Salud, Ciencia y Tecnología. 2024; 4:.929
doi: 10.56294/saludcyt2024.929

SALUD, CIENCIA Y TECNOLOGÍA

Check for updates

ORIGINAL

Novel Machine Learning Approach for Forecasting the Possibility of Recurrence in Community-Acquired Pneumonia

Novedoso enfoque de aprendizaje automático para predecir la posibilidad de recurrencia en la neumonía adquirida en la comunidad

Dnyaneshwar Prabhakar Bawane[1], Raja Ramalingam[2], M. Gopi[3], Vaibhav Kaushik[4], Prakhar Goyal[5], Yuvraj Parmar[6]

[1]Yeshwantrao Chavan College of Engineering, Department of Applied Mathematics and Humanitie, Nagpur, India.
[2]Mahatma Gandhi Medical College and Research Institute, Sri Balaji Vidyapeeth (deemed to be university), Department of pediatrics, Pondicherry, India.
[3]Dr.M.G.R Educational and Research Institute, Department of physiology, Chennai, India.
[4]Chitkara University, Department of Research Impact and Outcome, Punjab, India.
[5]Quantum University, Uttarakhand, India.
[6]Chitkara University, Department of Research and Development, SOLAN, India.

Cite as: Prabhakar Bawane D, Ramalingam R, Gopi M, Kaushik V, Goyal P, Parmar Y. Novel Machine Learning Approach for Forecasting the Possibility of Recurrence in Community-Acquired Pneumonia. Salud, Ciencia y Tecnología. 2024; 4:.929. https://doi.org/10.56294/saludcyt2024.929

Submitted: 28-12-2023 **Revised:** 20-04-2024 **Accepted:** 23-08-2024 **Published:** 24-08-2024

Editor: Dr. William Castillo-González

ABSTRACT

Antipsychotic drug usage is known to increase the risk of pneumonia, despite the fact that medications are commonly used to treat schizophrenia. By utilize machine learning (ML) to assemble a model for predicting community-acquired pneumonia (CAP) in schizophrenia patient. The beginning of pneumonia was predicted by eleven factors including gender, age, clozapine usage, drug-drug interactions, dose, length treatment, coughing, and changes in neutrophil and leukocyte counts, blood sugar levels, and body weight. To create the prediction models employed in this work, seven ML techniques were utilized in the study. To assess the overall performance of the model, we employed accuracy, sensitivity, specificity. In comparison to other seven ML methods, RF and DT have results the improved forecasting efficiency. Six other key risk variables were also found, including dose, clozapine usage, medication duration, change in neutrophil or leukocyte count, and drug-drug interaction. Our prediction model could be a helpful device for doctors caring for schizophrenic patients, even though these individuals still run the risk of pneumonia while using anti-psychotic medications.

Keywords: Schizophrenia; Community-Acquired Pneumonia (CAP); Co-Morbidities; Clozapine; Drug-Drug Interaction; Machine Learning (ML).

RESUMEN

Se sabe que el uso de fármacos antipsicóticos aumenta el riesgo de neumonía, a pesar de que los medicamentos se utilizan habitualmente para tratar la esquizofrenia. Utilizando el aprendizaje automático (ML) para montar un modelo de predicción de la neumonía adquirida en la comunidad (NAC) en pacientes con esquizofrenia. El inicio de la neumonía se predijo a partir de once factores, como el sexo, la edad, el uso de clozapina, las interacciones entre fármacos, la dosis, la duración del tratamiento, la tos y los cambios en los recuentos de neutrófilos y leucocitos, los niveles de azúcar en sangre y el peso corporal. Para crear los modelos de predicción empleados en este trabajo, se utilizaron en el estudio siete técnicas de ML. Para evaluar el rendimiento global del modelo, se emplearon la precisión, la sensibilidad y la especificidad. En comparación con los otros siete métodos ML, RF y DT presentan resultados la eficacia de predicción mejorada. También se encontraron otras seis variables clave de riesgo, incluyendo la dosis, el uso de clozapina, la duración de la medicación, el cambio en el recuento de neutrófilos o leucocitos, y la interacción fármaco-fármaco. Nuestro modelo de predicción podría ser un dispositivo útil para los médicos que atienden a pacientes esquizofrénicos, aunque estos individuos siguen corriendo el riesgo de neumonía mientras utilizan medicamentos antipsicóticos.

