



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



## Using artificial intelligence-based models to predict the risk of mucormycosis among COVID-19 survivors: An experience from a public hospital in India

Shabbir Syed-Abdul<sup>a,b,c</sup>, A. Shoban Babu<sup>d,\*</sup>, Raja Shekhar Bellamkonda<sup>e</sup>, Ramaiah Itumalla<sup>f</sup>, GVRK Acharyulu<sup>e</sup>, Surya Krishnamurthy<sup>g</sup>, Y. Venkat Santosh Ramana<sup>h</sup>, Naresh Mogilicharla<sup>d</sup>, Shwetambara Malwade<sup>b</sup>, Yu-Chuan Li Jack<sup>a,b,\*\*</sup>

<sup>a</sup> Graduate Institute of Biomedical Informatics, College of Medical Science and Technology, Taipei Medical University, 5F., No. 172-1, Sec. 2, Keelung Rd. Da'an, Taipei 106, Taiwan

<sup>b</sup> International Center for Health Information Technology, College of Medical Science and Technology, Taipei Medical University, Taipei, Taiwan

<sup>c</sup> School of Gerontology and Health Management, College of Nursing, Taipei Medical University, Taipei, Taiwan

<sup>d</sup> Department of ENT, Gandhi Medical College and Hospital, Musheerabad, Secunderabad, Telangana 500003, India

<sup>e</sup> School of Management Studies, University of Hyderabad, Hyderabad, Telangana, India

<sup>f</sup> Department of Health Management, College of Public Health and Health Informatics, University of Hail, Hail, Saudi Arabia

<sup>g</sup> Data Scientist, iQGateway, Bengaluru, Karnataka, India

<sup>h</sup> Department of Hospital Administration, Gandhi Hospital, Secunderabad, Telangana, India

### ARTICLE INFO

#### Article history:

Accepted 10 December 2021

Available online 23 December 2021

#### Keywords:

Artificial intelligence

Coronavirus

COVID-19

Fungal infection

India

Mucormycosis

### SUMMARY

**Introduction:** India reported a severe public health challenge not only due to the COVID-19 outbreak but also the increasing number of associated mucormycosis cases since 2021. This study aimed at developing artificial intelligence based models to predict the risk of mucormycosis among the patients at the time of discharge from hospital.

**Methods:** The dataset included of 1229 COVID-19 positive patients, and additional 214 inpatients, COVID-19 positive as well as infected with mucormycosis. We used logistic regression, decision tree and random forest and the extreme gradient boosting algorithm. All our models were evaluated with 5-fold validation to derive a reliable estimate of the model error.

**Results:** The logistic regression, XGBoost and random forest performed equally well with AUROC 95.0, 94.0, and 94.0 respectively. The best accuracy and precision (PPV) were  $0.91 \pm 0.026$  and  $0.67 \pm 0.0526$ , respectively achieved by XGBoost, followed by logistic regression. This study also determined top five variables namely obesity, anosmia, de novo diabetes, myalgia, and nasal discharge, which showed positive impact towards the risk of mucormycosis.

**Conclusion:** The developed model has the potential to predict the patients at high risk and thus, consequently initiating preventive care or aiding in early detection of mucormycosis infection. Thus, this study, holds potential for early treatment and better management of patients suffering from COVID-19 associated mucormycosis.

© 2022 The British Infection Association. Published by Elsevier Ltd. All rights reserved.

\* Corresponding author. Shoban Babu at Department of ENT, Gandhi Medical College and Hospital, Musheerabad, Secunderabad, Telangana 500003, India

\*\* Corresponding author at: Graduate Institute of Biomedical Informatics, College of Medical Science and Technology, Taipei Medical University, 5F., No. 172-1, Sec. 2, Keelung Rd. Da'an, Taipei 106, Taiwan.

E-mail addresses: [shobhanaachi@gmail.com](mailto:shobhanaachi@gmail.com) (A.S. Babu), [jack@tmu.edu.tw](mailto:jack@tmu.edu.tw) (Y.-C. Li Jack).

