ORIGINAL ARTICLE

A novel approach to predict the risk of invasive candidiasis using artificial neural networks and comparison with other models

Sheetal Gouda[1], R. Sathyajith[2], Basavaraj V. Peerapur[3]*

[1]Department of Microbiology, Karwar Institute of Medical Sciences, Karwar- 581301 (Karnataka) India, [2]Laboratory Director, Trinity Diagnostics, Palakkad -678013 (Kerala) India, [3]Department of Microbiology, Raichur Institute of Medical Sciences, Raichur-584102 (Karnataka) India

Abstract

Background: Invasive Candidiasis (IC) is the third common cause of neonatal sepsis associated with high mortality. Low Birth Weight (LBW), prematurity, diabetes, and multifocal colonization increase the risk of IC. Since the development of IC involves multidimensional risk factors, various models are available to predict IC, including colonization index, corrected colonization index, Ostrosky's clinical prediction rule, and candida score. *Aim and Objectives:* To develop a novel model using Artificial Neural Networks (ANN) that has a strong capability to handle complex multi-variate risk factors. *Material and Methods:* A prospective, hospital-based observational study was conducted at Raichur, India, among neonates of LBW (<1500 g). Swabs were collected from various body sites followed by blood culture from each neonate to assess colonization and IC, respectively. χ^2-tests of significance were applied to ascertain significant risk factors for IC. Various risk prediction models were compared with the ANN model. *Results:* The study population consisted of 103 neonates, of which 21 were diagnosed with IC. The most common isolate obtained was *C. albicans*, followed by *C. parapsilosis*. The factors significantly affecting IC were gestation age, mode of delivery, respiratory distress, diabetes, use of antibiotics, and multifocal colonization. ANN model predicted IC with a PPV of 83.3% and a Negative Predictive Value (NPV) of 98.7%. Various prediction models had poor values of PPV and sensitivity compared to the ANN model. *Conclusion:* ANN showed higher PPV, NPV, sensitivity, positive likelihood ratio with a very low value of negative likelihood ratio, suggesting that ANN is superior to the other models. With multi-centred data available from various geographic regions, the ANN approach can be further strengthened for usage in public and private health setups.

Keywords: Invasive Candidiasis, Candida, Artificial Neural Networks, Prediction Model, Neonates

Introduction

Invasive Candidiasis (IC) is a major health concern associated with excess attributable mortality of 10.1% in paediatric age groups [1]. Premature neonates having Very Low Birth Weight (< 1500 g) (VLBW) and Extremely Low Birth Weight (<1000 g) (ELBW) are at higher risk for developing IC with an incidence of 2.8% and 5.3%, respectively in India [2]. The use of broad-spectrum Antibiotics (AB), steroids, central venous catheter, diabetes, Total Parenteral Nutrition (TPN), prolonged mechanical ventilation, and indwelling catheters further increases the risk of developing IC [3]. Colonization of the skin or mucous membrane has also been suspected as a risk for developing IC as colonization precedes infection [3]. The progression from colonization to infection was established by describing the Colonization Index (CI) and Corrected Colonization Index (CCI) [4]. Though

CI and CCI provided a link between colonization and development of IC, the clinical parameters were ignored entirely. Further, the possibility of lack of colonization ruling out IC among neonates is also unanswered.

In this scenario, to predict the risk of IC, Ostrosky's Clinical Prediction Rule (OCPR) was suggested considering clinical parameters [5] but completely lacked microbiological parameters (colonization). Playford *et al*. conducted an external validation of OCPR and found a Positive Predicting Value (PPV) of 4% and a Negative Predicting Value (NPV) of 99%. So, he suggested including CCI in this OCPR where PPV was increased to 17% [6]. Similarly, Candida Score (CS) was put forth by Leon *et al*., considering both clinical and microbiological parameters [7]. Further, Leroy *et al*. externally validated this score, stating a PPV of 23.8% and NPV of 100%[8].

Prediction of IC involves multidimensional (both clinical and microbiological), multi-variate analysis that requires robust tools to establish a relationship between the risk factors and IC. Among the many statistical analytical techniques available, an Artificial Neural Network (ANN) has a strong capacity to handle multiple variables involving multidimensional complexities in medical diagnostics [9]. However, ANN has not been tried to predict the risk of IC in the literature. To address this gap, the present work aims to identify the most significant risk factors, develop a model based on ANN, and compare it with various risk prediction models to predict IC.