Palabras clave: Esquizofrenia; Neumonía Adquirida en la Comunidad (NAC); Comorbilidades; Clozapina; Interacción Fármaco-Fármaco; Aprendizaje Automático (ML).

INTRODUCTION

Community-acquired pneumonia (CAP) refers to pneumonia can occur outside healthcare institutions and impacts anybody like small children and the elderly individual as more probable to widen the disease.[1] Enormous patient datasets, with demographics, medical history and test outcome can be analyzed by ML techniques to identify similarity and predict results.[2] CAP symptoms like cough, fever, chest discomfort, breathing troubles and exhaustion, which can differ based on the severity of the disease.[3] Image tests like CT scans or chest X-rays can be utilized to evaluate the repetition and validate the diagnosis.[4] Repetitions in CAP particularly involve antibiotics, oxygen treatment and fluid executive as supportive care procedures.[5] Employing features related to the repetition of CAP, supervised ML techniques like the logistic regression (LR) or DT can be utilized for developing the forecast system.[6] An Unsupervised ML technique such as grouping and the principal component analyses can be employed for the prediction of subgroups in patients by altering the factors of risk and patterns of recurrence.[7]

The objective is to utilize ML techniques to develop the device for physicians for the classification of patients in risk of uncertain or persistent conditions, by lowering the CAP trouble and improving patient outcomes.[8] The recurrence of CAP is a frequent occurrence in patients with threat factors including fundamental lung infection, suppressive of the immune system usage and inventive period.[9] Protecting CAP recurrence includes predicting and depicting the threat factors like smoke termination and injection against infection and pneumococcal microbes.[10] The numerous datasets of patient's data is necessary to boost the ML technique for forecasting the recurrence of the CAP.[11] The effectiveness in the ML model's efficiency can be estimated through evaluated the illness in an independent data.[12] The objective of developing the ML model in predicting the CAP recurrence is useful for clinicians in forecasting the patients at managing the treatment and an excessive threat.[13] The possibility of CAP recurrence is influenced through infection amount, patient's age, fitness level and fundamental lung conditions or co-morbidities.[14] Clinicians evaluate the patient's medical symptoms, radiographic results and laboratory experiment outcomes to establish the possibility of CAP recurrence.[15] Preventive procedures, such as pulmonary treatment and supportive treatment can decrease the threat of CAP repetition in patients having fundamental lung infection.[16] The model can be influenced by demographics, medical record, laboratory test outcomes and image results.[17] Elderly individuals have extreme CAP and other risk factors having extreme risk of increasing inflammation-mediated severe cardiac actions that potentially incorporate antibiotic treatment efficiency. Figure 1 illustrates the bacterial CAP's pathophysiology process.

The study [18] proposes a new method for autonomously identifying pathogenic microorganisms of CAP using patient biological data from time-dependent body temperature, and conventional laboratory parameters. Treatment concerns include setting selection, initiating antiretroviral therapy, providing respiration assistance, transitioning to oral medicine, and short treatment duration.[19] The research presents a ML method for predicting future pneumonia risk using clinical data for the risk group.[20] The research aims to evaluate the efficacy of multiple ML techniques in predicting patient treatment outcomes.[21] The research aimed to assess the effectiveness of using ML techniques to predict 90-day passing rates in intensive care units.[22] During assessment, augmentation, and restoration, CAP in children is checked for both left and right lung functions.[23] CAP recurrence is most likely in the first few weeks, causing severe symptoms and longer hospital stays, and can occur from the same or different microorganism.[24] The objective to develop a tool for doctors to identify patients at high risk of recurrence, adjusts their treatment accordingly, and enhances patient outcomes.[25] This study explores the use of ML to predict CAP recurrence probability, aiding clinicians in identifying high-risk patients and developing personalized treatment methods.

METHOD

The Medical records from 200 patients were examined to compile research data. The computerized physician order entry (CPOE) system was introduced at the patient hospital and most records are still with document. The study suggests that electronic medical record systems may not fully capture all pneumonia factors, and recommends training competent term-workers for data extraction devices. The hospital's medical records were analyzed by patient list, assigning each patient based on pneumonia status during treatment. The study examined infected inpatient records at specialized mental hospitals, focusing on clozapine usage, drug-

drug interactions, and sneezing habits to determine the assumed clozapine, fluoxetine, carpine, or depatec combination. Fluxifoxamine, carpine, and depatec combined with clozapine can cause medication interactions and pneumonia development with continuous parameters tracking clozapine use and duration. The predictors were derived from pneumonia literature, expert doctor opinions, and patient records, resulting in 185 suitable instances with complete data.

