ORIGINAL RESEARCH PAPER

Volume - 15 | Issue - 04 | April - 2026 | PRINT ISSN No. 2277 - 8179 | DOI : 10.36106/ijsr

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH

THE AWARENESS ON TELEMEDICINE AMONG PATIENTS IN EASTERN PART OF INDIA: AN APPROACH THROUGH DATA MINING

Health Science

Kaushick Mishra* Department of Pharmacology, Malda Medical College and Hospital, Uma Roy Sarani, Singatala, Malda – 732101, West Bengal, India*Corresponding Author

Subhabrata Ghosh Department of Physiology, Institute of Post Graduate Medical Education and Research (I.P. G. M. E & R), 244, A.J.C. Bose Road, Kolkata – 700020, West Bengal, India

Mayukh Goswami Department of Opthalmology, Calcutta National Medical College and Hospital, 32 Gorachand Road, Beniapukur, Kolkata – 700014, West Bengal, India

ABSTRACT

Objective: The present study aims to predict predictive accuracy of dataset through machine learning (ML) algorithms to know the awareness among patients towards acceptance of telemedicine during the post-pandemic scenario. **Methods:** A cross-sectional hospital-based questionnaire survey was conducted randomly among 150 patients who received medicines through phone calls or internet calls in a tertiary care hospital and were treated as outpatients. Among 150 parents, there are two categories in which one-time users (group A, n = 75) and many times users (group B, n = 75) receive telemedicine. The knowledge about telemedicine was assessed by prefixed questionnaires. The datasets as per responses on telemedicine were studied through ML algorithms especially tree algorithms and determined the sensitivity, specificity, and accuracy on awareness performance.**Results:** The present results indicated that the correctly classified models, receiver operating characteristic (ROC) curve, Matthew's correlation coefficient (MCC) curve and Precision-recall curve (PRC) observed highest values in RF followed by J48 and LMT as per 10-folds cross validation test. Regarding sensitivity, specificity and accuracy, higher value was obtained for J48 (80.25%, 73.91% and 77.33%) followed by RF (77.01%, 74.60% and 76.00%), DST (72.38%, 84.44% and 76.00%), LMT (72.55%, 81.25% and 75.33%), REPT (73.40%, 75.00% and 74.00%) and lower value was obtained for RT (72.53%, 71.19% and 74.00%).**Conclusion:** The prediction accuracy on awareness regarding telemedicine was observed moderate as per data mining based on ML Tree algorithms. The present study may be informative related to these algorithm's prediction accuracy in the dataset.

KEYWORDS

Data mining, Machine learning, Technology acceptance, Telemedicine awareness.

INTRODUCTION

From novel coronavirus disease or nCOVID-19 pandemic, still there is a practice of telemedicine in different tertiary care hospitals.[1,2] World Health Organization (WHO) defined the telemedicine as the provision of healthcare services via all healthcare providers using information and communication technologies to exchange accurate data for illness and injury diagnosis, treatment, and prevention, research and evaluation, as well as for the ongoing education of healthcare providers, with a focus on the benefits of advancing the health of individuals and their communities.[3]

As a result of the COVID-19 pandemic making remote service delivery necessary, studies on the location of telemedicine in different geographic areas followed in 2020.[4] One of the most important aspects of the public health response was determined to be the creation of a more effective and integrated system for global public health readiness, facilitating cooperation and resource sharing.[5]

Thus, telemedicine might be useful for improving the delivery of health-care services using information and communication technologies for the exchange of valid information for diagnosis, treatment and prevention of disease and injuries, etc.[1] Moreover, it might also help in supporting training before referral to the hospitals when needed and reducing unnecessary hospital visits.[6] In modern telemedicine, additional technological breakthroughs that enable the transfer of images make it easier to share medical information like X-rays, scans, and real-time audio and video consultations.[7]

Although there had been few some studies which tried to determine doctor's knowledge and expertise about telemedicine consultations,[8,9] some studies have been conducted in India and abroad to establish the awareness among patients about telemedicine[10,11,12,13] where the patients receive as a tool of medical advice during post-pandemic period. So, this study is an incredibly unique study conducted to determine awareness of telemedicine for patient's perspective in a developing country, which is lacking.