Material and Methods

Study design

A prospective, hospital-based, observational study was undertaken in the Department of Microbiology and Neonatal Intensive Care Unit, Department of Paediatrics, Raichur Institute of Medical Sciences, over a period of one year from March 2016 to February 2017 after obtaining clearance from Institutional Ethics Committee, Raichur. All neonates whose parents/ guardians were willing to let their baby participate and whose birth weight was less than 1500g were included in the study after taking informed written consent. Infants with one minute Apgar score of less than 4, congenital disabilities, and skin disorders were excluded from the study. The study population was selected using the non-probability sampling technique. The gender, mode of delivery, complications during delivery, age of gestation, birth weight, respiratory distress, need for mechanical ventilation, surgery, and TPN were collected via questionnaire.

Methods

Commercially available pre-sterilized cotton swabs from HiMedia laboratory were used for sampling. The swabs from the skin (axilla), umbilicus, respiratory tract, and gastrointestinal tract (rectum) were inoculated on Sabouraud Dextrose Agar (SDA) and incubated aerobically at 37^0C for 48 hours. The growth was identified by morphology study, germ tube production, and carbohydrate fermentation and assimilation test [10]. Growth was enumerated semi-quantitatively as light (< 10 colonies), moderate (10-100 colonies), or heavy (> 100 colonies) [11]. Blood samples were collected from all neonates under aseptic precautions. The volume of blood taken for blood culture was 1 ml for neonates having a weight less than 1000 grams and 2 ml for weight between 1000 to 1500 grams [12]. These blood culture bottles were observed for at least one week

before declaring sterile. Any growth isolated from these bottles were subcultured on SDA and further identified by the conventional method[10].

Statistical analysis

χ^2 test

χ^2 test of significance was applied to know the association of various categorical variables like birth weight, gestation age, mode of delivery, respiratory distress, diabetes, use of steroids, prolonged use of antibiotics with colonization and IC. A p-value $\leqslant 0.05$ was considered significant at 95% confidence level.

IC Prediction models

Various models were used to predict IC, which include CI, CCI, OCPR, and CS. The scores and cut-off for the prediction models were calculated as shown in Table 1.

Table 1: Various models to predict invasive candidiasis

Risk Model	Components	Cut-off
CI[4]	$CI = \frac{No.of\ non-blood\ sites\ colonized}{No.of\ body\ sites\ cultured}$	Score 0.5
CCI[4]	$CCI = CI \times \frac{No.of\ sites\ heavily\ colonized}{No.of\ sites\ colonized}$	Score 0.4
OCPR[5]	ICU and mechanically ventilated (48 h), antibiotic use and CVC (1-3 days), and at least two of the following: (i) Surgery (0-7 days) (ii) TPN (1-3 days) (iii)Immunosuppressive use (0-7 days) (iv)Pancreatitis (0-7 days) (v) Dialysis (1-3 days) (vi)Use of steroids (3-7 days)	Satisfying conditions are termed as 'positive'
CS[7]	Clinical sepsis – 2 points TPN – 1 point Surgery – 1 point Multifocal colonization – 1 point	Score 3

CI – Colonization Index, CCI – Corrected Colonization Index, OCPR - Ostrosky's Clinical Prediction Rule, CS – Candida Score

ANN prediction model

ANN modelling was used in this work to derive a relationship between the various clinical and microbiological factors that were found significant from the χ^2 test, to predict the risk of IC. The neural network was developed with an input layer (number of nodes corresponding to significant risk factors), one hidden layer (10 nodes) and one output layer with a single node corresponding to IC. The values entered in the model were either one or zero, corresponding to a response of "yes" or a "no". The model was trained using the Levenberg-Marquardt algorithm and was retrained until the regression coefficients for training, testing, and validation were more than 0.8, and the mean square error was less than 0.01. Out of the 103 samples, 70% were used for training, 15% for testing, and the remaining 15% for validation. The developed model was then used to compare with the existing data, and any output showing a value 0.5 was considered as a risk to IC. The modelling was carried out using the neural network tool in MATLAB 2015 software.

Interpretation of risk prediction models

The agreement of the risk prediction models was determined by calculating sensitivity, specificity, PPV, NPV, positive and negative likelihood ratios (PLR, NLR)[13].

