REVISTA CUBANA
DE TRANSFORMACIÓN DIGITAL

RNPS 2487 • ISSN 2708-3411
Vol. 4 • Nro. 3 • julio-septiembre 2023 •e225

ARTÍCULO ORIGINAL

ACTUALIZACION SOBRE MODELOS DE INTELIGENCIA ARTIFICIAL, PARA PREDICCION DE EVOLUCION EN PACIENTES DE COVID-19 EN CUBA

UPDATE ON ARTIFICIAL INTELLIGENCE MODELS FOR PREDICTING OUTCOME IN COVID-19 PATIENTS IN CUBA

Eduardo Garea Llano
eduardo.garea@cneuro.edu.cu • ***https://orcid.org/ https://orcid.org/0000-0001-9101-1589***
Evelio Gonzalez Dalmau
evelio.gonzalez@cneuro.edu.cu • ***https://orcid.org/ 0000-0003-4569-2103***

Centro de Neurociencias de Cuba, Cuba

Recibido: 2023-07-03* • *Aceptado: 2023-11-06

RESUMEN

En este artículo presentamos una revisión de los modelos basados en Inteligencia Artificial (IA), orientados a la predicción de manifestaciones graves provocadas por el virus SARS-COV2. El objetivo de esta revisión fue evaluar los hallazgos más relevantes publicados entre 2020 y 2023, que puedan servir de base al desarrollo de un modelo propio ajustado a las condiciones de nuestro país, caracterizadas por la ausencia de un modelo propio ajustado a la ausencia de una base de datos de datos clínicos bien estructurada y con la presencia de estudios radiológicos apoyados por imágenes de rayos X de tórax (CXR) y tomografía computarizada (CT). Realizamos una revisión sistemática para resumir y evaluar críticamente los estudios disponibles que han desarrollado modelos de pronóstico de COVID-19 basados en IA, que predicen resultados de salud, sobre todo en modelos que utilizan como base las imágenes CXR y CT. Se realizaron búsquedas en tres bases de datos bibliográficas, para identificar artículos publicados sobre modelos de pronóstico que predijeran resultados adversos en pacientes adultos con COVID- 19, incluido el ingreso a la unidad de cuidados intensivos, la necesidad de ventilación mecánica y la mortalidad.

El estudio demostró que los modelos basados en aprendizaje profundo, que utilizan imágenes CXR o CT y su combinación con datos clínicos no complejos, pueden alcanzar un rendimiento significativo en la predicción. Por ello, aquí proponemos una estrategia para abordar este desafío, según las condiciones de nuestro país, combinando la clasificación del grado de gravedad de la afectación pulmonar en imágenes de CXR, datos clínicos de comorbilidades y datos biográficos.

Palabras clave: COVID-19, modelos predictivos, Inteligencia Artificial, datos clínicos, datos radiológicos.

ABSTRACT

In this work we present a review of models based on Artificial Intelligence (AI) aimed at predicting serious manifestations caused by the SARS-COV2 virus. The objective of this review was to evaluate the most relevant findings published between 2020 and 2023 that can serve as a basis for the development of our own model adjusted to the conditions of our country characterized by the absence of a well-structured clinical database with the presence of radiological studies based on chest x-ray (CXR) and computed tomography (CT) images. We conducted a systematic review to summarize and critically evaluate available studies that have developed AI-based COVID-19 prognostic models that predict health outcomes, especially models based on CXR and CT images. Three bibliographic databases were searched to identify published articles on prognostic models predicting adverse outcomes in adult patients with COVID-19, including intensive care unit admission, need for mechanical ventilation, and mortality.

The study demonstrated that Deep Learning-based models using CXR or CT images and their combination with non-complex clinical data can achieve significant prediction performance. On this basis, in the work we propose a strategy to address this challenge under the conditions of our country by combining the classification of the degree of severity of lung involvement in CXR images, clinical data on comorbidities and biographical data.

Keywords: *COVID-19; Predictive models; Artificial intelligence; Evolution; Clinical and radiological data.*

INTRODUCTION

With the decrease in COVID-19 cases in our country due to the development of our own vaccines and their massive application, as well as the health measures adopted, this disease still remains a national concern. The Cuban Neuroscience Center (CNEURO), as part of a project leaded by it and within the multidisciplinary scientific research program "Biotechnology, Pharmaceut ica l Industry and Medical Technology" (put in place by the Cuban government to respond to the CONVID19 pandemic), is developing a model based on AI for the prognosis of patient evolutio n affected by the disease. CNEURO is also improving the use of available CXR images in the diagnosis, prognosis and monitoring of

Covid-19 patients. This effort integrates specialists from the Cuban Society of Imaging, MINSAP and two Cuban universities (Technological University of Havana and Central University of Las Villas). The project includes the validation of artificial intelligence protocols and algorithms for predicting the trajectory of COVID-19 disease by combining patient health records (clinical data) and chest images (CXR, CT). Additionally, the timely translation of AI-related research into clinically validated and appropriately regulated systems that can be used in a real-life hospital environment.

As previous results obtained in this research, there are several works published by researchers of this project that have contributed to obtaining tools for the diagnosis and evaluation of the disease. In Portal et al (2022), a model was presented to automatically classify COVID-19 from CXR images. For this, the architecture of a resnet34 CNN was used with input images of size 512 × 512 pixels. Furthermore, training of the networks was limited to the internal regions of the lungs, using a segmented image and patch partitioning. This model demonstrated together with human specialists its usefulness in the identification of COVID-19 in clinical settings.

In Garea et al (2021a, 2021b, 2022, 2023), a model is presented and refined to estimate the degree of lung involvement and its severity in CXR images in patients diagnosed with COVID-19 in an advanced stage of the disease. The index is obtained from a method that combines image quality evaluation, digital image processing and deep learning for lung region segmentation.

As another important result of this project, in February 2023, the health registration was obtained by the Cuban regulatory authority (CECMED) of the Rx-COVID 19 system. Automated and weighted multiplatform software to record chest X-ray and CT and facilitate the diagnosis of COVID. -19. In this system, a significant amount of radiological data of patients affected by COVID-19 in our country has been documented.