In recent decades, big data mining by using machine learning (ML) algorithm models is demanding research in which performance accuracy can easily be recognized.[14-17] Many studies on big data mining by using ML algorithms related to biomedical and biological science, have been well recognized,[16-22] and few studies conducted the prediction of performance accuracy of telemedicine dataset[23,24] but on the awareness on telemedicine among patients' is a first-time study in eastern part of India in which easy predictive ML methods especially Tree algorithms were studied from the survey dataset.

Material and methods

Study design

A cross-sectional hospital-based survey was conducted from January 2025 to May 2025 on 150 patients received telemedicine of a tertiary care hospital post-pandemic phase. The survey identified their awareness towards acceptance of telemedicine. who received medicines through phone calls or internet calls in a tertiary care hospital who were treated as outpatients. Among 150 parents, there are two categories in which one-time users (group A, n = 75) and many times users (group B, n = 75) receive telemedicine. The awareness regarding telemedicine was assessed by prefixed questionnaires. The datasets as per responses on telemedicine were studied through ML algorithms especially Tree algorithms and determined the sensitivity, specificity, and accuracy on awareness performance.

Each group had been given a different set of questionnaires prepared in local languages and English, which determined their demographic features and socioeconomic status as well as their awareness towards telemedicine consultation.

Tool and Dataset selection

The data mining through ML modelling algorithm was executed by using WEKA (Waikato Environment for Knowledge Analysis) tool (version, 3.8.5) developed by Frank et al.[25] In the WEKA tool, we explored the data for pre-processing followed by classifying 4 inbuilt ML Tree algorithms to obtain different statistical results. The various attributes such as patient types (Group A and Group B), age, gender (males and females), educational level (P, HS, G, I, Pr), residence (urban or rural), Mother tongue (Bengali, Others), socioeconomic status (U, UM, UL, L, LM), telemedicine term (Yes or No), medicine received through (phone calls or internet calls), hospital_vist_difficulties (Yes or No), telemedicine_comfortable (Yes or No), beneficial_towards_chamber_consultation (Yes or No), can_save_hospital_visit_cost (Yes or No),can_save_hospital_visit _time (Yes or No), and class (good and bad) in the dataset were selected.

Data mining approach through machine learning Tree algorithms

The predictive accuracy of data for two categories (Group A and Group B) of patients were predicted through ML modelling especially Tree algorithms as per different classifiers viz. Pruned and unpruned decision tree C4 (J48), Random Forest (RF), Random tree (RT), and Class implementing minimal cost-complexity pruning (REPT) were predicted from dataset to know the overall awareness performance accuracy.

Prediction of awareness accuracy

The performance of model accuracy of above-mentioned ML Tree algorithm classifications as per correctly and incorrectly classified instances, Kappa (K) statistics, mean absolute error (MAE) and root mean squared error (RMSE) were studied for 10 folds cross validation.[26] As per Bouckaert et al.,[26] the modelling summary of results such as precision value, recall value, F-measure value, Matthew's correlation coefficient (MCC), receiver operating characteristic (ROC) curve and precision-recall curve (PRC), respectively as per class of good and bad response were retrieved from the studied tool.

RESULTS

The present study evaluated awareness on telemedicine among 150 patients (Group-A and Group-B) 75 numbers in each group through data mining approach. Fig 1 exhibits the graphical representation of above-mentioned fourteen attributes of patients' awareness on telemedicine. The value of sum of weight is 150 for all nominal and numeric attributes.