Results

Sociodemographic and clinical data

The study population consisted of 103 neonates; 60 were males and the remaining 43 females. VLBW neonates were 87.4%, and the remaining 12.6% were ELBW. Seventy-five (72.8%) were preterm, and 28(27.2%) were term neonates. Twenty-one were delivered by lower segment caesarean section and 82 by vaginal delivery. Among 13 neonates who developed respiratory distress, 9(69.2%) were preterm, and 5(38.5%) were ELBW. Seven had diabetes mellites, and among them, 5(71.4%) had respiratory distress. Fifteen neonates were on prolonged use of antibiotics.

Distribution of colonization

The swabs collected from the skin, umbilicus, respiratory tract, and gastrointestinal tract showed 78 (75.7%) neonates colonized with Candida species. Ten out of 13 ELBW and 65 out of 90 VLBW revealed colonization. Among 82 who had vaginal delivery, 78% were colonized. Among 78 neonates who were colonized, 71.8% had multi-focal colonization, 14.1% had RD, 16.7% were on antibiotics, and 9% were diabetic.

Various Candida species isolated from the colonized sites of 78 neonates included *C. albicans* (43.8%), *C. parapsilosis* (25.8%), *C. tropicalis* (12.4%), *C. glabrata* (10.7%) and *C. krusei* (7.3%). The highest growth was obtained from the skin (66) and the least from the respiratory tract (14).

Invasive candidiasis

The blood cultured from 103 neonates revealed IC among 21(20.4%) neonates. The distribution of various species of Candida obtained from blood cultures are depicted in fig. 1. The highest isolate obtained was *C. albicans* - 9(43.8%), followed by *C. parapsilosis* - 5 (23.8%).

χ^2 test results

The results and the association between various variables applied using χ^2 tests are shown in Table 2.

Prediction of IC

From the χ^2 test, a significant association with IC was noted for gestation age, mode of delivery, respiratory distress, diabetes, prolonged use of

antibiotics, and colonization. These factors were used as input nodes to the ANN model, where colonization was represented by CI. The neural network diagram of the model is shown in fig. 2. The sensitivity, specificity, PPV, NPV, PLR, and NLR were determined for each risk prediction model, as shown in Table 3. Among all the models, the highest PPV was found for the ANN model (83.3%), followed by CS (76.5%). ANN satisfied agreeable values for PPV, NPV, sensitivity, and specificity.

Table 2: Association of different factors with colonization and IC

χ^2 test variables		Probability value	Association
Factors contributing to candidiasis	**Candida growth**		
Birth weight	Colonization	0.459	Not Significant
Gestation age	Colonization	0.013	Significant
Mode of delivery	Colonization	0.000	Significant
Diabetes	Colonization	0.000	Significant
Prolonged use of antibiotics	Colonization	0.002	Significant
Respiratory distress	Colonization	0.010	Significant
Use of steroids	Colonization	0.118	Not Significant
Gestation age	IC	0.001	Significant
Mode of delivery	IC	0.000	Significant
Diabetes	IC	0.000	Significant
Prolonged use of antibiotics	IC	0.000	Significant
Respiratory distress	IC	0.001	Significant
Multifocal colonization	IC	0.000	Significant

IC – Invasive Candidiasis

Table 3: Results of various risk prediction models

Prediction Model	PPV (%)	NPV (%)	Sensitivity (%)	Specificity (%)	PLR	NLR
CI	37.5	100.0	100.0	57.3	2.3	0.0
CCI	72.7	93.8	76.2	92.7	10.4	0.3
OCPR	66.7	80.2	9.5	98.8	7.8	0.9
CS	76.5	90.7	61.9	95.1	12.7	0.4
ANN model (present work)	83.3	98.7	95.2	95.1	19.5	0.1

(Data presented as Number (%);PB: Polymyxin B; SXT: Co-trimoxazole; AN: Amikacin; IPM: Imipenem; CP: Ciprofloxacin; GM: Gentamicin; CAZ: Ceftazidime; NA: Nalidixic Acid; AM: Ampicillin; FM, Nitrofurantoin