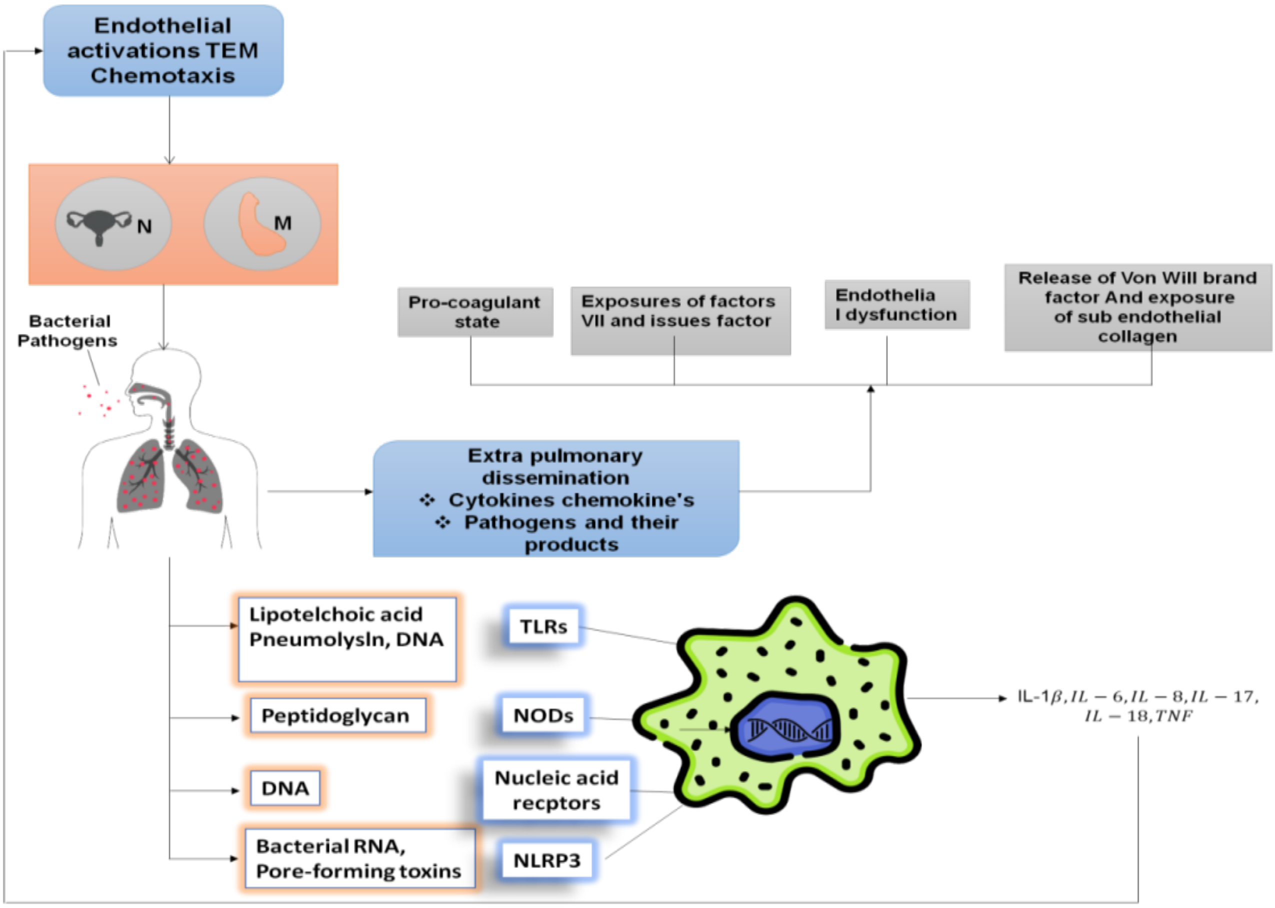

Figure 1. Pathophysiology of bacterial CAP

Test Setup

To forecast risk factors for individuals with schizophrenia, we utilized democracy and an open-source statistical software R 3.5.1. The research on pneumonia has the variety of ML techniques, including SVM, KNN, NB, RF, DT and C5.0. Therefore, we chose seven algorithms: LGR, CART, DT, NB, RF, SVM and KNN to build the forecasting the model and compare the effectiveness to the ML techniques with the outcomes of this study. The substring 6, 0-80 was utilized for automatically optimize the perfect representation parameter combinations is shown in table 1 to improve the performance of seven ML algorithms.