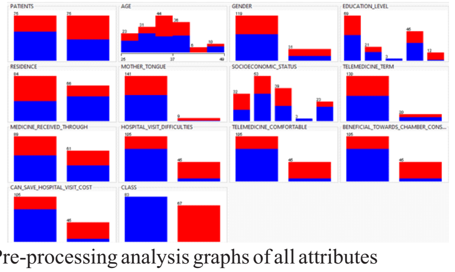

Fig. 1. Pre-processing analysis graphs of all attributes

In the case of age of patients, the mean ± SD value was obtained 35.31 ± 5.65 with minimum and maximum values of 25-49 years in which the higher 44 instances with a range of 33.00-37.00 years, followed by 36 instances of 37.00-41.00 years, 31 instance of 29.00-33.00 years, 23 instances of 25-29 years and 10 instances 45-49 years while lower value (41-45 years) of 6 instances, respectively were recorded as numeric data. In the case of gender, 119 males and 31 females, education level (Graduate 69, higher secondary 45, primary 21, professional 12 and illiterate 3 nos.), residence (urban 84 and rural 66 nos.), mother tongue (Bengali 141 and Hindi 9 nos.), socioeconomic status (upper middle 53, upper lower 39, upper 32, lower middle 23 and lower 3 nos.) as nominal data recorded. Regarding the telemedicine term, maximum 130 cases known while 20 nos. did not hear. Maximum cases of about 89 nos. were used internet calling while 61 nos. were used phone calling. For hospital visit difficulties, maximum numbers (105) reported difficulties while minimum numbers (45) did not tell difficulties. Maximum cases (105) reported telemedicine comfortable while minimum cases (45) reported non-comfortable. Maximum cases (105) reported beneficial towards chamber consultation while minimum cases (54) reported none. Maximum cases (105) reported to save hospital visit costs while minimum cases (45) reported none. As per the responses of the patients, the data was classified as good (83) and bad (67), respectively.

Table 1 describes the summary results of studied 6 ML algorithms especially Tree models such as Decision Stum Tree (DST), Pruned and Unpruned Decision Tree C4 (J48), Logistic Model Tree (LMT), Random Forest (RF), Random Tree (RT), and Class implementing minimal cost-complexity pruning (REPT) related to 14 attributes to predict the overall awareness prediction accuracy. Overall prediction of algorithms classification, the correctly classified models were observed the highest values in J48 (77.33), followed by DST and RF (76.00), LMT (75.33), REPT (74.00), while the lowest value in RT (72.00), respectively as per 10-folds cross validation test (Table 1). The statistical values for Kappa statistic (KS), mean absolute error (MAE) and root mean squared error (RMSE) of studied models, better predicted J48 followed by followed by DST and RF, LMT, REPT and RT (Table 1).

Table 1: Results on different classified instances and statistical values for different Tree algorithms

Sl No.	Classifier algorithms	CCI (%)	ICI (%)	KS	MAE	RMSE
1.	DST	76.00	24.00	0.50	0.36	0.43
2.	J48	77.33	22.67	0.54	0.26	0.39
3.	LMT	75.33	24.67	0.49	0.32	0.41
4.	RF	76.00	24.00	0.51	0.27	0.41
5.	RT	72.00	28.00	0.43	0.26	0.46
6.	REPT	74.00	26.00	0.46	0.31	0.42

CCI = Correctly classified instances; ICI = Incorrectly classified instances; DST = Decision Stum Tree; J48 = Pruned and Unpruned Decision Tree C4; LMT = Logistic Model Tree; RF = Random Forest; RT = Random Tree; REPT = Class Implementing Minimal Cost-Complexity Pruning; KS = Kappa Statistics; MAE = Mean Absolute Error; RMSE = Root Mean Squared Error

In the case of predictive statistical analysis data for 6 ML Tree algorithms based on good and bad class, the F-score values higher for DST (81.0% and 68.0%), followed by LMT (80.0% and 68.0%), J48 (79.3% and 75.0%), RF (78.0% and 72.3%), REPT (78.0% and 68.3%) and lower for RT (76.0% and 67.0%), respectively predicted. For MMC based on good and bad class, higher value for J48 (54.3%) followed by DST (52.4%), RF (51.2%), RF (51.2%), LMT (50.5%), REPT (47.1%) and lower for RT (43.0%), respectively predicted. For ROC based on good and bad class, higher value for RF (85.0%) followed by J48 (83.1%), LMT (82.5%), REPT (79.8%), RT (79.0%) and lower for DST (68.9%), respectively predicted. For ROC (area under curve) based good and bad class, higher value for RF (90.0% and 79.0%) followed by J48 (87.9% and 70.1%), LMT (86.0% and 75.0%), REPT (82.6% and 70.5%), RT (78.3% and 69.3%) and lower for DST (67.0% and 64.0%), respectively predicted. The ROC curves are exhibited for each algorithm viz. RF (Fig 2), J48 (Fig 3), LMT (Fig 4), REPT (Fig 5), RT (Fig 6) and DST (Fig 7).