### Introduction

Mucormycosis (previously called zygomycosis) is a life-threatening but rare fungal infection caused by a group of molds called mucormycetes.<sup>1</sup> It is an opportunistic infection characterized by sinusitis and orbital cellulitis that penetrates the arteries to produce thrombosis of the ophthalmic and internal carotid arteries, and later invades the veins and lymphatics.<sup>2</sup> India, the second most populated country in the world has been facing public health challenge due to the outbreak of the COVID-19.<sup>3</sup> As of July 8th 2021, in India 4,60,704 active COVID-19 cases and

4,05,028 deaths were reported.<sup>4</sup> At this crucial juncture, increasing number of the mucormycosis cases in country are alarming the policy makers and healthcare experts emphasizing the great necessity of preventive measures to detect and treat the patients with mucormycosis. Until 28th June 2021, India reported 40,845 cases of mucormycosis, and considering hundreds of thousands of COVID-19 active cases this number is expected to increase significantly. As of 21st July 2021, India recorded more than 4300 deaths from COVID-19 related mucormycosis.<sup>5</sup> Difficult diagnosis and relapse of the disease are among the factors hampering treatment success. Therefore, this study is aimed to develop artificial intelligence (AI)-based models to predict the risk of mucormycosis among the COVID-19 patients at the time of their discharge from hospital.

## Methods

### Data pre-processing

This single-center retrospective study was conducted at a tertiary public hospital, designated for COVID-19 and mucormycosis treatment at the Telangana state in India. The dataset comprised of 1229 COVID-19 positive patients, and additional 214 inpatients, those who were tested positive for COVID-19 and subsequently infected with mucormycosis. All these patients were hospitalized and received treatment between 19th March 2021 and 30th June 2021. Since the patient group with mucormycosis included only those in the age group between 30 and 75 years, we filtered all the patients in the dataset to that age group to ensure that the models are unbiased. The original dataset from medical records contains 74 features from which relevant variables were filtered and encoded for predictive modeling. We manually removed all features that could contribute to data leak such as symptoms of mucormycosis, time related features etc. The categorical variables were binary/one-hot encoded. After encoding, 35 features were input to our prediction models. The final set of features used for modeling is as below:

- Nominal features:
  - Sex: male, female;
  - Place of Residence: urban, rural;
  - Duration of diabetes: chronic, de novo, none;
  - Covid-19 symptoms: severity, asymptomatic, throat pain, anosmia, nasal obstruction, loss of taste, headache, nasal discharge, myalgia, cough, fever, shortness of breath;
  - Comorbidities: dialysis, obesity (based on physician's assessment of Body Mass Index), lung disease, hypothyroid, heart disease, renal disease, hypertension, diabetes, none;
  - Medications: antivirals, mineral supplements, anti-parasitics, blood thinners, steroids, vitamin supplements, broad spectrum antibiotics;
- Ordinal features:
  - Covid-19 severity: mild, moderate, severe (determined as patient assessment done by the physician based on the guidelines defined by the government of India<sup>6</sup>);
- Numerical features:
  - Age: 30 to 75;

### Modeling

We used logistic regression, decision tree and random forest from the scikit-learn<sup>7</sup> library and the extreme gradient boosting algorithm from the XGBoost<sup>8</sup> library. To prevent overfitting and selection bias from our skewed dataset we tuned the models to generalize well on unseen data. Regularization was adapted (L2 for logistic regression and XGBoost), wherever applicable and the model parameters were chosen to reduce the model complexity (limiting the depth of the trees to 5, reducing XGBoost to 100 trees and increasing the forest to 500 trees). We used class weights in all our

models to handle the class imbalance. We evaluated all our models with 5-fold validation to derive a reliable estimate of the model error.

## Results

The AUROC, recall, precision and accuracy from the different models used with the 5-fold validation are shown in Fig. 1. The logistic regression, XGBoost and random forest performed equally well with AUROC 95.0, 94.0, and 94.0 respectively. The best accuracy and precision (PPV) were  $0.91 \pm 0.026$  and  $0.67 \pm 0.0526$ , respectively achieved by XGBoost, followed by logistic regression. Fig. 2 shows box plot of coefficients for all the features across the 5 folds indicating variables having positive and negative impact on the risk of mucormycosis.