On this basis, the project proposes to move to a higher phase by developing a computational tool for the prognosis of evolution towards adverse outcomes in adult patients with COVID-19. The need to continue developing new technological capabilities for the evaluation of patient follow - up has been identified in the face of resource limitations in the current Cuban context to respond effectively to similar challenges in the future. This development aims at a more integrative and interdisciplinary approach to prognostic evaluation protocols so that patients admitted with COVID-19 have a better chance of survival and improve their health status by optimizing protocols and treatments.

In this work we will carry out a systematic literature review (SLR) with the following objectives:

(1) Systematically review the models that have been developed to predict COVID-19 outcomes based on the use of RX, CT images and clinical data; (2) Analyze the most relevant predictive models that are currently available; (3) Synthesize and extract useful results and conclusions about the prediction models of COVID-19 outcomes that can serve as a basis for the development of our own predictive model adjusted to the conditions of our country (characterized by the absence of a well-known structured clinical database and with the presence of radiological studies based on CXR and CT images given their low cost and the portability of the existing equipment in healthcare centers).

The work is structured as follows. The "Methodology" Section provides a description of the method of searching and selecting papers for this review and present various classifications of the selected papers. In the "Results and Discussion" section present a detailed analysis of the 6 significant selected works. On the basis of the analysis of the performance of the studied methods we carried out a general discussion. Finally we offer the conclusions of the work.

METHODOLOGY

To carry out this review study we used the SLR methodology. SLR is a methodology used to evaluate studies through primary analyzes on published articles based on specific search criteria. An SLR is performed based on previous similar studies through a systematic review in national and international publication databases. The ultimate purpose of the SLR is to summarize the studies conducted and identify advantages and limitations between previous studies and current studies.

According to Okoli (2015), SLR is “a systematic, explicit, detailed and repeatable approach to identifying, evaluating and analyzing the body of existing work carried out by researchers, academics and practitioners”. On the other hand, Tranfield et al. (2003), considers SLR a “fundamental scientific activity”. In Moher et al. (2009) a checklist of Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) was presented.

The research period for this study was from December 2020 to May 2023. This study was conducted in 4 phases: (1) the development of literature search strategies, (2) the formulation of inclusion and exclusion criteria, (3) quality evaluation, and (4) analysis and conclusion.

Research Questions

For the development of this documentary research, we stated a research question (RC) to achieve a clear definition of the main topic. The motivation and frequently asked questions of this study were as follows:

Motivation: Identify models based on artificial intelligence for the prediction of COVID-19 outcomes and that use RX images, CT and clinical data or their combination.

RQ1: What factors support the prediction of COVID-19 outcomes based on artificial intelligence techniques?

RQ2: What methods and techniques do the models identify to predict COVID-19 outcomes?

Inclusion and exclusion criteria

Current search engines provide a high level of retrieval, which leads to a large number of irrelevant resources being retrieved. Therefore, to obtain effective results, we followed a systematic search strategy. This stage of SLR examines the literature to find relevant literature based on particular criteria. In this study, 3 inclusion and exclusion criteria were adopted to identify relevant content and restrict irrelevant content. The first inclusion criterion was the type of document: only published documents were included, excluding manuscripts under review and unpublished ones. Domain (i.e., the topic area identified for the study) was the second selection criterion; we then included papers containing predictive models developed or used in the COVID-19 domain, while other papers were excluded. The last selection criterion was the language in which the document was published. To avoid confusion and complexity related to translation, only documents available in English were included, while documents in other languages were excluded (Table 1).

Table 1. Inclusion exclusion criteria

Criteria	Inclusion	Exclusion
Document type	Published documents	Under review, unpublished or upcoming documents
Domain	Predictive AI models of COVID-19 outcomes	Other predictive models of COVID-19 Other predictive models of COVID-19 outcomes not based on AI
Language	English	Other than English

Databases and search strategies

The terms were searched in several databases (Google Scholar, Scopus, Publish or Perish and Web of Science [WoS]). The search terms are as follows: predictive AI models, COVID-19 outcomes, Coronavirus, SARS-CoV, SARS-CoV-2, healthcare, healthcare system, survival model, healthcare. Various combinations of search terms were used to retrieve resources in particular databases. Some of the search strings used are the following: “Predictive AI models” and “COVID-19 outcomes”; “COVID-19 RX and CT Datasets” and “Prediction Modeling”; “Predictive analysis" and "COVID-19 clinical data".

After applying the inclusion and exclusion criteria, 574 documents of original articles and 16 relevant reviews were recovered. Therefore, a total of 590 documents went to the second stage of scrutiny and quality evaluation (Table 2). The percentages of articles from various databases in the initial, selection and acceptance stages of the document selection process are illustrated.

Table 2. Document selection

Database	Initial	Screened	Accepted
Google Scholar	366	17	8
Scopus	136	9	7
Web of Sciences	88	9	3
Total selected	590	35	18

In the initial phase, of the total number (i.e., 590 documents) of retrieved documents, Google Scholar accounted for 62%, Scopus for 23%, and WoS for 15% (Figure 1A). After initial screening, 35 papers were included for further consideration. During the selection phase, 48% of the initially included documents were retrieved from Google Scholar. Of the remaining 52%, Scopus and WoS had a share of 26% each (Figure 1B). Of the total accepted documents, 44% were retrieved from Google Scholar, 39% from Scopus, and 17% from WoS (Figure 1C).