Table 2: Summary of statistical values for prediction accuracy of studied algorithms

Sl. No.	Classifier algorithms	Class	F-Measure	MCC	ROC area	PRC area
1.	DST	Good	0.809	0.524	0.689	0.668
		Bad	0.679	0.524	0.689	0.637
2.	J48	Good	0.793	0.543	0.831	0.879
		Bad	0.750	0.543	0.831	0.701
3.	LMT	Good	0.800	0.505	0.825	0.857
		Bad	0.678	0.505	0.825	0.750
4.	RF	Good	0.788	0.512	0.850	0.900
		Bad	0.723	0.512	0.850	0.789
5.	RT	Good	0.759	0.430	0.790	0.783
		Bad	0.667	0.430	0.790	0.693
6.	REPT	Good	0.780	0.471	0.798	0.826
		Bad	0.683	0.471	0.798	0.705

DST = Decision Stum Tree; J48 = Pruned and Unpruned Decision Tree C4; LMT = Logistic Model Tree; RF = Random Forest; RT = Random Tree; REPT = Class Implementing Minimal Cost-Complexity Pruning; MCC = Matthew's correlation coefficient; ROC = Receiver operating characteristic; PRC = Precision-recall curve

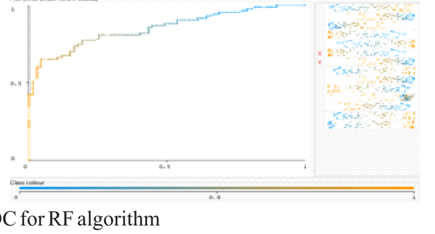

Fig 2: ROC for RF algorithm

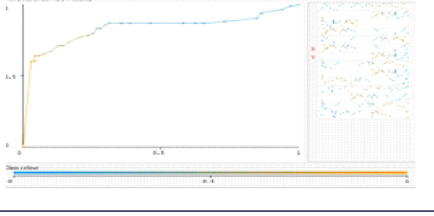

Fig 3: ROC for J48 algorithm

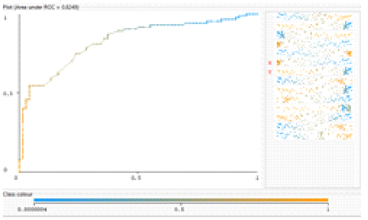

Fig 4: ROC for LMT algorithm

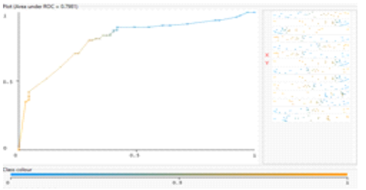

Fig 5: ROC for REPT algorithm

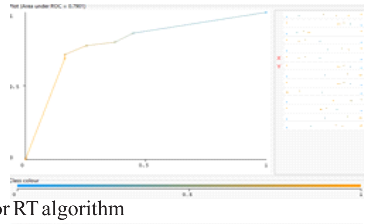

Fig 6: ROC for RT algorithm

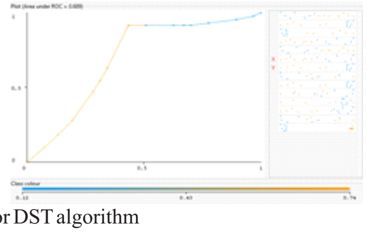

Fig 7: ROC for DST algorithm

Table 3 describes the sensitivity, specificity, and predictive accuracy for the suitable Tree algorithms in which the higher value was obtained for J48 (80.25%, 73.91% and 77.33%) followed by RF (77.01%, 74.60% and 76.00%), DST (72.38%, 84.44% and 76.00%), LMT (72.55%, 81.25% and 75.33%), REPT (73.40%, 75.00% and 74.00%) and lower value was obtained for RT (72.53%, 71.19% and 74.00%), respectively.

Table 3: Overall sensitivity, specificity, and predictive accuracy values of best algorithms

Sl. No.	Classifier algorithms	Sensitivity (%)	Specificity (%)	Accuracy (%)
1.	DST	72.38	84.44	76.00
2.	J48	80.25	73.91	77.33
3.	LMT	72.55	81.25	75.33
4.	RF	77.01	74.60	76.00
5.	RT	72.53	71.19	72.00
6.	REPT	73.40	75.00	74.00

DISCUSSION

It is well-known fact that telemedicine consultation service is primarily based on patients' awareness and doctors' knowledge and proficiency.[7,8,12,27-29] Furthermore, it is securely required to know the patients' awareness followed by satisfaction regarding recovery of disease.[6,7,10,13,27-30]

In recent periods, many researchers are emphasizing big data mining by using ML algorithms in which the prediction accuracy can easily be accomplished in medical sciences[15,16,23,24] but the performance accuracy of telemedicine dataset on the awareness of patients' is the first-time approach in eastern part of India in which easy predictive accuracy is achieved by using ML Tree algorithms from the survey dataset.