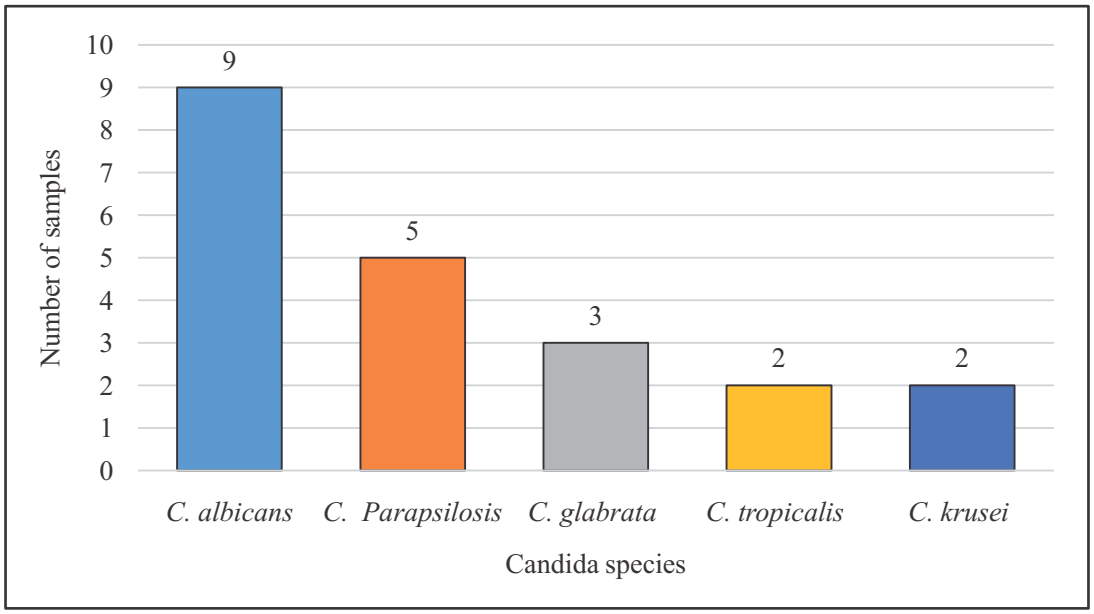

Figure 1: Distribution of various species of Candida obtained from the blood culture

