ultrasound. IEEE Trans. Med. Imaging 2020;39(8):2676–87. https://doi.org/10.1109/TMI.4210.1109/TMI.2020.2994459.

[73] M. Zhou et al., "Improved deep learning model for differentiating novel coronavirus pneumonia and influenza pneumonia," 2020, doi: 10.1101/2020.03.24.20043117.

[74] Babukarthik RG, Adiga VAK, Sambasivam G, Chandramohan D, Amudhavel J. Prediction of COVID-19 Using Genetic Deep Learning Convolutional Neural Network (GDCNN). IEEE Access 2020;8:177647–66. https://doi.org/10.1109/Access.628763910.1109/ACCESS.2020.3025164.

[75] Bai HX, Wang R, Xiong Z, Hsieh B, Chang K, Halsey K, et al. Artificial Intelligence Augmentation of Radiologist Performance in Distinguishing COVID-19 from Pneumonia of Other Origin at Chest CT. Radiology 2020;296(3):E156–65. https://doi.org/10.1148/radiol.2020201491.

[76] Castiglioni I, Schiaffino S, Polidori L, Gandola D, Messa C. "Artificial intelligence applied on chest X-ray can aid in the diagnosis of COVID-19 infection: a first experience from Lombardy. Italy" 2020.

[77] Wang S et al. "A deep learning algorithm using CT images to screen for Corona Virus Disease 2020;COVID-19)":1–27. https://doi.org/10.1101/2020.02.14.20023028.

[78] Cheng F-Y, Joshi H, Tandon P, Freeman R, Reich DL, Mazumdar M, et al. Using Machine Learning to Predict ICU Transfer in Hospitalized COVID-19 Patients. J. Clin. Med. Jun. 2020;9(6):1668. https://doi.org/10.3390/jcm9061668.

[79] Kim H-J, Han D, Kim J-H, Kim D, Ha B, Seog W, et al. An Easy-to-Use Machine Learning Model to Predict the Prognosis of Patients With COVID-19: Retrospective Cohort Study. J. Med. Internet Res. 2020;22(11):e24225. https://doi.org/10.2196/24225.

[80] Bertsimas D, Lukin G, Mingardi L, Nohadani O, Orfanoudaki A, Stellato B, et al. COVID-19 mortality risk assessment: An international multi-center study. PLoS ONE Dec. 2020;15(12):e0243262. https://doi.org/10.1371/journal.pone.0243262.

[81] E. Jimenez-Solem et al., "Developing and Validating Covid-19 Adverse Outcome Risk Prediction Models From a Bi-National European Cohort of 5594 Patients," medRxiv, p. 2020.10.06.20207209, 2020, [Online]. Available: https://doi.org/10.1101/2020.10.06.20207209.

[82] Burdick H, Lam C, Mataraso S, Siefkas A, Braden G, Dellinger RP, et al. Prediction of respiratory decompensation in Covid-19 patients using machine learning: The READY trial. Comput. Biol. Med. 2020;124:103949. https://doi.org/10.1016/j.compbiomed.2020.103949.

[83] Cai W, Liu T, Xue X, Luo G, Wang X, Shen Y, et al. CT Quantification and Machine-learning Models for Assessment of Disease Severity and Prognosis of COVID-19 Patients. *Acad. Radiol.*, Sep. 2020;27(12):1665–78. https://doi.org/10.1016/j.acra.2020.09.004.

[84] Wu G, Zhou S, Wang Y, Lv W, Wang S, Wang T, et al. A prediction model of outcome of SARS-CoV-2 pneumonia based on laboratory findings. Sci. Rep. 2020;10(1). https://doi.org/10.1038/s41598-020-71114-7.

[85] Ma X, Ng M, Xu S, Xu Z, Qiu H, Liu Y, et al. Development and validation of prognosis model of mortality risk in patients with COVID-19. Epidemiol. Infect. Aug. 2020;148. https://doi.org/10.1017/S0950268820001727.

[86] Izquierdo J, Ancochea J, Soriano J. Clinical Characteristics and Prognostic Factors for ICU Admission of Patients with COVID-19 Using Machine Learning and Natural Language Processing. SSRN Electron. J. 2020. https://doi.org/10.2139/ssrn.3582866.

[87] G. Wu et al., "Development of a clinical decision support system for severity risk prediction and triage of COVID-19 patients at hospital admission: An international multicentre study," Eur. Respir. J., vol. 56, no. 2, 2020, doi: 10.1183/13993003.01104-2020.

[88] X. Jiang et al., "Towards an artificial intelligence framework for data-driven prediction of coronavirus clinical severity," Comput. Mater. Contin., vol. 63, no. 1, pp. 537–551, 2020, doi: 10.32604/cmc.2020.010691.