Table 1. Settings for model parameters

Method	Parameter	Best parameter
-	c	581,883
SVM	Sigma	0,1 786 673
CART	Cp	0,005 952 381
NB	fL	0
KNN	K	27
-	Use kernel	FASLE
RF	mtry	12
C5.0	Model	rules
-	Winnow	FALSE

The disparity in classification of samples can significantly impact ML performance. To ensure accurate measurement and prevent over-fitting, we used the holdout technique, dividing 70 % of training and 30 % of testing. The training dataset was repeatedly applied to seven classifiers using the ten-fold technique for evaluating model performance.

RESULTS

Table 2 presents predictive and descriptive data on schizophrenia, 106 patients has pneumonia and 79 patients has no pneumonia.

Table 2. Data Description of patients having pneumonia in patients and has no pneumonia in patients

Variables	Patients has no pneumonia (N=79)		Pneumonia Patients (N=106)	
-	Quantity	Description of data	Quantity	Description of data
Sex	Male or Female	Male or Female	Male or Female	Male: 53 and Female: 26
Clozapineuse	Yes/no	Yes/no	Yes/no	M=51, 28, SD=14, 66
Age	22~81	26~82	26~82	No=74, Yes=5
Link of Drug-drug	Yes/No	Yes/No	Yes/No	Yes=1 and No=78
Dosage	0~350	0~800	0~800	Yes=16, No=63
Coughing	Yes/No	Yes/No	Yes/No	M=10, 13, SD=47, 62
Change of leukocyte count	-8840~5500	-5200~5021	-5200~5021	M=8, 61, SD=42, 77
Duration of medication	0~295	0~377	0~377	SD=1996, 90, M=658, 63
Change of neutrophil count	−18, 6~32, 6	−29, 4~27, 3	−29, 4~27,3	M=0, 29, SD=21, 52 M=1, 05, SD=4, 53
Change of blood sugar level	-	-318~153	-318~153	Male: 53, Female: 26
Change of weight	-49~133 −11, 5~13	−16, 5~12, 5	−16, 5~12,5	M=51, 28, SD=14, 66

Performance assessment

The performance was evaluated using specificity, statistics with accuracy, accuracy and sensitivity, assessing the percentage of correctly categorized records, specificity indicating the recognition of negative reports. AUC evaluates binary classifier performance by graphing sensitivity against specificity, with 0,8 and 0,9 values indicating good and exceptional performance, modifying accuracy by considering random prediction possibility.

Model executions

Table 3 presents the performance of seven ML methods, evaluating accuracy, sensitivity, and specificity using crossover validation and obtaining averages and standard deviations for the training dataset.

Table 3. Performance assessment of the used models

Methods	Accuracy (%)		Sensitivity (%)		Specificity (%)	
	Test	Train	Test	Train	Test	Train
KNN	66	64	74	62	56	65
SVM	89	87	96	83	80	92
CART	83	84	90	73	73	85
LGR	73	67	79	62	66	70

The accuracy rate for all seven algorithms is greatest for, followed by support vector machine (SVM) (0,87), and CART (0,80). Less than 0, 7 percent accuracy was achieved by the remaining classifiers. The AUC values of random forest (RF), C5.0, and SVM are better than 0,9, suggesting good classifier performance. Classification and regression tree (CART) outperforms logistic regression (LGR) with AUC values over 0, 8, while k-nearest neighbors (KNN) performs poorly. SVM and CART outperform each other based on AUC. With the use of the testing dataset, it estimated our models to avoid over-fitting. By calculating, sensitivity and specificity, figure 2 and 3 demonstrates performances of various methods. The testing dataset shows higher accuracy for the remaining seven classifiers than the training dataset, indicating no over-fitting.

Predictor importance

The study used information gain and gain ratio criteria to evaluate predictor importance, revealing six factors influencing pneumonia frequency in schizophrenia patients: dosage, clozapine use, therapy duration, neutrophil and leukocyte count changes, and drug-drug interactions. The investigation factors are ranked based on their growth ratios is shown in table 4.