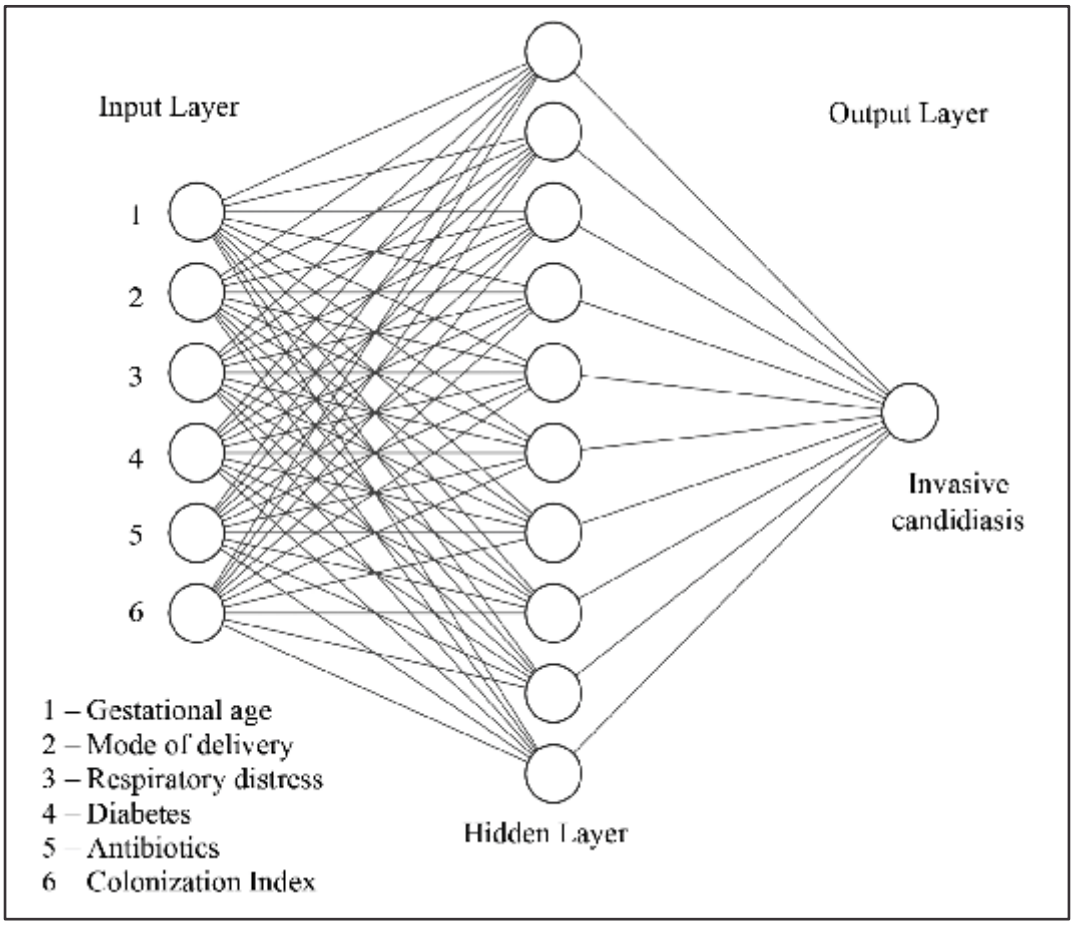

Figure 2: The neural network model used in the current work

Discussion

Factors affecting colonization

The present study shows 78 neonates colonized with Candida species. These findings are similar to studies where the colonization rate ranged from 60-76.2%[14,15]. Higher colonization in neonates could be attributed to the increased trans-epidermal water loss from neonatal skin, creating a moist environment that facilitates fungal colonization[16].

The colonization was highest among preterms (77.3%) and vaginal delivery (78%). χ^2 tests showed a significant association of gestation age, mode of delivery with colonization proving that colonization is inversely proportional to gestation age and directly related to colonization of birth canal. The findings are similar to studies where colonization was higher among preterm [17] and vaginal delivery [18]. Higher colonization among preterm could be due to functionally incompetent epidermis barrier, absence of vernix caseosa, developmental defects in immune function, and weak cutaneous defences[19]. Higher colonization by vaginal delivery could be due to colonization of the birth canal which transmits vertically to the mouth of newborn during labour[20].

A significant association was noted among prolonged use of antibiotics, respiratory distress, and diabetes with colonization, suggesting that these factors play an important role on colonization and are similar to studies in the literature[3,21].

Colonization pattern

C. albicans (43.8%) was the most common isolate obtained from colonized sites, followed by *C. parapsilosis* (25.8%), as seen in the literature [21-23]. Further study by David *et al*. [24] also suggested that centres that do not use antifungal prophylaxis, *C. albicans* and *C. parapsilosis* account for the majority of cases. Skin (37.1%) was the most commonly colonized site, followed by the gastrointestinal tract (31.5%), as seen in the literature [25]. The higher colonization on the skin could be due to increased moisture or diffused glucose on the epithelial surface of the skin[26].

Invasive candidiasis

Among 103 neonates, 56 had multifocal colonization, and among them, 21 developed IC. Statistical analysis showed a significant association between multifocal colonization and IC, suggesting multifocal colonization as a risk marker for determining IC similar to studies in the literature [11,28]. This could be because Candida species produce proteases and phospholipases that lyse the thin keratin and lipid membrane of the neonatal skin, which facilitates epithelial invasion [28]. A significant association was also noted with IC for factors of gestation age, mode of delivery, diabetes, prolonged use of antibiotics, and respiratory distress suggesting these factors also contribute to the development of IC.

C. albicans (43.8%) was the most common isolate obtained from blood cultures, followed by *C. parapsilosis* (23.8%), which is comparable to the results as seen in the literature [21,29]. This could be attributed as a consequence of increased colonization, together with a local or generalized defect in host defences leading to invasion into the bloodstream [25]. Further, mechanical disruption of fragile premature skin during handling and medical procedures associated with barrier perturbation increases the risk of IC following colonization[30].

Prediction of IC

For an ideal predictive model, PPV, NPV, sensitivity, and specificity values should be the highest (100%). The only model that appropriately satisfies all the parameters is the ANN model, with NPV, Sensitivity and Specificity >95%. PPV is 83.3% but is still the highest compared to the other models. PPV and sensitivity depict a model's closeness to the number of true cases having a disease, which is not satisfactorily predicted by CI, CCI, OCPR, and CS. Although CI has a sensitivity of 100%, its PPV is relatively poor (37.5%). Similarly, the values of PLR and NLR are required to be maximum and minimum, respectively, which is greatly satisfied by the ANN model compared to the others. The main reason that ANN can more accurately predict the disease is its comprehensive modelling design that accommodates various weights and biases to all the risk factors to carefully predict the IC with low errors[9].