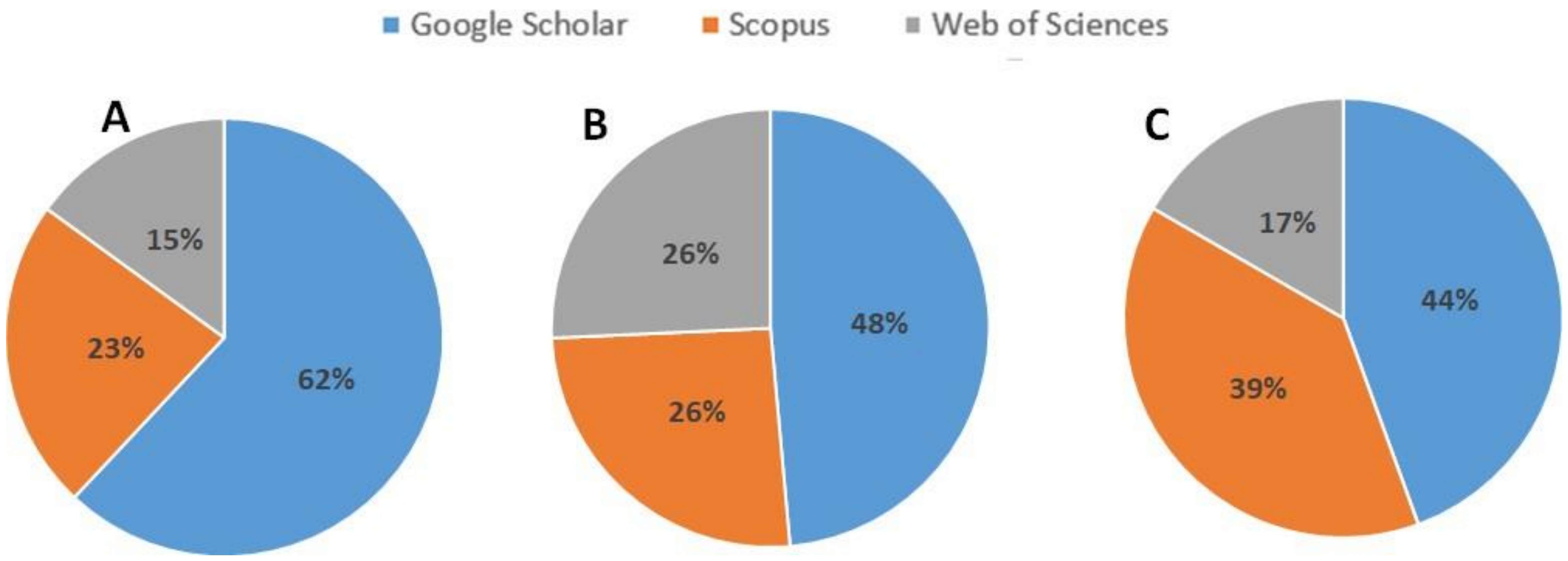

Figura. 1. Database's percentage (May 2023). (A) Document selection (initial). (B) Document selection (screened). (C) Document selection (included). Document selection was carried out based on selection criteria.

A total of 16 relevant review articles published in the period 2020 to 2023 were detected (Albahri, O. S. et al., 2020; Bansal, A. et al., 2020; Jalaber, C. et al., 2020; Kumar, A. et al., 2020; Naudé,W., 2020; Shi, F. et al., 2020; Swapnarekha, H. et al., 2020; Gallo Marin B. et al., 2021; Mbunge, E. et al., 2021; Rahimi, I. et al., 2021; Rasheed, J. et al., 2021; Hassan, A. et al., 2022; Jamshidi, M. B. et al., 2022; Almotairi, K. H. et al., 2023; Buttia, C. et al., 2023; Heidari, A. et al, 2023).

For the primary review of the 16 identified surveys we found 368 relevant papers based on title and abstract. From them after an additional review, 166 were excluded en the initial phase because they were not directly related to predictive models of disease evolution, but rather to predictive models of evaluation of the development of the epidemic in different countries or to the direct diagnosis of the disease, 202 paper were chosen for a more in-depth analysis in the second phase, of which we finally chose 6 reports.

The rest of the initial selected 590 papers (388) 12 were selected in the second phase. All selected papers are based on the use of chest clinical data alone, X-ray images (CXR) and/or Tomography (TC) alone, or their combination with clinical data for the prediction of severe manifestations and mortality from COVID-19.

Quality Assessment and Coding

Quality evaluation of a phenomenon is conducted as a systematic way to avoid biases and errors. 590 documents were chosen. Based on their titles, these documents were further analyzed and 35 documents were screened. The content was scrutinized on the basis of the title, abstract, introduction, and conclusion and 18 studies were finally selected for the review. The distribute ion by year of publication of the selected papers is shown in Table 3.

Table 3. Distribution by year of publication of the elected report

Year	References	Outcomes	Used data	Model
2020	Bae, J. *et al.,* 2020 (1)	Mechanic ventilation (MV), mortality	CXR	Machine learning (ML), Deep learning (DL)
	Liang W. *et al.,* 2020 (2)	MV, mortality, ICU.	CXR, clinical and demographic data	Machine Learning (ML)
	Zhang B. *et al.,* 2020 (3)	Severe Outcome	CT, clinical data	ML
2021	Heo J *et al.,* 2021 (4)	ICU	Clinical data	ML

	Lieveld A. et al, 2021 (5)	ICU, Death, Hospitalization	CT	ML
	Quiroz J. C. *et al.*, 2021 (6)	Severity	CT, clinical data	ML
	Tanboga I. *et al.*, 2021 (7)	Mortality	Clinical data	ML
	Ho T. *et al.*, 2021 (8)	Respiratory failure	CT, clinical data	DL
2022	Ortiz A. *et al.*, 2022 (9)	Death	CT, clinical data	DL
	Duanmu, H. *et al.*, 2022 (10)	Death	CXR, clinical data	DL
	Gourdeau, D. *et al.*, 2022 (11)	MV	CXR	DL
	Matsumoto, T. *et al.*, 2022 (12)	Mortality and time to death at admission	CXR, clinical data	DL
2023	Olowolayemo A. *et al.*, (13)	Infection severity	CXR	DL
	Asteris *et at.* (14)	ICU	Clinical data	DL
	Wendland. *et al.* (15)	mortality, ICU and MV	Clinical data	ML, DL
	Verzellesi *et al.* .(16)	Mortality	CT, clinical data	ML, DL
	Panagiotis *et al.* (17)	ICU	Clinical data	DL
	Fu *et al.* (18)	Severe Outcomes	CT, clinical data	DL

As a first approximation to a taxonomic classification of the selected models, we simply took the year of its publication (Table 3), then on this basis we carried out the analyzes of various aspects that allowed us to reach conclusions about the nature of the models, the type of data from which they are fed to make the predictions and the types of predictions they are capable of making along with their effectiveness. In this sense we decided to make four large groupings of the reviewed methods. In the following subsection we present the proposed taxonomic divisions. The numbers in parentheses correspond to the index assigned to each report in Table 3.

