Acıbadem Univ. Sağlık Bilim. Derg. 2022; 13 (2): 213-219
https://doi.org/10.31067/acusaglik.994754

ORIGINAL ARTICLE / ARAŞTIRMA YAZISI Biochemistry and Molecular Biology / Biyokimya ve Moleküler Biyoloji

Gram Stain Prediction with Machine Learning Techniques Using Biochemical Parameters in ICU Patients with Urinary Tract Infection

Hakan Ayyıldız[1], Seda Arslan Tuncer[2], Mehmet Kalaycı[1], Rojda Aslan[3]

[1]Fethi Sekin City Hospital, Biochemistry Department, Elazig, Turkey
[2]Firat University, Software Engineering, Elazig, Turkey
[3]Fethi Sekin City Hospital, Microbiology Department, Elazig, Turkey

Hakan AYYILDIZ
Seda ARSLAN TUNCER
Mehmet KALAYCI
Rojda ARSLAN

Correspondence: Hakan Ayyıldız
Fethi Sekin City Hospital, Biochemistry Department, Elazig, Turkey
Phone: +905057791672
E-mail: hknayyildiz@hotmail.com

Received: 13 September 2021
Accepted: 30 January 2022

ABSTRACT

Purpose: The aim of this study was to develop a useful algorithm based on complete blood count (CBC), urinalysis, and biochemical parameters that could be an alternative to Gram staining in the prediction of UTI and the determination of initial antibiotic treatment in ICU patients.

Methods: All the specimens included in the study were obtained from ICU patients and were subjected to Gram staining in the laboratory. Simultaneously, CBC, urinalysis, and biochemical tests were performed for each specimen. A classification based on biochemical parameters was performed for the estimation of Gram-negative and Gram-positive bacteria, as an alternative to Gram staining.

Results: Classification was achieved using multiple classification systems including Artificial Neural Networks (ANN), Support Vector Machine (SVM), the K-Nearest Neighbors (KNN), and Decision Tree Language (DTL) and the best classification performance was achieved by ANN, with an accuracy of 84.6%, a sensitivity of 88.5%, and a specificity of 73.5%.

Conclusion: The high specificity and accuracy of the algorithm indicated that this algorithm can be effectively used in the selection of empirical antibiotic treatment for ICU patients with UTI and can provide more advanced and technological opportunities by combining laboratory parameters with machine learning techniques.

Keywords: Gram Stain, Laboratory, Machine Learning, Urinary Tract Infections

İdrar Yolu Enfeksiyonu olan Yoğun Bakım Hastalarında Biyokimya Parametreleri Kullanarak Makine Öğrenmesi ile Gram boyama tahmini

ÖZET

Amaç: Bu çalışma, yoğun bakımda yatan ve idrar yolu enfeksiyonu bulunan hastalarda başlangıç tedavi seçimine rehberlik etmek için Gram boyamaya alternatif olabilecek, biyokimya tetkikleri ile oluşturulmuş makine öğrenmesi algoritmalarının etkililiğine odaklanmıştır.

Yöntem: Çalışmaya alınan örnekler laboratuvardan Gram boyama isteği yapılmış ve Gram boyamaya eş zamanlı TİT, CBC ve CRP isteği yapılmış 203 yoğun bakım hastasından geriye yönelik oluşturulmuştur. Çalışmada, biyokimya laboratuvar parametreleri kullanarak Gram boyamaya alternatif, Gram-negatif ve Gram-pozitif bakteri tahmini için sınıflandırma yapılmıştır.

Bulgular: Sınıflandırmada ANN (Artificial Neural Networks), SVM (Support vector machine), KNN (K-Nearest Neighbors), DTL (Decision Tree Learning) gibi birçok sınıflandırıcı kullanılmış olup, en yüksek sınıflandırma başarısına ANN ile ulaşılmıştır. Sınıflandırma sonucu ANN ile %84.6 Acc, %88.5 Sn, %73.5 Sp değerlerine ulaşılmıştır.

Sonuç: Özellikle algoritmanın yüksek özgüllüğü ve doğruluğu, yoğun bakım hastalarında idrar yolu enfeksiyonu ampirik antibiyotik tedavisi seçiminde kullanılabileceğini, laboratuvar testlerinin makine öğrenmesi metotları ile harmanlanarak daha gelişmiş ve teknolojik imkanlar sağlayabileceğini göstermiştir.