Conclusion

The present work focused on assessing risk factors involved in predicting Invasive Candidiasis among neonates using a novel ANN model and was compared with various prediction models like Colonization Index, Corrected Colonization Index, Ostrosky's Clinical Prediction Rule, and

Candida Score. High values of Positive Predicting Value, Negative Predicting Value, sensitivity, specificity, and positive likelihood ratio were seen with the present ANN model suggesting it to be superior to other models.

Acknowledgement

The authors would like to thank Dr Sooraj Mohan, National Institute of Technology Goa, for his valuable support in statistical analysis and ANN modelling.

References

1. Zaoutis TE, Argon J, Chu J, Berlin JA, Walsh TJ, Feudtner C. The epidemiology and attributable outcomes of candidemia in adults and children hospitalized in the United States: A propensity analysis. *Clin Infect Dis* 2005;41(9):1232-1239.
2. Amboiram P, Balakrishnan U, Ninan B, Ramaswamy S, Ashok C, Kumar KS. Incidence of invasive candidal infection in very low birth weight neonates over a period of 5-year: A single institutional study. *Indian J Child Health* 2016;3(3):191-195.
3. Ahmed A, Azim A, Baronia AK, Marak KRSK, Gurjar M.Risk prediction for invasive candidiasis. *Indian J Crit Care Med* 2014;18(10):682-688.
4. Pittet D, Monod M, Suter PM, Frenk E, Auckenthaler R. Candida colonization and subsequent infections in critically III surgical patients. *Ann Surg* 1994; 220(6)751-758.
5. Ostrosky-Zeichner L, Sable C, Sobel J, Alexander BD, Donowitz G, Kan V, *et al*. Multicenter retrospective development and validation of a clinical prediction rule for nosocomial invasive candidiasis in the intensive care setting. *Eur J Clin Microbiol Infect Dis* 2007;26(4):271-276.
6. Playford EG, Lipman J, Kabir M, McBryde ES, Nimmo GR, Lau A, *et al*. Assessment of clinical risk predictive rules for invasive candidiasis in a prospective multicentre cohort of ICU patients. *Intensive Care Med* 2009;35(12):2141-2145.
7. León C, Ruiz-Santana S, Saavedra P, Almirante B, Nolla-Salas J, Álvarez-Lerma F, *et al*. A bedside scoring system ("Candida score") for early antifungal treatment in nonneutropenic critically ill patients with Candida colonization. *Crit Care Med* 2006;34(3): 730-737.
8. Leroy G, Lambiotte F, Thévenin D, Lemaire C, Parmentier E, Devos P, *et al*. Evaluation of "Candida score" in critically ill patients: a prospective, multicenter, observational, cohort study. *Ann Intensive Care* 2011;1(1):50.
9. Amato F, López A, Peña-Méndez EM, Vaňhara P, Hampl A, Havel J. Artificial neural networks in medical diagnosis. *J Appl Biomed* 2013;11(2):47-58.
10. Walsh TH, Hayden, Randall T, Larone DH. Larone's Medically Important Fungi: A guide to identification. 6th ed. Washinton, D. C.: ASM Press; 2018.
11. Lau AF, Kabir M, Chen SCA, Playford EG, Marriott DJ, Jones M, *et al*. Candida colonization as a risk marker for invasive candidiasis in mixed medical-surgical intensive care units: development and evaluation of a simple, standard protocol. *J Clin Microbiol* 2015;53(4):1324-1330.
12. Schelonka RL, Chai MK, Yoder BA, Hensley D, Brockett RM, Ascher DP. Volume of blood required to detect common neonatal pathogens. *J Pediatr* 1996;129(2):275-278.
13. Shreffler J, Huecker MR. Diagnostic testing accuracy: sensitivity, specificity, predictive values and likelihood ratios. StatPearls Publishing LLC; 2021.
14. Hamzavi SS, Amanati A, Badiee P, Kadivar MR, Jafarian H, Ghasemi F, *et al*. Changing face of Candida colonization pattern in pediatric patients with hematological malignancy during repeated hospitalizations, results of a prospective observational study (2016–2017) in Shiraz, Iran. *BMC Infect Dis* 2019;19(1):759.