Taxonomic classifications of the studied models

Taking into account the type of outcome prediction:

- Models that estimate the need for mechanical ventilation: (1), (8), (11), (15)
- Models that estimate the need for patient admission to the ICU: (2), (4), (5), (14), (15), (17)
- Models that estimate the possibility of death of the patient: (5), (7), (9), (10), (12), (15), (16)
- Models that estimate the possibility of patient evolution towards more severe forms of the disease: (3), (6), (13).

By the number of predictive outputs:

- Simple models that predict only one condition: (1), (3), (4), (6), (7), (9), (10), (11), (13), (16), (17)
- Combined models that predict two or more conditions: (2), (5), (8), (12), (14), (15)

By the type of data used in the prediction process:

- Models using only CXR images: (1), (11), (13)
- Models using only CT images: (5)
- Models using only clinical data: (4), (7), (14), (15), (17)
- Mixed models using CXR images and clinical data: (2), (10), (12)
- Mixed models using CT images and clinical data: (3), (6), (8), (9), (16)

By the nature of the algorithms used in the predictive process:

- Models based on classic machine learning algorithms: (2), (3), (4), (5), (6), (7)
- Models based on deep learning algorithms: (8), (9), (10), (11), (12), (13), (14)
- Mixed models that combine classic machine learning and deep learning algorithms: (1), (15), (16).

RESULTS

For the development of this report, due the limited number of pages, we made a selection of the 6 more representative papers for a deeper analysis in this work. Papers were chosen under the criteria of the biggest impact factor of the publication. In the following section we carry out a detailed analysis of the selected papers in terms of scope and results. The order of the analyzed reports responds to the classification based on the nature of the algorithms used in the predictive process.

Analysis and discussion of the selected methods

Models based on classic machine learning algorithms

In Zhang, B et al (2020), a study for the early detection of COVID-19 patients is carried out, for which they developed and validated a nomogram to predict a poor outcome at 30 days in patients with COVID-19. The prediction model was developed on a primary cohort consisting of 233 patients with laboratory-confirmed COVID-19, and data was collected from January 3 to March 20, 2020. Significant prognostic factors for poor outcome were identified and integrated at 30 days to build the nomogram. The model underwent internal validation and external validation with two separate cohorts of 110 and 118 cases, respectively. The performance of the nomogram was evaluated with respect to its predictive accuracy, discriminatory ability, and clinical utility. The model was externally validated in two cohorts achieving an AUC of 0,946 and 0,878, sensitivity of 100 and 79 %, and specificity of 76,5 and 83,8 %, respectively.

This study has some limitations. The retrospective nature of the sample may introduce potential risks of bias in the data. The CT score is subjective with large intra- and inter-observer variability obtained from the initial CT examination; CT-based radiomics or deep learning and follow-up CT may provide further prognostic information. This model is not applicable for patients with critical illness on admission, which may result in inclusion bias. COVID-19 triage could lead to less severe cases having delayed testing, which would skew the inclusion set to a more critical status; therefore, the performance of the model may be overestimated. Finally, the identification of predictors depends on the features available, the feature selection method used, and the sample size of the studies.

Models based on deep learning algorithms

In Ho. T et al. (2021), the authors developed deep learning models to rapidly identify high-risk COVID-19 patients based on CT images and clinical data. To do this, data from 297 patients with COVID-19 from five hospitals in Daegu, South Korea were analyzed. A deep learning model combining an artificial neural network for clinical data and a convolutional neural network for 3D CT imaging data, was developed to classify these cases as high risk of severe progression (i.e. , event) or low risk (i.e. no events). A total of 19 clinical features from were concatenated with a 64- dimensional feature vector from the CT image. The developed model achieved high classification performance using novel images of coronavirus pneumonia lesions (ie, 93,9 % accuracy, 80,8 % sensitivity, 96,9 % specificity, and 0,916 AUC) and lung segmentation images (ie, 94,3 % accuracy, 74,7 % sensitivity, 95,9 % specificity, and AUC score of 0,928) for event versus non-event groups. However, the study has several limitations. First, it was a retrospectively designed study, where the size of the data set (ie the number of patients) was small. Furthermore, the number of patients who progressed to the severe stage was relatively small. Therefore, the precision and sensitivity of the CNN model were based on CT images that can be affected by variation in unbalanced data sets. Second, only COVID-19 data was included in this study. However, a true diagnostic model must contain the features to distinguish COVID-19 from other types of pneumonia. Third, the authors compared their model with three typical 2D models that were developed in a 3D domain. As these models were designed for 2D imaging, this comparison did not present the alternatives for the same application domain. Therefore, 3D context is important to differentiate between evented and eventless COVID-19 structures, which requires the development of pretrained 3D models.

In Ortiz A. et al (2022), the authors evaluated the value of aggregated chest CT data for COVID- 19 prognosis compared to clinical metadata. They developed a patient-level algorithm to aggregate chest CT volume into a 2D representation that can be integrated with clinical metadata to distinguish COVID-19 pneumonia from chest CT volumes of healthy participants and participate ants with other viral pneumonia.

Furthermore, the authors present a multitask model for the joint segmentation of different classes of lung lesions present in lungs infected with COVID-19 with the aim of extracting highly relevant local features. From this they create prognostic models using the extracted characteristics along with patient demographics, comparing the performance of the models using different combinat ions of relevant data to predicting mortality. The overall objective of this work was to enable automated extraction of relevant features from chest CT volumes that can be incorporated with clinical data for risk stratification of COVID-19 patients. They compared this multitask segmentation approach with combining feature-independent volumetric CT classification feature maps with clinical metadata to predict mortality. For the prediction task, the authors compared the performance of five different machine learning models for predicting mortality in patients with COVID-19.

Experimental results showed that the combination of features derived from chest CT volumes improves the AUC performance to 0.80 from the 0.52 obtained using only clinical patient data. Olowolayemo. A et al (2023) presented a model to determine the mortality of an infected patient upon arrival at health facilities to determine whether or not admission to intensive care is necessary. A CNN model based on the ResNet-18 architecture was trained on CXRs of COVID-19 patients to estimate their mortality risk. Training of the proposed model was performed using the Stony Brook University (SBU) dataset. The first stage of the study used the IEEE8023 dataset, obtained from a public GitHub repository. The model using the original unimputed data with class weight adjustment achieved the best performance among the six compared models. This model was used in the presented study using the SBU data set. Subsequently, the first model, that is, the best model from the first training phase, and the second model, that is, the model that was fitted on the SBU data set, were compared using a test set created from a part of the SBU data set.