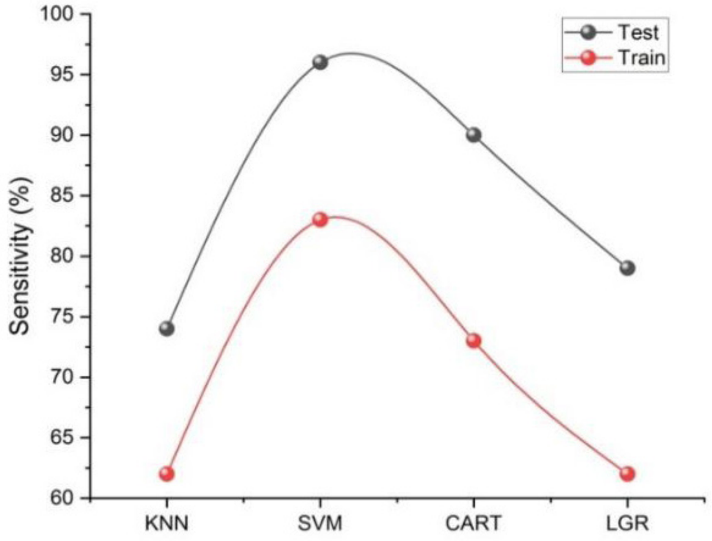

Figure 2. Comparison of sensitivity

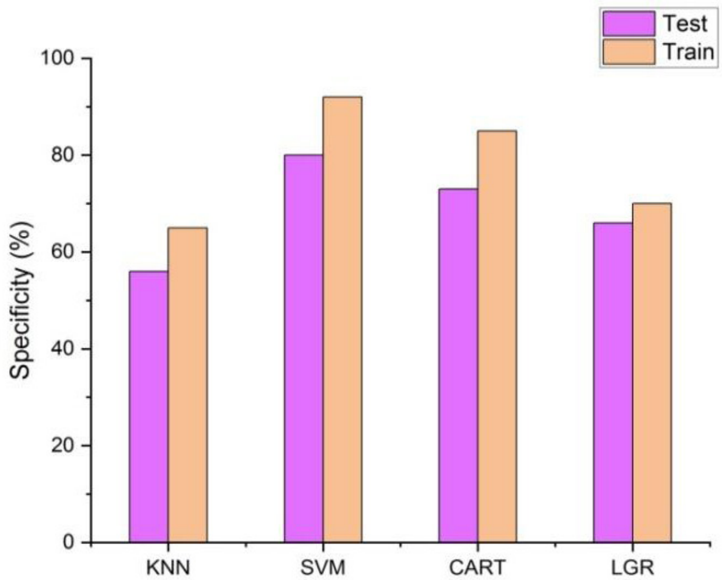

Figure 3. Comparison of specificity

Table 4. Ranking of the variables based on grow ratios

Variables	Grow ratio	Ranking	Information grow	Ranking
Clozapineuse	0,150	2	0,093	2
Dosage	0,217	1	0,124	1
Neutrophil Change Count	0,068	6	0,046	4
Medication Duration	0,115	3	0,072	3
Drug-drug Interaction	0,076	5	0,028	6
Leukocyte Change Count	0,107	4	0,039	5

Accepting the pneumonia risk factors in individuals have schizophrenia is crucial, but only a small percentage of studies use ML for prediction. The study's seven ML techniques, including SVM and CART, showed superior prediction accuracy, as shown in figure 4.