Anahtar sözcükler: İdrar yolu enfeksiyonu, Makine Öğrenmesi, Biyokimya, Gram boyama, Yoğun Bakım Hastaları

Copyright © 2021 the Author(s). Published by Acibadem University. This is an open access article licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives (CC BY-NC-ND 4.0) International License, which is downloadable, re-usable and distributable in any medium or format in unadapted form and for noncommercial purposes only where credit is given to the creator and publishing journal is cited properly. The work cannot be used commercially without permission from the journal.

Urinary tract infections (UTIs) are the primary cause of sepsis in an intensive care unit (ICU) and urosepsis can be seen in up to 16% of ICU patients. Catheter-associated UTI is the most common cause of ICU-acquired infection and accounts approximately for 40% of hospital-acquired infections (1-5). Additionally, UTI has been shown to be the most common infection in ICU patients (4).

Prompt treatment with an effective antibiotic therapy will decrease the mortality risk in patients suspected of (not with) sepsis (6). In this regard, appropriate antibiotic selection criteria are highly important in terms of both treatment and antibiotic resistance.

Gram-negative bacteria (GNB) are the most common organisms causing ICU-acquired UTI. Primary risk factors in these patients include catheterization, length of stay in ICU, ineffective antibiotic use, and gender. In the empirical antibiotic therapy in patients with suspected ICU-acquired UTI, the initial choice of antibiotics usually covers GNB and includes aminoglycosides, β-lactam, and a β-lactamase inhibitor or carbapenem antibiotics. However, it often does not cover Gram-positive bacteria (7).

In the patients with ICU-acquired UTI, the selection of appropriate organism-based antibiotics can be life-saving at the early stages of the infection. Moreover, it can also preserve the effectiveness of existing antibiotics, thereby slowing down the development of antibiotic resistance which has increased in recent years. The World Health Organization (WHO) published a global action to optimize the use of antimicrobial agents and to improve awareness and understanding of antibiotic resistance. In addition, although administering empirical antibiotics within a few hours after hospital admission reduces mortality in critically ill patients, it may have adverse effects on broad-spectrum antibiotic resistance (7-10).

Although culture specimens obtained from ICU patients are the golden standard for identifying organisms responsible for UTI, it takes approximately 48-72 hours to obtain the results of culture and antibiotic susceptibility tests. Accordingly, shortening this time period and initiating the initial effective and rapid treatment according to Gram staining results is highly important.

Transportation of the specimens obtained for Gram staining to the laboratory and their processing and microscopic examination are time-consuming activities that increase laboratory workload. For this reason, careful and early detection of the infection and prompt initiation of appropriate antibiotic treatment for the causative agent will increase the success of the treatment and will also reduce the complications.

Artificial intelligence can be effective in the improvement of healthcare services due to its effect on modeling and decision-making processes. In addition to the growing importance of evidence-based medicine applications, various machine learning systems can perform clinical diagnosis and even offer treatment recommendations (11). The number of promising studies conducted on this issue is increasing continuously (12,13).

The aim of this study was to develop a useful algorithm based on complete blood count (CBC), urinalysis, and biochemical parameters that could be an alternative to Gram staining in the prediction of UTI and the determination of initial antibiotic treatment in ICU patients and could provide faster results and thus could reduce the workload of laboratories. This algorithm will be a cost-effective technological alternative to Gram staining and culture analysis and even to the algorithms used for predicting antibiotic resistance and will also shed light on studies investigating personalized prediction algorithms. Additionally, to our knowledge, there has been no study reporting on the determination of bacterial species in UTI by machine learning methods using CBC, urinalysis, and biochemical parameters.

MATERIALS AND METHODS

This study was conducted through a collaboration between Elazig City Hospital Biochemistry and Microbiology Departments and Fırat University Software Engineering Department. This retrospective study was performed by analyzing the data obtained from the laboratory information system in Elazig Fethi Sekin City Hospital that had been recorded between March 2019 and May 2020.