[89] Liang W, Yao J, Chen A, Lv Q, Zanin M, Liu J, et al. Early triage of critically ill COVID-19 patients using deep learning. Nat. Commun. 2020;11(1). https://doi.org/10.1038/s41467-020-17280-8.

[90] J. Gong et al., "A tool for early prediction of severe coronavirus disease 2019 (covid-19): A multicenter study using the risk nomogram in Wuhan and Guangdong, China," Clin. Infect. Dis., vol. 71, no. 15, pp. 833–840, 2020, doi: 10.1093/cid/ciaa443.

[91] A. Abdulaal, A. Patel, E. Charani, S. Denny, N. Mughal, and L. Moore, "Prognostic modeling of COVID-19 using artificial intelligence in the United Kingdom: Model development and validation," J. Med. Internet Res., vol. 22, no. 8, Aug. 2020, doi: 10.2196/20259.

[92] J. Chen et al., "Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography," vol. 10, Nov. 2020, Accessed: Nov. 09, 2020. [Online]. Available: /articles/s41598-020-76282-0.

[93] A. Ter-Sarkisov, "COVID-CT-Mask-Net: Prediction of COVID-19 from CT Scans Using Regional Features," medRxiv, vol. 50, p. 2020.10.11.20211052, 2020, [Online]. Available: https://doi.org/10.1101/2020.10.11.20211052.

[94] Yan T, Wong PK, Ren H, Wang H, Wang J, Li Y. Automatic distinction between COVID-19 and common pneumonia using multi-scale convolutional neural network on chest CT scans. Chaos, Solitons Fractals 2020;140:110153. https://doi.org/10.1016/j.chaos.2020.110153.

[95] R. Lokwani, A. Gaikwad, V. Kulkarni, A. Pant, and A. Kharat, "Automated Detection of COVID-19 from CT Scans Using Convolutional Neural Networks," Jun. 2020, [Online]. Available: https://arxiv.org/pdf/2006.13212.pdf.

[96] C. Jin et al., "Development and evaluation of an artificial intelligence system for COVID-19 diagnosis," Nat. Commun., vol. 11, no. 1, 2020, doi: 10.1038/s41467-020-18685-1.

[97] Li Y, Wei D, Chen J, Cao S, Zhou H, Zhu Y, et al. Efficient and Effective Training of COVID-19 Classification Networks with Self-Supervised Dual-Track Learning to Rank. IEEE J. Biomed. Heal. Informatics 2020;24(10):2787–97. https://doi.org/10.1109/JBHI.622102010.1109/JBHI.2020.3018181.

[98] Sun L, Mo Z, Yan F, Xia L, Shan F, Ding Z, et al. Adaptive Feature Selection Guided Deep Forest for COVID-19 Classification with Chest CT. IEEE J. Biomed. Heal. Informatics 2020;24(10):2798–805. https://doi.org/10.1109/JBHI.622102010.1109/JBHI.2020.3019505.

[99] L. Wang, Z. Q. Lin, and A. Wong, "COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images," Sci. Rep., vol. 10, no. 1, Dec. 2020, doi: 10.1038/s41598-020-76550-z.

[100] K. Zhang et al., "Erratum: Clinically Applicable AI System for Accurate Diagnosis, Quantitative Measurements, and Prognosis of COVID-19 Pneumonia Using Computed Tomography (Cell (2020) 181(6) (1423–1433. e11), (S0092867420305511), (10.1016/j.cell.2020.04.045))," Cell, vol. 182, no. 5, p. 1360, 2020, doi: 10.1016/j.cell.2020.08.029.

[101] A. S. Yadaw, Y. chak Li, S. Bose, R. Iyengar, S. Bunyavanich, and G. Pandey, "Clinical features of COVID-19 mortality: development and validation of a clinical prediction model," Lancet Digit. Heal., vol. 2, no. 10, pp. e516–e525, Oct. 2020, doi: 10.1016/S2589-7500(20)30217-X.

[102] J. G. Klann et al., "Validation of a Derived International Patient Severity Algorithm to Support COVID-19 Analytics from Electronic Health Record Data," medRxiv, p. 2020.10.13.20201855, 2020, [Online]. Available: http://medrxiv.org/content/early/2020/10/14/2020.10.13.20201855.abstract.

[103] S. Navlakha, S. Morjaria, R. Perez-Johnston, A. Zhang, and Y. Taur, "Projecting COVID-19 disease severity in cancer patients using purposefully-designed machine learning," medRxiv, p. 2020.10.16.20214098, 2020, [Online]. Available: https://doi.org/10.1101/2020.08.23.20179838.