The result showed that the second model trained on the SBU dataset performed much better in terms of accuracy, significantly outperforming the first model trained only on the IEEE8023 dataset, with an accuracy of 86 %. This was despite the fact that chest radiographs were not discriminated based on the duration of infection of the patients at the time of the radiograph. The model also achieved a high recovery value of 81,5 % with a false negative rate of 18,4 % on the test set.

The main limitation of this work is the number of CXR images used. The limited amount of data, especially CXRs of non-surviving cases, forced them to use oversampling methods, such as image augmentation, to increase the number of images in the minority class.

Mixed models that combine classic machine learning and deep learning algorithms

Bae J et al., (2020). The aim of this study was to predict mechanical ventilation requirement and mortality using computational modeling of CXRs for COVID-19 patients. A total of 530 unidentified CXRs from 515 COVID-19 patients treated between March and August 2020 were analyzed. They trained and evaluated random forest (RF) machine learning classifiers to predict mechanical ventilation requirement and mortality using radiomic features extracted from patients` CXRs and explored deep learning (DL) approaches to the clinic.

The authors studied 143 radiomic features as those derived from the energy calculation. In this study, the authors performed an exploratory analysis of these radiomic features to determine their relative value in predicting clinical outcomes for patients with COVID-19. For each radiomic feature, statistics including measures of median, skewness, standard deviation, and kurtosis were calculated. These statistics and clinical factors, including expert scores and patient age/sex, were used for the construction of the classifier.

The authors demonstrated that radiomic features can provide information about which features of a patient's CXR are significant for making clinical predictions and may be more informative to a clinician than DL-only approaches. From the results achieved, it can be seen that radiomic characteristics play an interesting role in predicting results for COVID-19.

This work has some limitations. First, the study used reference CXRs that are likely to be inconsistent in the interval between COVID-19 infection and imaging. While this is representative of the clinical reality that patients receive baseline chest x-rays at different time points in the course of their illness, future studies could build better time-to-event prediction models using data with a distribution more uniform temporary, although in the context of the COVID-19 pandemic this is not easy to achieve given the urgency and the large number of cases to be evaluated. However, for the study of other respiratory infections, it could be taken into account.

It is also important to note that the two clinical outcomes studied in this paper are not independent or mutually exclusive; in general, a patient who requires mechanical ventilation is more likely to succumb to his illness than one who does not. On the other hand, the study took into account a limited number of clinical characteristics, so the inclusion of data on comorbidities such as a history of cancer, chronic obstructive pulmonary disease, hypertension, etc., could improve the performance of the proposed models. Finally, further validation would be needed to demonstrate the robustness of the classification models in the broader context of treating COVID-19 and other respiratory conditions at other hospitals and locations.

In Wendland. P. et al. (2023), the study was aimed to improve early risk stratification of hospitalized Covid-19 patients by predicting in-hospital mortality, transfer to intensive care unit (ICU) and mechanical ventilation from electronic health record data of the first 24 h after admission.

For model development they conducted an observational retrospective cohort study using electronic health record data from a hospital of medium level of care located in the federal state of Rhineland-Palatinate in the west of Germany initially collected for billing purposes. In the study were included 520 patients with a positive RT-PCR for SARS-CoV- 2 identified admitted to the hospital from March 2020 until December 2021.

For the model development and prior to any preprocessing steps authors performed a random train- test split using 80 % of the data as training set and 20 % as test set. They defined three Covid-19 associated endpoints:

1. Death during hospital stay, short “in-hospital mortality” 2. Admission to intensive care unit (ICU), short “transfer to the ICU” 3. Necessity for mechanical ventilation (al OPS beginning with “8–71“), short “mechanical ventilation”

In the experiment authors compared three supervised classifiers: Logistic regression (LR), Random forest (RF) and XGBoost. To select predictive features for each of these three model classes they employed 5-fold cross validation. For LR they performed forward-backward selection. For the random forest classifier and the XGBoost classifier they used the mean feature importance as a criterion for feature selection and in addition also trained these tree based classifiers using the same features as identified for the LR models. Further, for RF and XGBoost they performed a hyperparameter optimization on the training set. The model (including selected features) with the highest receiver operator characteristics area under the curve (ROC-AUC) averaged over the cross validation folds from the training data set was selected as the final model for the respective endpoint.

As result, the machine learning model predicts in-hospital mortality (AUC = 0,918), transfer to ICU (AUC = 0,821) and the need for mechanical ventilation (AUC = 0,654) from a few laboratory data of the first 24 h after admission. Models based on dichotomous features indicating whether a laboratory value exceeds or falls below a threshold perform nearly as good as models based on numerical features. Authors devise completely data-driven and interpretable machine-learning models for the prediction of in-hospital mortality, transfer to ICU and mechanical ventilation for hospitalized Covid-19 patients within 24 h after admission. Numerical values of CRP and blood sugar and dichotomous indicators for increased partial thromboplastin time (PTT) and gluta mic oxaloacetic transaminase (GOT) were amongst the best predictors.

The study only included patients admitted to hospitals from the beginning of the pandemic until the end of 2021. Due to the rapidly changing epidemiological circumstances of the pandemic the authors were not able to test the generalizability of the proposed models to a population, where the Omicron mutation is the dominating virus mutation. From 2020 until December 2021 the Wildtype, Alpha, Beta and Delta mutations were the dominating Covid-19 variants in Germany. The study do not included vital parameters, pre-existing comorbidities and vaccination. The cohort of 520 patients is relatively small. Further, the studied data was imbalanced, because only 50 to 90 patients with poor outcomes were observed for each respective endpoint.

Performance of the studied models

Given the conditions in which the COVID-19 pandemic develops, the studied models have been tested on a wide variety of conditions in terms of number of patients, availability, completeness, diversity and quality of images and clinical data, character of the study, ethnic and age diversity of the patients participating in the studies, balance of the classes taken into account for the training of the predictive models. This great diversity does not allow for a comparison based on equal performance conditions. However, we have made a summary of the results that they offer in terms of the most common metrics used in the studied reports. Table 4 offers a summary of these results for each of the studied models.