15. Bansiwal K, Binnani A, Gupta PS. Study of Candida colonization and candidemia in neonates admitted in neonatal intensive care unit at a tertiary care institute of Bikaner, Rajasthan, India. *Int J Curr Microbiol Appl Sci* 2016;5(10).
16. Cartlidge P. The epidermal barrier. *Semin Neonatol* 2000;5(4):273-280.
17. Kaufman DA, Gurka MJ, Hazen KC, Boyle R, Robinson M, Grossman LB. Patterns of fungal colonization in preterm infants weighing less than 1000 grams at birth. *Pediatr Infect Dis J* 2006; 25(8):733-737.
18. Ali GY, Algohary EHSS, Rashed KA, Almoghanum M, Khalifa AA. Prevalence of Candida colonization in preterm newborns and VLBW in neonatal intensive care unit: role of maternal colonization as a risk factor in transmission of disease. *J Matern Neonatal Med* 2012;25(6):789-795.
19. Fluhr JW, Darlenski R, Taieb A, Hachem JP, Baudouin C, Msika P, *et al*. Functional skin adaptation in infancy – almost complete but not fully competent. *Exp Dermatol* 2010;19(6):483-492.
20. Al-Rusan RM, Darwazeh AMG, Lataifeh IM. The relationship of Candida colonization of the oral and vaginal mucosae of mothers and oral mucosae of their newborns at birth. *Oral Surg Oral Med Oral Pathol Oral Radiol* 2017;123(4):459-463.
21. Wadile R, Bhate V. Study of clinical spectrum and risk factors of neonatal candidemia. *Indian J Pathol Microbiol* 2015;58(4):472-474.
22. Rafat S. Comparision of species disribution, antifungal susceptibility and virulence factor in pregnant and non-pregnant women with vulvo-vaginal candidiasis. *J Krishna Inst Med Sci Univ* 2021;10(4):80-88.
23. Rajmane V, Rajmane S, Mohite S, Patil V, Ghatole M. Incidence, risk factors and susceptibility profile of candida species isolated from blood of non-neutropenic medical intensive care unit patients in a tertiary care centre. *J Krishna Inst Med Sci Univ* 2016;5(1):45–33.
24. David T, Attila G, D. NJ. Candida parapsilosis, an Emerging Fungal Pathogen. *Clin Microbiol Rev* 2008;21(4):606-625.
25. Pappas PG, Lionakis MS, Arendrup MC, Ostrosky-Zeichner L, Kullberg BJ. Invasive candidiasis. *Nat Rev Dis Prim* 2018;4(1):18026.
26. Chiou YB, Blume-Peytavi U. Stratum Corneum Maturation. *Skin Pharmacol Physiol* 2004;17(2):57-66.
27. Alenazy H, Alghamdi A, Pinto R, Daneman N. Candida colonization as a predictor of invasive candidiasis in non-neutropenic ICU patients with sepsis: A systematic review and meta-analysis. *Int J Infect Dis* 2021;102:357-362.
28. Evans NJ, Rutter N. Development of the Epidermis in the Newborn. *Neonatology* 1986;49(2):74-80.
29. Rajmane V, Rajmane S, Mohite S, Raje V. Study of candida bloodstream infections in surgical intensive care unit patients and susceptibility profile of the isolates. *J Krishna Inst Med Sci Univ* 2015;4(3):82-88.
30. Visscher MO, Carr AN, Narendran V. Premature infant skin barrier maturation: status at full-term corrected age. *J Perinatol* 2021;41(2):232-239.

Author for Correspondence:
Dr. Sheetal Gouda, Department of Microbiology, Karwar Institute of Medical Sciences, Karwar- 581301 (Karnataka) India
Email: sheetalkgouda@gmail.com Cell: 8747909300

How to cite this article:
Gouda S, Sathyajith R, Peerapur BV. A novel approach to predict the risk of invasive candidiasis using artificial neural networks and comparison with other models. *J Krishna Inst Med Sci Univ* 2022; 11(4):10-19

Submitted: *17-July-2022* ***Accepted:*** *15-Sep-2022* ***Published:*** *01-Oct-2022*