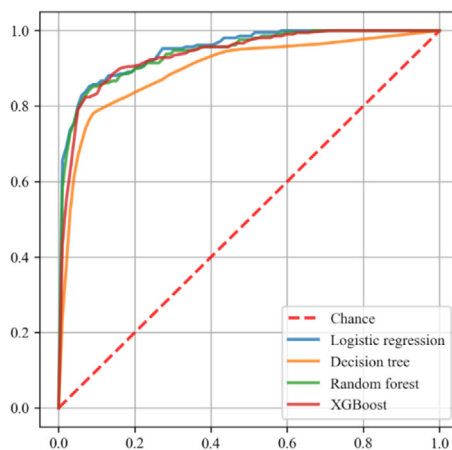
## Discussion

To the best of our knowledge, this is the first study to develop an artificial intelligence-based model to predict the risk of mucormycosis among COVID-19 survivors. Our developed models achieved a high AUROC of  $0.95 \pm 0.023$ . Moreover, this study also indicated the probable risk factors having predisposition towards mucormycosis. Features such as obesity, anosmia, de novo diabetes, myalgia, and nasal discharge were the top five variables showed positive impact towards risk of mucormycosis. Previous research by Mahalakshmi et al., Singh et al. and Pakdel et al. also indicated that uncontrolled diabetes can be a strong pre-disposing factor for post COVID mucormycosis.<sup>9–11</sup> Some studies also pointed out other risk factors for mucormycosis, such as treatment with corticosteroids, those who underwent organ transplantation, or those suffering from hematological malignancies.<sup>12,9,13</sup> Our study was the first one to reveal de novo diabetes obesity, anosmia, myalgia and nasal discharge as the factors associated with development of mucormycosis in COVID-19 patients.

On the other hand, absence of diabetes mellitus, under antiparasitic medications, nasal obstruction were the symptoms negatively associated with the mucormycosis. These models can be used to predict the risk and aid healthcare providers to take preventive care for the patients who are at high risk to develop mucormycosis. Mortality from the disease is relatively high and early diagnosis along with preventive measures would be the key to survival.<sup>10</sup> Our study, indicates potential for early detection, early treatment and better management of patients suffering from COVID-19 associated mucormycosis. The AI models can not only detect high risk patients, but also indicate the possible pre-disposing factors that need to be taken care of post COVID-19 treatment.

However, there are some limitations for generalization of these models, which need further validations with dataset from other states of India. This study was conducted in a public hospital in India, overcrowded with patients, operated with limited infrastructures and resources. Therefore, the prediction model lacks inclusion of granular patient related information such as diabetic ketoacidosis, hematological malignancy and solid organ tumors, neutropenia, and deferoxamine therapy. Despite this, the study's model has the potential to predict the patients at high risk and thus, consequently initiating preventive care or aiding in early detection of mucormycosis infection. Moreover, including a dataset for a longer time span would enable more accurate results with a larger dataset.

Future studies could include predicting the number of days after discharge, when a patient has a possibility of developing mucormycosis.



	AUROC	Recall	Precision	Accuracy
Logistic regression	<b>0.95 ± 0.023</b>	<b>0.87 ± 0.057</b>	0.64 ± 0.052	0.90 ± 0.019
XGBoost	0.94 ± 0.0172	0.85 ± 0.049	<b>0.67 ± 0.082</b>	<b>0.91 ± 0.026</b>
Random forest	0.94 ± 0.018	0.86 ± 0.078	0.61 ± 0.052	0.88 ± 0.021
Decision tree	0.90 ± 0.012	0.80 ± 0.07	0.60 ± 0.06	0.88 ± 0.022

Fig. 1. Area under receiver operating characteristic curve (AUROC) for the models developed in the study.