[104] Shashikumar SP et al. Development and Prospective Validation of a Deep Learning Algorithm for Predicting Need for Mechanical Ventilation. Chest 2020;no. 405:Dec. https://doi.org/10.1016/j.chest.2020.12.009.

[105] A. Youssef et al., "Development and Validation of Early Warning Score Systems for COVID-19 Patients," medRxiv, no. October, p. 2020.11.04.20225904, 2020, [Online]. Available: https://doi.org/10.1101/2020.11.04.20225904.

[106] A. Karthikeyan, A. Garg, P. K. Vinod, and U. Deva Priyakumar, "Machine learning based clinical decision support system for early COVID-19 mortality prediction," medRxiv, vol. 2019, no. December 2019, p. 2020.08.19.20177477, 2020, [Online]. Available: https://doi.org/10.1101/2020.08.19.20177477.

[107] E. Casiraghi et al., "Explainable machine learning for early assessment of COVID-19 risk prediction in emergency departments," IEEE Access, no. December 2019, pp. 1–1, Oct. 2020, doi: 10.1109/access.2020.3034032.

[108] L. Yan et al., "A machine learning-based model for survival prediction in patients with severe COVID-19 infection," 2020, doi: 10.1101/2020.02.27.20028027.

[109] P. Schwab, A. DuMont Schütte, B. Dietz, and S. Bauer, "Clinical Predictive Models for COVID-19: Systematic Study," J. Med. Internet Res., vol. 22, no. 10, p. e21439, Oct. 2020, doi: 10.2196/21439.

[110] Nemati M, Ansary J, Nemati N. Machine-Learning Approaches in COVID-19 Survival Analysis and Discharge-Time Likelihood Prediction Using Clinical Data. Patterns Aug. 2020;1(5):. https://doi.org/10.1016/j.patter.2020.100074100074.

[111] J. S. Zhu et al., "Deep-learning artificial intelligence analysis of clinical variables predicts mortality in COVID-19 patients," J. Am. Coll. Emerg. Physicians Open, no. May, pp. 1–10, 2020, doi: 10.1002/emp2.12205.

[112] Vaid A et al. Machine Learning to Predict Mortality and Critical Events in COVID-19 Positive New York City Patients: A Cohort Study (Preprint). J. Med. Internet Res. 2020;22(11):Sep. https://doi.org/10.2196/24018.

[113] S. Wollenstein-Betech, C. G. Cassandras, and I. C. Paschalidis, "Personalized predictive models for symptomatic COVID-19 patients using basic preconditions: Hospitalizations, mortality, and the need for an ICU or ventilator," Int. J. Med. Inform., vol. 142, no. May, p. 104258, 2020, doi: 10.1016/j.ijmedinf.2020.104258.

[114] L. Sun et al., "Combination of four clinical indicators predicts the severe/critical symptom of patients infected COVID-19," J. Clin. Virol., vol. 128, no. May, p. 104431, 2020, doi: 10.1016/j.jcv.2020.104431.

[115] N. Razavian et al., "A validated, real-time prediction model for favorable outcomes in hospitalized COVID-19 patients," npj Digit. Med., vol. 3, no. 1, 2020, doi: 10.1038/s41746-020-00343-x.

[116] Assaf D et al. Utilization of machine-learning models to accurately predict the risk for critical COVID-19. Intern. Emerg. Med. 2020;15(8):1435–43. https://doi.org/10.1007/s11739-020-02475-0.

[117] A. Trivedi et al., "Effective Deep Learning Approaches for Predicting COVID-19 Outcomes from Chest Computed Tomography Volumes," medRxiv, p. 2020.10.15.20213462, 2020, [Online]. Available: https://doi.org/10.1101/2020.10.15.20213462.

[118] Ryan L et al. Mortality prediction model for the triage of COVID-19, pneumonia, and mechanically ventilated ICU patients: A retrospective study. Ann. Med. Surg. 2020;59(July):207–16. https://doi.org/10.1016/j.amsu.2020.09.044.

[119] E. Tjoa and C. G. Fellow, "A Survey on Explainable Artificial Intelligence (XAI): Towards Medical XAI," arXiv, pp. 1–21, 2019, doi: 10.1109/tnnls.2020.3027314.

[120] M. Athanasiou, K. Sfrintzeri, K. Zarkogianni, A. C. Thanopoulou, and K. S. Nikita, "An explainable XGBoost-based approach towards assessing the risk of cardiovascular disease in patients with Type 2 Diabetes Mellitus," Sep. 2020, Accessed: Apr. 27, 2021. [Online]. Available: http://arxiv.org/abs/2009.06629.