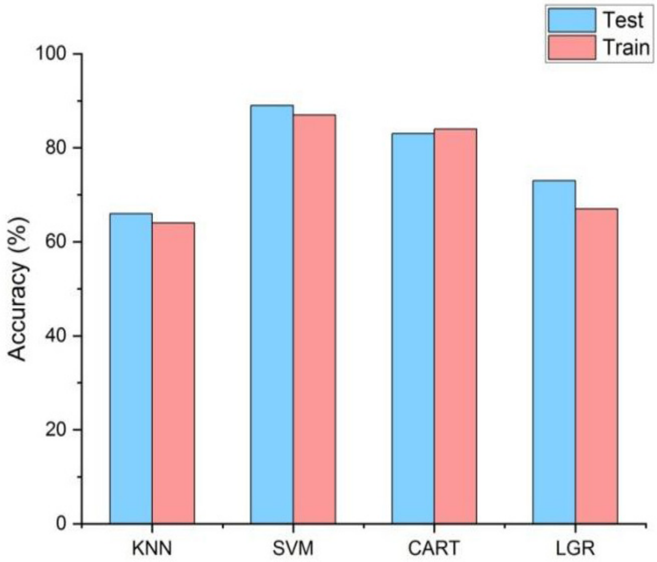

Figure 4. Comparison of accuracy

The study highlights the increased pneumonia risk associated with common antipsychotics, highlighting the need for careful medical monitoring and prescribing. Used a boosting naive Byes student to more precisely forecast the amount of time pneumonia patients would need to spend in the hospital and 95, 2 % of accuracy is the model. With an accuracy of 67, 5 %, their model surpassed our non-boosting naive Bayes model. Clinical decision system for pneumonia hospitalization was developed using an SVM learner, demonstrating the effectiveness of boosting ensemble strategies in enhancing model predictive power. The six indicators in the model with the best predictive accuracy (83,9 %) were ages, sex, amount of medicines, span of stay, co-morbidities amount and cost of total admission. Predictive accuracy was somewhat higher (87,1 %) with our suggested model using a SVM learner. ML is being used to identify risk factors for pneumonia in schizophrenia patients, revealing potential indicators that standard statistical models often overlook. ML helps identify pneumonia risk factors in schizophrenia patients, reducing class imbalance issues in health data. Psychiatrists should prioritize risk variables in diagnosis and treatment. Future research should consider the limitations of our inquiry and gather more data from a wider range of hospitals.

CONCLUSION

The effectiveness of ML approaches in predicting recurrence in CAP is not definitively determined based on available information. Some studies have utilized ML algorithms to forecast the CAP recurrence risk by the factors like age, co-morbidities, and laboratory test results. Algorithms have the potential to offer more precise and personalized risk assessments than conventional statistical methods. The study developed the classification framework by seven ML techniques to detect pneumonia possibility factors in schizophrenia patients. The ML approaches used were validated using 185 suitable examples, with RF and DT shows the improved prediction efficiency. An effectiveness of ML algorithms for CAP recurrence prediction relies on the quality and quantity of data used for training. The ML approaches show promise in predicting recurrence in CAP, but further research is needed to confirm their effectiveness and establish their clinical utility.

REFERENCES

1. Graham, Frances F., et al. “Global perspective of Legionella infection in community-acquired pneumonia: a systematic review and meta-analysis of observational studies.” International journal of Environmental research and public health 19.3 (2022): 1907.

2. Abdel-Basset, Mohamed, et al. “Two-stage deep learning framework for discrimination between COVID-19 and community-acquired pneumonia from chest CT scans.” Pattern recognition letters 152 (2021): 311-319.

3. Carpenter, Christopher R., et al. “Diagnosing COVID-19 in the emergency department: a scoping review of clinical examinations, laboratory tests, imaging accuracy, and biases.” Academic Emergency Medicine 27.8 (2020): 653-670.

4. Afonso, P. Diana, Amanda Isaac, and José Martel Villagrán. “Chondroid tumors as incidental findings and differential diagnosis between enchondromas and low-grade chondrosarcomas.” Seminars in musculoskeletal

radiology. Thieme Medical Publishers, 23.01 (2019): 003-018.

5. Swenson, Kai E., and Dean L. Winslow. “Impact of sepsis mandates on sepsis care: unintended consequences.” The Journal of Infectious Diseases 222.Supplement_2 (2020): S166-S173.

6. Shamrat, FM Javed Mehedi, et al. “Implementation of machine learning algorithms to detect the prognosis rate of kidney disease.” 2020 IEEE international conference for innovation in technology (INOCON). IEEE, (2020): 1-7.

7. Islam, Md Khairul, et al. “Brain tumor detection in MR image using superpixels, principal component analysis and template based K-means clustering algorithm.” Machine Learning with Applications 5 (2021): 100044.

8. Chekroud, Adam M., et al. “The promise of machine learning in predicting treatment outcomes in psychiatry.” World Psychiatry 20.2 (2021): 154-170.

9. Klein, Allan L., et al. “Efficacy and safety of rilonacept for recurrent pericarditis: results from a phase II clinical trial.” Heart 107.6 (2021): 488-496.