From the table it is possible to observe that the models that achieve the best performance in their respective evaluations do not depend on the type of outcome they predict, however they have a high dependence on the type of data with which they work. The models with the best results are those that use clinical data and CT images and their combination. On the other hand, models that combine CXR images and clinical data have shown acceptable performance.

Regarding the type of model, there is no clear differentiation between models based on classical methods of machine learning and those based on deep learning, however everything seems to indicate that in the case of machine learning algorithms, those based on regression have offered more accurate results.

Table 4. Summary of performance results of the studied models

Ref. in table3	Used data	Model type	AUC	Sensiblity.	Specifity.	Accuracy.(%)	c- score
(1)	CXR	ML +DL	0.794	0.722	0.732	-	-
(2)	CXR, clinical	ML	0.880	-	-	-	-
(3)	CT, clinical	ML	0.912	0.895	0.8015	-	-
(4)	Clinical	ML	0.895	-	-	-	-
(5)	CT	ML	0.916	0.808	0.969	-	-
(6)	CT, clinical	ML	0.910	0.894	0.872	-	-
(7)	Clinical	ML	0.950	0.764	0.919	-	-
(8)	CT, clinical	DL	0.942	-	-	-	-
(9)	CT, clinical	DL	0.800	-	-	-	-
(10)	CXR, clinical	DL	0.870	-	-	-	-
(11)	CXR	DL	0.743	-	-	-	-
(12)	CXR, clinical	DL	-	-	-	-	0.75
(13)	CXR	DL	-	-	-	86	-
(14)	Clinical	DL	-	0.935	-	95.97	-
(15)	Clinical	ML+ DL	0.797	-	-	-	-
(16)	CT, clinical	ML+ DL	0.864	-	-	-	-
(17)	Clinical	DL	-	-	93.55	95.97	-
(18)	CT, clinical	DL	0.880	-	-	-	-

Regarding the limitations of the models studied, in the vast majority of them they are related to the size of the training, validation and test sets, the imbalance of the samples of each class used in the training and the origin of the samples taken in one or a limited number of hospital institutions. Another limitation is related to the time of the pandemic in which the samples were taken, where most of them were considered in the first 2 years (2020, 2021) when the highest peaks of patients were recorded and more lethal strains of the disease (Alpha, Beta and Delta mutations).

None of the analyzed studies takes into account the post-vaccination scenario and its influence on the decrease in lethality, as well as the predominance of the more transmissible but less lethal variants of the Omicron strain.

General discussion

Based on the results of the study carried out, we have been able to determine that predictive models based on clinical data and their combination with CT or CXR images can achieve significant performance in the evolutionary prediction of the state of severity of the disease, death or the need for specialized treatments such as mechanical ventilation, hospital admission or ICU. The trend that is observed is the development of models based on Deep learning, however from the analysis developed we see that the methods based on regression have had a good performance.

The development of a predictive model under the conditions of Cuba would be in accordance with the availability of CXR images and non-standardized clinical data. The development of an investigation in this sense could be aimed at obtaining a prediction model for severe manifestations from COVID-19 by combining the classification of the degree of severity of lung involvement on CXR images, clinical data on comorbidities, and biographical data. The model would be developed based on the use of deep neural networks and taking as a starting point the developments achieved so far by our research team in the evaluation of the level of severity of lung involvement in CXR images by means of classification (Garea et al., 2021a, 2021b, 2022, 2023). The inclusion of radiomic features for the evaluation of the index of involvement of the pulmonary region could lead to an increase in the accuracy of the prediction of the degree of severity caused by the disease. As a baseline for comparison, we could then first take the work presented in Bae J. et al. (2020), where it would be convenient to implement the extraction of radiomic features, and their combination in the last three experimental schemes presented by the authors (P3, P4 and P5) with CNN deep learning models. Another method that can be taken as a baseline for comparison would be those presented in Olowolayemo, A. et al. (2023) based on deep learning.

ACKNOWLEDGEMENT

The VLIR-UOS has partially financed this research under the South Initiative: Toward Precision Medicine for the Prediction of Treatment Response to Covid-19 in Cuba (COVID-19 PROMPT).

CONCLUSIONS

In this work, a review study was presented with the objective of evaluating the most relevant findings published between 2020 and 2023 that could serve as a basis for the development of an own model adjusted to the conditions of Cuba.

As part of the methodology applied to the study, 18 relevant articles were identified, whose analysis demonstrated that predictive models based on clinical data and their combination with CT or CXR images can achieve significant performance in the evolutionary prediction of COVID-19 outcomes as death or the need for specialized treatments such as mechanical ventilation, hospital admission or ICU.

The development of an investigation in this sense could be aimed at obtaining a prediction model for severe manifestations from COVID-19 by combining the classification of the degree of severity of lung involvement on CXR images, clinical data on comorbidities, and biographical data.

As the main insufficiencies of the reviewed articles, it was found that the models that achieve the best performance in their respective evaluations do not depend on the type of outcome they predict, however they have a high dependence on the type of data with which they work.

The development of an investigation in this sense could be aimed at obtaining a prediction model for severe manifestations from COVID-19 by combining the classification of the degree of severity of lung involvement on CXR images, clinical data on comorbidities, and biographical data.

The model would be developed based on the use of deep neural networks and taking as a starting point the developments achieved in the evaluation of the level of severity of lung involvement in CXR images by means of classification.

REFERENCES

Albahri, O.S.; Zaidan, A.A.; Albahri, A.S.; Zaidan, B.B.; Abdulkareem, K.H.; Al-Qaysi, Z.T.; Alamoodi, A.H.; Aleesa, A.M.; Chyad, M.A.; Alesa, R.M.; et al.(2020). Systematic review of artificial intelligence techniques in the detection and classification of COVID-19 medical images in terms of evaluation and benchmarking: Taxonomy analysis, challenges, future solutions and methodological aspects. J. Infect. Public Health 2020, 13, 1381–1396.

Almotairi, K.H.; Hussein, A.M.; Abualigah, L.; Abujayyab, S.K.M.; Mahmoud, E.H.; Ghanem, B.O.; Gandomi, A.H.(2023). Impact of Artificial Intelligence on COVID-19 Pandemic: A Survey of Image Processing, Tracking of Disease, Prediction of Outcomes, and Computational Medicine. Big Data Cogn. Comput. 2023, 7, 11. https://doi.org/10.3390/bdcc7010011

Asteris P. G., Kokoris S., Gavriilaki E. et al (2023). Early prediction of COVID-19 outcome using artificial intelligence techniques and only five laboratory indices. Clinical Immunology 246 (2023) 109218. https://doi.org/10.1016/j.clim.2022.109218.

