

## **Supplementary material**

### **Systematic review and critical appraisal of prediction models for diagnosis and prognosis of COVID-19 infection**

#### **Contents:**

Search strings

Data extraction

Supplementary Table 1. Overview of the primary datasets used in included studies.

Supplementary Table 2. Overview of modelling techniques for diagnosis and prognosis of COVID-19 infection.

## **Supplementary material: search strings**

Search string bioRxiv, medRxiv, arXiv: ncov OR corona OR wuhan OR COVID

Search string PubMed: (covid-19[All Fields]) OR ("Wuhan coronavirus" [Supplementary Concept] OR "2019 ncov"[tiab] OR ("novel coronavirus"[tiab] OR "new coronavirus"[tiab]) AND (wuhan[tiab] OR 2019[tiab])) OR 2019-nCoV[All Fields] OR (wuhan[tiab] AND coronavirus[tiab]))))

Search string EMBASE via Ovid: undisclosed (<https://ispmbern.github.io/covid-19/living-review/index.html>)

## **Supplementary material: data extraction**

Data were extracted on the following items:

- Population (China, other).
- Intended timing of model use (screening of patients, ICU admission, etc.).
- Setting (inpatients, outpatients, suspected cases).
- Participants: study design, recruitment method, number of centres, inclusion criteria, exclusion criteria, patient age (mean and standard deviation or median and interquartile range), patient sex (n males and percentage).
- Predictors: list of candidate predictors, number of candidate predictors, number of additional degrees of freedom for candidate predictors (e.g. for categorical variables with more than two categories or for modelling continuous variables non-linearly), predictors in the final model, number of predictors in the final model, number of additional degrees of freedom in the final model.
- Outcome: definition of outcome, timing of outcome.
- Analysis:
  - Total number of participants, total number of participants with the outcome.
  - Total number of participants with any missing predictor or outcome values.
  - Method for handling missing data.
  - Method for prediction model development (logistic regression, Cox regression, neural network, tree-based, etc.).
  - Method for selection of predictors during multivariable analysis and the criteria used (e.g., p-value used for selection).

- Handling of categorical and continuous variables.
  - Method(s) for validation (e.g. apparent, internal or external) and for optimism adjustment.
  - Performance measures (calibration, discrimination, other) resulting from validation.
  - Model presentation (coefficients and confidence intervals, final model, alternative presentation formats of the model such as a web-based tool).
- Standard signalling questions and items to assess risk of bias according to PROBAST (Moons, Wolff, et al.), on four domains:
    - participants
    - predictors
    - outcome
    - analysis

If risk of bias was high in at least one of the subdomains, overall risk of bias was judged to be high, as per PROBAST guidance (Moons, Wolff, et al.).

**Supplementary Table 1.** Overview of the primary datasets used in included studies.

	Authors	Description of dataset	Study design	Age mean $\pm$ se, median (IQR), or range	Sex (% male)
<b>Hospital admission in general population</b>					
	DeCaprio, Gartner, et al.	Medicare claims data 2015 to 2016	Administrative records	$\geq 18$	Unclear
<b>Diagnosis</b>					
	Song, Xu, et al.	Zhijiang District of the First Affiliated Hospital of Zhejiang University School of Medicine	Unclear	38 (IQR 30 to 55)	49%
	Yu, Shao, et al.	Wuhan's Children's Hospital, February 1 to March 3	Retrospective cohort	0 to 15 (range)	61%
<b>Diagnostic imaging</b>					
	Gozes, Frid-Adar, et al.	COVID cases: Wenzhou hospital (China), controls: chainz.cn registry, El Camino Hospital, California, LIDC registry (7 USA hospitals)	Nonnested case-control	Unclear	Unclear
	Chen, Wu, et al.*	Renmin Hospital of Wuhan University	Unclear	Cases: 52 (IQR 38 to 69) Controls: 48 (IQR 35 to 55)	Cases: 35% Controls: 56%
	Shan, Gao, et al.	Shanghai Public Health Clinical Center and other centers in Shanghai	Unclear	Unclear	High
	Wang, Kang, et al.	Xi'an Jiaotong University First Affiliated Hospital, Nanchang University First Hospital and Xi'an No.8 Hospital of Xi'an Medical College	Unclear	Unclear	Unclear
	Xu, Jiang, et al.	First Affiliated Hospital of Zhejiang University, No. 6 People's Hospital of Wenzhou, No. 1 People's Hospital of Wenling, January 19 to February 14	Nonnested case control	Unclear	Unclear
	Ying, Zheng, et al.	Renmin Hospital of Wuhan University, Third Affiliated Hospital, Sun Yat-Sen Memorial Hospital	Unclear	Unclear	Unclear
<b>Prognosis</b>					
	Caramelo, Ferreira, et al.	Simulated based on data from Wuhan, China, December 8 to February 11	Simulation	Unclear	41%
	Lu, Hu, et al.	Wuhan Hankou Hospital, January 21 to February 5	Retrospective cohort	Unclear	44%
	Qi, Jiang, et al.	5 hospitals from Ankang, Lishui, Zhenjiang, Lanzhou, Linxia between January 23 to February 8; date of last follow-up February 20	Prospective cohort	38 (IQR 26 to 47)	55%
	Xie, Hungerford, et al.	Tongji Hospital and Jinyintan Hospital, admitted between January and February	Retrospective cohort	Tongji: 65 (IQR 54 to 73) Jinyintan: 56 (IQR 47 to 68)	Tongji: 52% Jinyintan: 35%
	Yan, Zhang, et al.	Tongji Hospital Wuhan, January 10 to February 18	Retrospective cohort	59 $\pm$ 16	59%
	Yuan, Yin, et al.	Hubei Public Health Clinical Center; central Hospital of Wuhan, January 1 to January 25	Retrospective cohort	60 (IQR 47 to 69)	44%