Although the present study was performed with the small sets of data as per first-time users compared to many-times users of telemedicine, in which present dataset predicted moderate awareness based on Tree algorithms as per the sensitivity, specificity and accuracy of these algorithms viz. higher value was obtained for J48 (80.25%, 73.91% and 77.33%) followed by RF (77.01%, 74.60% and 76.00%), DST (72.38%, 84.44% and 76.00%), LMT (72.55%, 81.25% and 75.33%), REPT (73.40%, 75.00% and 74.00%) and lower value was obtained for RT (72.53%, 71.19% and 74.00%) by using WEKA tool. The ML algorithms for classification are usually estimated through common methods, which are similar to previous studies in the field of medical sciences.[15,16,23,24]

Interestingly, WEKA tool supports several types of datasets for establishing big data mining through ML algorithms especially on disease diagnosis and prevention.[15,16,23,24] The present algorithms can be suitable to determine easily the prediction accuracy on moderate awareness among patients which may be due to the many-times user till date through the higher education, better socio-economic status, ability to perform internet or virtual calls compared to first-time users.

CONCLUSION

In the present predictive findings, it was known that awareness prediction accuracy is moderate performance regarding telemedicine between the groups of Group-A and Group-B during the post-pandemic of n-COVID-19 as per ML Tree algorithms by using WEKA tool. Out of 6 algorithms, best performance was obtained in three models viz. J48, DST and RF. The present predictive study could be informative based on these Tree algorithms, which recognized moderate accuracy. It is suggested with more sample size, which may be observed better performance with these and other algorithms to confirm better awareness among patients on telemedicine.

Acknowledgements

Authors convey thanks to all patients who participated and helping during survey. Also, thanks to the developers of WEKA tool used in the present study.

Declaration of conflicting interests

All the authors declared no potential conflicts of interest.

Funding

This is non-funded project.

REFERENCES

1. Agarwal N, Jain P, Pathak R, Gupta R. Telemedicine in India: A tool for transforming health care in the era of COVID-19 pandemic. J Educ Health Promot. 2020;9:190.
2. Losorelli SD, Vendra V, Hildrew DM, Woodson EA, Brenner MJ, Sirjani DB. The future of telemedicine: revolutionizing health care or flash in the pan?. Otolaryngology–Head and Neck Surgery. 2021;165(2):239-43.
3. World Health Organization, editor. Telemedicine: Opportunities and developments in member states: Report on the second Global survey on eHealth [Internet]. Published 2010. Geneva, Switzerland: World Health Organization. 93 p. (Global Observatory for eHealth Series). http://www.who.int/goe/publications/goe_telemedicine_2010.pdf (Accessed on 20 July 2026).
4. Ndwabe H, Basu A, Mohammed J. Post pandemic analysis on comprehensive utilization of telehealth and telemedicine. Clinical eHealth. 2024;7:5-14.
5. Bhaskar S, Bradley S, Chattu VK, Adisesh A, Nurtazina A, Kyrykbayeva S, et al. Telemedicine across the globe-position paper from the COVID-19 pandemic health system resilience PROGRAM (REPROGRAM) international consortium (Part 1). Frontiers in Public Health. 2020;8:556720.
6. Searle B, Barker RO, Stow D, Spiers GF, Pearson F, Hanratty B. Which interventions are effective at decreasing or increasing emergency department attendances or hospital admissions from long-term care facilities? A systematic review. BMJ Open. 2023;13(2):e064914.
6. Blozik E, Wildeisen I.E, Fueglistaler P, Overbeck J.V. Telemedicine can help to ensure that patients receive timely medical care. Journal of Telemedicine and Telecare. 2012; 18(2):119-21.
7. Meher SK , Tyagi R.S, Chaudhry T. Awareness and attitudes to telemedicine among doctors and patients in India. Journal of Telemedicine and Telecare. 2009;15(3):139-41.
8. Bali S, Gupta A, Khan A, Pakhare A, Evaluation of Telemedicine centres in Mahya Pradesh, Central India. Journal of Telemedicine and Telecare. 2016:22(3):183-1881B0B.
9. Bagayoko CO, Müller H, Geissbuhler A. Assessment of internet-based tele-medicine in Africa (the RAFT project). Computerized Medical Imaging and Graphics. 2006:30(6):407-416.
10. Meher S.K , Tyagi R.S, Chaudhry T. Awareness and attitudes to telemedicine among doctors and patients in India. Journal of Telemedicine and Telecare. 2009;15(3):139-41.
11. Lee PI, Hu YL, Huang YC, Hsueh PR. Are children less susceptible to COVID-19? J Microbiol Immunol Infect. 2020:53(3):371-2.
12. Chauhan V, Galwankar S, Arquilla B, Garg M, Somma SD, El-Menyar A, Krishnan V, Gerber J, Holland R, Stawicki SP. Novel coronavirus (COVID-19): Leveraging telemedicine to optimize care while minimizing exposures and viral transmission. J Emerg Trauma Shock. 2020;13:20-4.
13. Parameshwarappa PM, Jose OJ. Telemedicine awareness and the preferred digital healthcare tools: A community-based cross-sectional study from rural Karnataka, India. Indian Journal of Community Medicine. 2023;8(6):915-9.
14. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-444.
15. Mishra S, Mallick PK, Tripathy HK, Bhoi AK, González Briones A. Performance evaluation of a proposed machine learning model for chronic disease datasets using an integrated attribute evaluator and an improved decision tree classifier. Applied Sciences. 2020; 10(22):8137.
16. Mishra S, Dash A, Jena L. Use of deep learning for disease detection and diagnosis, In: Bhoi A, Mallick P, Liu CM, Balas V, (eds) Bio-inspired Neurocomputing. Studies in