Bae, J., Kapse, S., Singh, G., Gattu, R., Ali, S., Shah, N., Marshall, C., Pierce. J., Phatak T., Gupta, A., Green, J., Madan, N., Prasanna, P.(2021). Predicting Mechanical Ventilation and Mortality in COVID-19 Using Radiomics and Deep Learning on

Chest Radiographs: A Multi- Institutional Study. Diagnostics (Basel). 2021 Sep 30;11(10):1812. doi: 10.3390/diagnostics11101812. PMID: 34679510; PMCID: PMC8535062.

Bansal, A.; Padappayil, R.P.; Garg, C.; Singal, A.; Gupta, M.; Klein, A. (2020). Utility of artificial intelligence amidst the COVID-19 pandemic: A review. J. Med. Syst. 2020, 44, 1–6.

Buttia, C., Llanaj, E., Raeisi-Dehkordi, H. et al. (2023). Prognostic models in COVID-19 infection that predict severity: a systematic review. Eur J Epidemiol 38, 355–372 (2023). https://doi.org/10.1007/s10654-023-00973-x

Duanmu, H., Ren, T., Li, H. et al. (2022). Deep learning of longitudinal chest X-ray and clinical variables predicts duration on ventilator and mortality in COVID-19 patients. BioMed Eng OnLine 21, 77. https://doi.org/10.1186/s12938-022-01045-z

Fu. Y, Zeng L., Huang P., Liao M. et al (2023). Severity-onset prediction of COVID-19 via artificial-intelligence analysis of multivariate factors. Heliyon. V 9, I 8, E18764, 2023. https://doi.org/10.1016/j.heliyon.2023.e18764

Gallo Marin, B., Aghagoli, G., Lavine, K., Yang, L., Siff, E.J., Chiang, S.S., Salazar-Mather, T.P., Dumenco, L., Savaria, M.C., Aung, S.N., Flanigan, T., Michelow, I.C. (2021).Predictors of COVID-19 severity: A literature review. Rev Med Virol. 2021 Jan; 31(1):1-10. doi: 10.1002/rmv.2146. Epub 2020 Jul 30. PMID: 32845042; PMCID: PMC7855377.

Garea Llano, E., Castellanos Loaces, H. A., Martínez Montes, E., & González Dalmau, E. (2021a). Estimación del Grado de Afectación Pulmonar por COVID-19 Mediante la Clasificación Supervisada de la Imagen de Rayos X. Revista Cubana De Transformación Digital, 2(3), 4–18. https://doi.org/10.5281/zenodo.5545908.

Garea-Llano, E., Castellanos-Loaces, H.A., Martinez-Montes, E., Gonzalez-Dalmau, E. (2021b). A Machine Learning Based Approach for Estimation of the Lung Affectation Degree in CXR Images of COVID-19 Patients. In: Hernández Heredia, Y., Milián Núñez, V., Ruiz Shulcloper, J. (eds) Progress in Artificial Intelligence and Pattern Recognition. IWAIPR 2021. Lecture Notes in Computer Science(), vol 13055. Springer, Cham. https://doi.org/10.1007/978-3-030-89691-1_2

Garea-Llano, E., Martinez-Montes, E., Gonzalez-Dalmaus, E. (2022). Affectation index and severity degree by COVID-19 in Chest X-ray images using artificial intelligence. Int Rob Auto J. 2022; 8(3):103–107. DOI: 10.15406/iratj.2022.08.00252.

Garea-Llano, E., Diaz-Berenguer, A., Sahli, H., Gonzalez-Dalmau, E. (2023). Chest X-Ray Imaging Severity Score of COVID-19 Pneumonia. In: Rodríguez-González, A.Y., Pérez- Espinosa, H., Martínez-Trinidad, J.F., Carrasco-Ochoa, J.A., Olvera-López, J.A. (eds) Pattern Recognition. MCPR 2023. Lecture Notes in Computer Science, vol 13902. Springer, Cham. https://doi.org/10.1007/978-3-031-33783-3_20.

Gourdeau, D., Potvin, O., Biem, J.H. et al. (2022). Deep learning of chest X-rays can predict mechanical ventilation outcome in ICU-admitted COVID-19 patients. Sci Rep 12, 6193 (2022). https://doi.org/10.1038/s41598-022-10136-9

Hassan, A.; Prasad, D.; Rani, S.; Alhassan, M. (2022). Gauging the Impact of Artificial Intelligence and Mathematical Modeling in Response to the COVID-19 Pandemic: A Systematic Review. BioMed Res. Int. 2022, 2022, 7731618.

Heidari, A., Jafari Navimipour, N., Unal, M. et al. (2022). Machine learning applications for COVID-19 outbreak management. Neural Comput & Applic 34, 15313–15348 (2022). https://doi.org/10.1007/s00521-022-07424-w.

Heo, J., Han, D., Kim, H. J., Kim, D., Lee, Y. K., Lim, D. Seog, W. (2021). Prediction of patients requiring intensive care for COVID-19: development and validation of an integerbased score using data from Centers for Disease Control and Prevention of South Korea. J Intensive Care, 9(1), 16. doi:10.1186/s40560-021-00527-x

Ho, T. T., Park, J., Kim, T., Park, B., Lee, J., Kim, J. Y., Choi, S. (2021). Deep Learning Models for Predicting Severe Progression in COVID-19-Infected Patients: Retrospective Study. JMIR Med Inform, 9(1), e24973. doi:10.2196/24973

Jalaber, C.; Lapotre, T.; Morcet-Delattre, T.; Ribet, F.; Jouneau, S.; Lederlin, M.(2020). Chest CT in COVID-19 pneumonia: A review of current knowledge. Diagn. Interv. Imaging 2020, 101, 431–437.