All the specimens included in the study were obtained from ICU patients and were subjected to Gram staining in the laboratory. Simultaneously, complete blood count (CBC), urinalysis, and biochemical tests were performed for each specimen. Patients with negative Gram staining results and patients whose biochemistry tests were not conducted at the same time with Gram staining were excluded from the study. The 202 patients included in the study comprised 108 women (age, 18-97 years) and 94 men (age, 18-95 years). The demographical characteristics of the patients are shown in Table 1.

Table 1. Demographical characteristics of the patients

	Min	Max	Median	Interquartiles (25%-75%)
AGE	18	97	76	59,2-83
WBC	4,6	31,2	10	8,0-12,6
NEU	1,86	29,8	6,57	4,96-9,47
LYM	0,2	8,51	1,66	1,10-2,48
MON	0,16	3,24	0,71	0,55-1
EOS	0,001	4,15	0,14	0,04-0,3
BAS	0,001	0,69	0,06	0,04-0,1
USGT	1.002	1.033	1.013	1.011-1.018
Ph	5	9	6,5	5,5-7,5
CRP	1,07	458	62,05	23,6-102

WBC: White Blood Cell, NEU: Neutrophil, LYM: Lymphocyte, MON: Monocyte, EOS: Eosinophil, BAS: Basophile, USGT: Urine Spesific Gravite Test, CRP: C Reactive Protein

Complete blood count (CBC) was performed using a Unicel DXH-800 hematology analyzer (Beckman Coulter, Inc., Brea, CA, USA). CRP levels were analyzed by the spectrophotometric method using an Image 800 analyzer (Beckman Coulter, Inc., Brea, CA, USA). Urinalysis were performed using the Iris iQ200 Sprint automated urine microscopy analyzer (Beckman Coulter, Inc., Brea, CA, USA). Table 2 presents the reference ranges for CBC, urinalysis, and biochemical parameters.

Table 2. Reference ranges

Parameter	Variable	Reference Range
CBC	WBC	3.6-11 (10^9/L)
	Neutrophil count	1.7-7.6 (10^9/L)
	Lymphocyte count	1.0-3.2 (10^9/L)
	Monocyte count	0.3-1.1 (10^9/L)
	Eosinophil count	0-0.5 (10^9/L)
	Basophile count	0-0.1 (10^9/L)
Urinalysis	Density	1.005-1.025
	pH	4.5-8
	Nitrite	Negative
	Erythrocyte	0-4
	Leukocyte	0-4
Biochemical	CRP	0-8 mg/L

WBC: White blood cell count, CRP: C-reactive protein

The extent of positive Gram staining was estimated using the measurements of CBC, urinalysis, and biochemical parameters. To achieve this, a database was created based on the collected data and then classification was performed according to the extent of intensity of Gram staining using 12 input parameters (CBC, urinalysis, and CRP). We selected features that are known to be associated with urinary tract infection and that are made up of routine and easily accessible biochemistry laboratory parameters that are studied daily. The classification was achieved using multiple classification systems including Artificial Neural Networks (ANN), Support Vector Machine (SVM), the K-Nearest Neighbors (KNN), and Decision Tree Language (DTL) and the best classification performance was achieved by ANN. Figure 1 illustrates the flow diagram used for the classification.