\* The study also included a prospective validation cohort of 27 consecutive patients (11 with COVID-19 pneumonia)

**Supplementary Table 2.** Overview of modelling techniques for diagnosis and prognosis of COVID-19 infection.

Study	Outcome	Predictors in final model	Modelling technique
<b>Hospital admission in general population</b>			
Decaprio, Gartner, et al.	Hospital admission for COVID-19 pneumonia (proxy events)*1	Age; sex; number of previous hospital admissions; 11 diagnostic features; interactions between age and diagnostic features	Logistic regression
Decaprio, Gartner, et al.	Hospital admission for COVID-19 pneumonia (proxy events)*1	Age and 500+ features related to diagnosis history	Tree-based (XGBoost)
Decaprio, Gartner, et al.	Hospital admission for COVID-19 pneumonia (proxy events)*1	500+ undisclosed features, including age, diagnostic history, social determinants of health, Charlson comorbidity index	Tree-based (XGBoost)
<b>Diagnosis</b>			
Song, Xu, et al.	COVID-19 diagnosis	Fever; history of close contact; signs of pneumonia on CT; neutrofil-to-lymphocyte ratio; highest body temperature; sex; (age, meaningful respiratory syndromes)	Logistic regression
Yu, Shao, et al.	Severe disease (yes/no) defined based on clinical symptoms	Direct Bilirubin; Alaninetransaminase	Tree-based (decision tree)
<b>Diagnostic imaging</b>			
Gozes, Frid-Adar, et al.	COVID-19 diagnosis	Not applicable	Neural net (deep convolutional neural network)
Chen, Wu, et al.	COVID-19 pneumonia	Not applicable	Neural net (Unet ++)
Xu, Jiang, et al.	COVID-19 diagnosis	Not applicable	Neural net (convolutional neural network)
Wang, Kang, et al.	COVID-19 diagnosis	Not applicable	Neural net (convolutional neural network)
Ying, Zheng, et al.	diagnosis of COVID-19 vs healthy controls	Not applicable	Neural net (DRENet vs. VGG16, DenseNet, ResNet for comparison)
Ying, Zheng, et al.	diagnosis of COVID-19 vs bacterial pneumonia	Not applicable	Neural net (DRENet vs. VGG16, DenseNet, ResNet for comparison)
Shan, Gao, et al.	Segmentation and quantification of infection regions in lung from chest CT scans.	Not applicable	Neural net (VB-Net)
<b>Prognosis</b>			
Caramelo, Ferreira, et al.	Mortality (period unspecified) *2	Age; sex; presence of any comorbidity (hypertension, diabetes, cardiovascular disease, chronic respiratory disease, cancer) *3	Logistic regression
Lu, Hu, et al.	Mortality (12 day)	Age; C-reactive protein	Cox regression
Qi, Jiang, et al.	Hospital stay >10 days	6 features derived from CT images *3	Logistic regression
Qi, Jiang, et al.	Hospital stay >10 days	6 features derived from CT images *3	Tree-based (random forest)
Xie, Hungerford, et al.	Mortality (in hospital)	Age, LDH, lymphocyte count, SPO2	Logistic regression

Yan, Zhang, et al.	Mortality (period unspecified)	Lactic dehydrogenase; lymphocyte count; high-sensitivity C-reactive protein	Tree-based (XGBoost)
Yuan, Yin, et al.	Mortality (period unspecified)	Clinical scorings of CT images (zone, left/right, location, attenuation, distribtion of affected parenchyma)	Ad hoc scoring rule

\*1 Proxy events used: pneumonia (except from TB), influenza, acute bronchitis, or other specified upper respiratory infections (no COVID-19 pneumonia cases in data).

\*2 Outcome and predictor data were simulated.

\*3 Wavelet-HLH\_gldm\_SmallDependenceLowGrayLevelEmphasis, wavelet-LHH\_gldm\_Correlation, wavelet-LHL\_glszm\_GrayLevelVariance, wavelet-LLH\_glszm\_SizeZoneNonUniformityNormalized, wavelet-LLH\_glszm\_SmallAreaEmphasis, wavelet-LLH\_gldm\_Correlation.