Jamshidi, M.B.; Roshani, S.; Talla, J.; Lalbakhsh, A.; Peroutka, Z.; Roshani, S.; Sabet, A.; Dehghani, M.; Lotfi, S.; Hadjilooei, F.; et al.(2022). A Review on Potentials of Artificial Intelligence Approaches to Forecasting COVID-19 Spreading. AI 2022, 3, 493–511.

Kumar, A.; Gupta, P.K.; Srivastava, A. (2020). A review of modern technologies for tackling COVID-19 pandemic. Diabetes Metab. Syndr. Clin. Res. Rev. 2020, 14, 569–573.

Liang, W., Liang, H., Ou, L., Chen, B., Chen, A., Li, C. China Medical Treatment Expert Group for, C. (2020). Development and Validation of a Clinical Risk Score to Predict the Occurrence of Critical Illness in Hospitalized Patients With COVID-19. JAMA Intern Med, 180(8), 1081-1089. doi:10.1001/jamainternmed.2020.2033 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7218676/)

Lieveld, A. W. E., Azijli, K., Teunissen, B. P., van Haaften, R. M., Kootte, R. S., van den Berk, I.

A. H., Nanayakkara, P. W. B. (2021). Chest CT in COVID-19 at the ED: Validation of the COVID-19 Reporting and Data System (CO-RADS) and CT Severity Score: A Prospective, Multicenter,Observational Study. Chest, 159(3), 1126-1135. doi:10.1016/j.chest.2020.11.026.

Matsumoto, T., Walston, S.L., Walston, M. et al. (2022). Deep Learning–Based Time-to-Death Prediction Model for COVID-19 Patients Using Clinical Data and Chest Radiographs. J Digit Imaging (2022). https://doi.org/10.1007/s10278-022-00691-y

Mbunge, E.; Akinnuwesi, B.; Fashoto, S.G.; Metfula, A.S.; Mashwama, P. (2021). A critical review of emerging technologies for tackling COVID-19 pandemic. Hum. Behav. Emerg. Technol. 2021, 3, 25–39.

Moher D, Liberati A, Tetzlaff J, et al. (2009). Reprint--preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Phys Ther 2009;89:873−80.

Naudé,W.(2020). Artificial Intelligence against COVID-19: An Early Review; Institute of Labor Economics: Bonn, Germany, 2020.

Okoli C (2015). A guide to conducting a standalone systematic literature review. Commun Assoc Inf Syst 2015;37:43.

Olowolayemo, A., Yasin, M., Raashid S. M. (2023). Predicting Mortality Risk of Covid-19 Patients Using Chest X-Rays. International Journal on Perceptive and Cognitive Computing, 9(1), 33–43.

Ortiz, A., Trivedi, A., Desbiens, J. et al. (2022). Effective deep learning approaches for predicting COVID-19 outcomes from chest computed tomography volumes. Sci Rep 12, 1716 (2022). https://doi.org/10.1038/s41598-022-05532-0

Panagiotis G. Asteris, Styliani Kokoris, Eleni Gavriilaki, et al (2023). Early prediction of COVID- 19 outcome using artificial intelligence techniques and only five laboratory indices, Clinical Immunology, Volume 246, 2023, 109218, ISSN 1521-6616, https://doi.org/10.1016/j.clim.2022.109218.

Portal-Diaz, J.A., Lovelle-Enríquez, O., Perez-Diaz, M. et al. (2022). New patch-based strategy for COVID-19 automatic identification using chest x-ray images. Health Technol. 12, 1117–1132 (2022). https://doi.org/10.1007/s12553-022-00704-4

Quiroz, J. C., Feng, Y. Z., Cheng, Z. Y., Rezazadegan, D., Chen, P. K., Lin, Q. T., Cai, X. R. (2021). Development and Validation of a Machine Learning Approach for Automated Severity Assessment of COVID-19 Based on Clinical and Imaging Data: Retrospective Study. JMIR Med Inform, 9(2), e24572. doi:10.2196/24572 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7879715/).

Rasheed, J.; Jamil, A.; Hameed, A.A.; Al-Turjman, F.; Rasheed, A. (2021). COVID-19 in the Age of Artificial Intelligence: A Comprehensive Review. Interdiscip. Sci. Comput. Life Sci. 2021, 13, 153–175

Rahimi, I.; Chen, F.; Gandomi, A.H. (2021). A review on COVID-19 forecasting models. Neural Comput. Appl. 2021, 1–11.

Shi, F.; Wang, J.; Shi, J.; Wu, Z.; Wang, Q.; Tang, Z.; He, K.; Shi, Y.; Shen, D. (2020). Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for COVID-19. IEEE Rev. Biomed. Eng. 2020, 14, 4–15

Swapnarekha, H.; Behera, H.S.; Nayak, J.; Naik, B. (2020). Role of intelligent computing in COVID-19 prognosis: A state-of-the-art review. Chaos Solitons Fractals 2020, 138, 109947.

Tanboga, I. H., Canpolat, U., Cetin, E. H. O., Kundi, H., Celik, O., Caglayan, M., . . .Topaloglu, S. (2021). Development and validation of clinical prediction model to estimate the probability of death in hospitalized patients with COVID-19: Insights from a nationwide database. J Med Virol, 93(5), 3015-3022. doi:10.1002/jmv.26844.

Tranfield D, Denyer D, Smart P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. Br J Manag 2003;14:207−22.

Verzellesi L, Botti A, Bertolini M, Trojani V, Carlini G, Nitrosi A, Monelli F, Besutti G, Castellani G, Remondini D, et al. Machine and Deep Learning Algorithms for COVID-19 Mortality Prediction Using Clinical and Radiomic Features. Electronics. 2023; 12(18):3878. https://doi.org/10.3390/electronics12183878

Wendland. P., Schmitt V., Zimmermann J , Hager L., Gopel S., Schenkel-Hager C., Kschischo M (2023). Machine learning models for predicting severe COVID-19 outcomes in hospitals. Informatics in Medicine Unlocked 37 (2023) 101188. https://doi.org/10.1016/j.imu.2023.101188

Zhang, B., Liu, Q., Zhang, X., Liu, S., Chen, W., You, J., Zhang, S. (2020). Clinical Utility of a Nomogram for Predicting 30-Days Poor Outcome in Hospitalized Patients With COVID- 19: Multicenter External Validation and Decision Curve Analysis. Front Med (Lausanne), 7, 590460. doi:10.3389/fmed.2020.590460