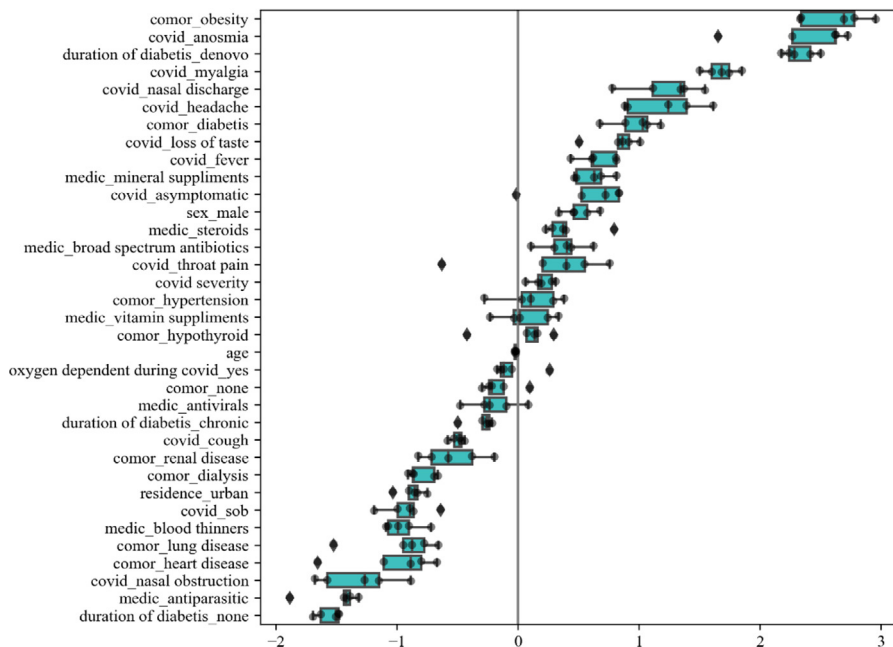


Fig. 2. Box plot of coefficients for all the features across the 5 folds showing positive and negative impact on the risk of mucormycosis.

**Acknowledgements**

This work is supported in part by Ministry of Science and Technology, Taiwan [Grant Nos. 108–2221-E-038–013, 110–2923-E-038 –001 -MY3]; Taipei Medical University, Taiwan [grant numbers 108–3805–009–110, 109–3800–020–400]; Ministry of Education, Taiwan [Grant No. 108–6604–002–400]; Wanfang hospital, Taiwan [Grant No. 106TMU-WFH-01–4].

**References**

- do Monte Junior ES, Dos Santos MEL, Ribeiro IB, de Oliveira Luz G, Baba ER, Hirsch BS, et al. Rare and fatal gastrointestinal mucormycosis (Zygomycosis) in a COVID-19 patient: a case report. *Clin Endosc* 2020;**53**(6):746.
- Baker RD. Mucormycosis—a new disease? *J Am Med Assoc* 1957;**163**(10):805–8.
- Jha AK, Jha R. India’s Response to COVID-19 Crisis. *Indian Econ J* 2020;**68**(3):341–51.
- Government of India COVID-19 cases across India 2021 [cited 2021 8 July]. Available from: <https://www.mygov.in/covid-19>.
- BBC News Mucormycosis: India records more than 4,300 'black fungus' deaths 2021 [cited 2021 21 July]. Available from: <https://www.bbc.com/news/world-asia-india-57897682>.
- Directorate General of Health Services M, GOI, Comprehensive guidelines for management of COVID-19 patients 2021 [cited 2021 10 December]. Available from: <https://www.dghs.gov.in/WriteReadData/News/202105270436027770348ComprehensiveGuidelinesforManagementofCOVID-1927May2021DteGHS.pdf>.
- Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res* 2011;**12**:2825–30.
- Chen T, Guestrin C Xgboost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016.
- Mahalaxmi I, Jayaramayya K, Venkatesan D, Subramaniam MD, Renu K, Vijayakumar P, et al. Mucormycosis: an opportunistic pathogen during COVID-19. *Environ Res* 2021;**201**:111643.

10. Singh Y, Ganesh V, Kumar S, Patel N, Aggarwala R, Soni KD, et al. Coronavirus disease-associated mucormycosis from a tertiary care hospital in India: a case series. *Cureus* 2021;**13**(7).
11. Pakdel F, Ahmadikia K, Salehi M, Tabari A, Jafari R, Mehrparvar G, et al. Mucormycosis in patients with COVID-19: a cross-sectional descriptive multicentre study from Iran. *Mycoses* 2021.
12. Ibrahim AS, Spellberg B, Walsh TJ, et al. Pathogenesis of Mucormycosis. *Clin Infect Dis* 2012;**54**(suppl\_1):S16–22.
13. Mishra Y, Prashar M, Sharma D, Kumar VP, Tilak T. Diabetes, COVID 19 and mucormycosis: clinical spectrum and outcome in a tertiary care medical center in Western India. *Diabetes Metab Syndr Clin Res Rev* 2021:102196.