10. Santoro, Stephanie L., et al. “Pneumonia and respiratory infections in Down syndrome: A scoping review of the literature.” American Journal of Medical Genetics Part A 185.1 (2021): 286-299.

11. Yang, Jialiang, et al. “Prediction of HER2-positive breast cancer recurrence and metastasis risk from histopathological images and clinical information via multimodal deep learning.” Computational and structural biotechnology journal 20 (2022): 333-342.

12. Elavarasan, Dhivya, and PM Durairaj Vincent. “Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications.” IEEE access 8 (2020): 86886-86901.

13. Sutton, Elizabeth J., et al. “A machine learning model that classifies breast cancer pathologic complete response on MRI post-neoadjuvant chemotherapy.” Breast Cancer Research 22 (2020): 1-11.

14. Valent, Peter, et al. “Diagnosis, classification and management of mast cell activation syndromes (MCAS) in the era of personalized medicine.” International journal of molecular sciences 21.23 (2020): 9030.

15. Poongodi, M., Mounir Hamdi, and Huihui Wang. “Image and audio caps: automated captioning of background sounds and images using deep learning.” Multimedia Systems 29.5 (2023): 2951-2959.

16. Chebib, Najla, et al. “Pneumonia prevention in the elderly patients: the other sides.” Aging clinical and experimental research 33 (2021): 1091-1100.

17. Corny, Jennifer, et al. “A machine learning–based clinical decision support system to identify prescriptions with a high risk of medication error.” Journal of the American Medical Informatics Association 27.11 (2020): 1688-1694.

18. Rothberg, Michael B. “Community-acquired pneumonia.” Annals of internal medicine 175.4 (2022): ITC49-ITC64.

19. Kuo, Kuang Ming, et al. “Predicting hospital-acquired pneumonia among schizophrenic patients: a machine learning approach.” BMC Medical Informatics and Decision Making 19 (2019): 1-8.

20. Goodwin, Travis R., and Dina Demner-Fushman. “Deep learning from incomplete data: detecting imminent risk of hospital-acquired pneumonia in ICU patients.” AMIA Annual Symposium Proceedings. Vol. 2019. American Medical Informatics Association, 2019 (2019): 467.

21. Giang, Christine, et al. “Predicting ventilator-associated pneumonia with machine learning.” Medicine 100.23 (2021): e26246.

22. Thorsen-Meyer, Hans-Christian, et al. “Dynamic and explainable machine learning prediction of mortality

in patients in the intensive care unit: a retrospective study of high-frequency data in electronic patient records." The Lancet Digital Health 2.4 (2020): e179-e191.

23. Huang, Dongmin, Lingwei Wang, and Wenjin Wang. "A multi-center clinical trial for wireless stethoscope-based diagnosis and prognosis of children community-acquired pneumonia." IEEE Transactions on Biomedical Engineering 70.7 (2023): 2215-2226.

24. Feldman, Charles, et al. "Pathogenesis and prevention of risk of cardiovascular events in patients with pneumococcal community-acquired pneumonia." Journal of internal medicine 285.6 (2019): 635-652.

25. Azabou, Eric, et al. "Vagus nerve stimulation: a potential adjunct therapy for COVID-19." Frontiers in medicine 8 (2021): 625836.

FINANCING
None.

CONFLICT OF INTEREST
None.

AUTHORSHIP CONTRIBUTION
Conceptualization: Dnyaneshwar Prabhakar Bawane, Raja Ramalingam, M. Gopi, Vaibhav Kaushik, Prakhar Goyal, Yuvraj Parmar.

Data curation: Dnyaneshwar Prabhakar Bawane, Raja Ramalingam, M. Gopi, Vaibhav Kaushik, Prakhar Goyal, Yuvraj Parmar.

Formal analysis: Dnyaneshwar Prabhakar Bawane, Raja Ramalingam, M. Gopi, Vaibhav Kaushik, Prakhar Goyal, Yuvraj Parmar.

Drafting - original draft: Dnyaneshwar Prabhakar Bawane, Raja Ramalingam, M. Gopi, Vaibhav Kaushik, Prakhar Goyal, Yuvraj Parmar.

Writing - proofreading and editing: Dnyaneshwar Prabhakar Bawane, Raja Ramalingam, M. Gopi, Vaibhav Kaushik, Prakhar Goyal, Yuvraj Parmar.