Computational Intelligence. 903 Springer: Singapore, 2021.
17. Talapatra SN, Chaudhuri R, Ghosh S. CellProfiler and WEKA Tools: Image analysis for fish erythrocytes shape and machine learning model algorithm accuracy prediction of dataset, World Scientific News. 2021;154:101-116.
18. Bhuvaneswari E, Sarma Dhulipala VR. The study and analysis of classification algorithm for animal kingdom dataset. Information Engineering. 2013;2(1):6-13.
19. Merelli I, Pérez-Sánchez H, Gesing S, D'Agostino D. Managing, analysing, and integrating big data in medical bioinformatics: open problems and future perspectives. Biomed Res Int. 2014;2014:134023.
20. Chakraborty I, Choudhury A, Banerjee TS. Artificial intelligence in biological data. J Inform Tech Softw Eng. 2017;7: 1000207.
21. Almryad A, Kutucu H. Automatic identification for field butterflies by convolutional neural networks. Engineering Science and Technology, an International Journal. 2020;23:189-195.
22. Attwal KPS, Dhiman AS. Exploring data mining tool - Weka and using Weka to build and evaluate predictive models. Advances and Applications in Mathematical Sciences. 2020;19(6): 451-469.
23. Reategui-Rivera CM, Cui W, Escobar-Agreda S, Rojas-Mezarina L, Finkelstein J. Machine Learning-Based Prediction of No-Show Telemedicine Encounters. Telemed Rep. 2025;6(1):109-119.
24. Kokulavani K, Manthena KV, Sakthisaravanan B, Sreelatha T, Visumathi J, Guruprakash KS, et al. Optimizing telemedicine patient care with machine learning for disease progression and treatment planning. Eng Technol Appl Sci Res. 2025;15(4):25783-8.
25. Frank E, Hall MA, Witten IH. *The WEKA workbench, Online appendix for data mining: Practical machine learning tools and techniques*. 4th ed. Morgan Kaufmann, 2016.
26. Bouckaert RR, Frank E, Hall M, et al. WEKA manual for version 3-8-5. University of Waikato, Hamilton, New Zealand, 2020, December 21.
27. Bagayoko CO, Müller H, Geissbuhler A. Assessment of internet-based tele-medicine in Africa (the RAFT project). Computerized Medical Imaging and Graphics. 2006;30(6):407-416.
28. Sakurai K, Suzuki Y, Mori K, Imai T, Hasegawa Y. Awareness of telemedicine among patients with neurological disorders. Journal of the Neurological Sciences. 2017;381(Suppl.): 375.
29. Kumari Rani P, Raman R, Manikandan M, Mahajan S, Paul PG, Sharma T. Patient satisfaction with tele-ophthalmology versus ophthalmologist-based screening in diabetic retinopathy. J Telemed Telecare. 2006;12(3):159-60.
30. Lunova T, Hurndall KH, Kirk UB, Franklin BD, Darzi A, Neves AL. Patient safety measures for virtual consultations in primary care: a systematic review. BMJ Qual Saf. 2025;bmjqs-2025-018712.