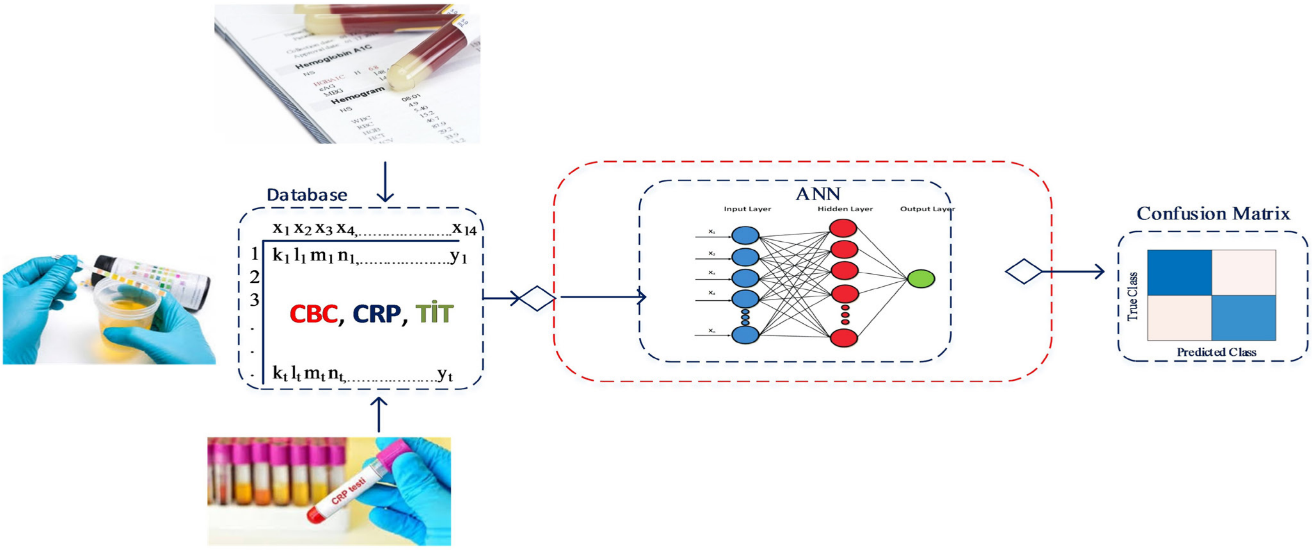

Figure 1. Flow diagram used for classification

Classification

Artificial Neural Networks (ANN) is an information processing technology imitating the working and learning skills of the human brain (14). To date, numerous algorithms related to ANN have been developed and used in many areas. In the present study, the CBC, urinalysis, and CRP values were classified according to the families of the bacteria that caused infection in the urine culture using the Backpropagation Neural Network.

Backpropagation is a widely used method for training parameters in ANN. The generalized delta rule is a method that allows some of the difference (error) between the target values and the network output (error) to be reflected back to each training cell and thereby to change the weights according to the error and to repeat this process for a certain number of times during the training of backpropagation networks, in attempts to obtain the lowest error possible (15). Figure 2 illustrates the structure of backpropagation neural network.

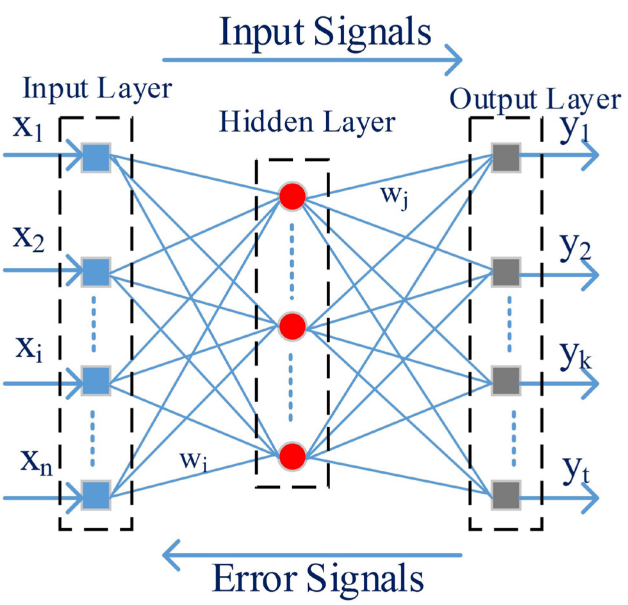

Figure 2. Multilayer backpropagation neural network

The k-fold cross validation was used to minimize distribution-related errors in the training and testing phase of the model proposed in the study. To achieve this, the training dataset was randomly divided into k number of subsets, whereby k-1 number of subsets were used for training and the remaining one subset was used for testing. The same procedure was repeated k times and the values obtained at each repetition were summed up and averaged and then the performance of the model was calculated. Figure 3 illustrates the structure of ANN.

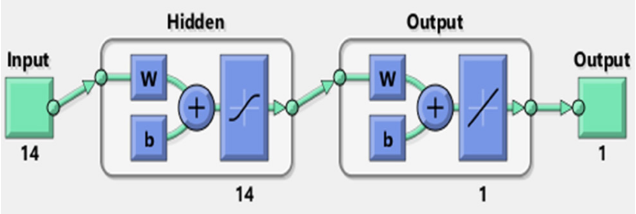

Figure 3. Structure of the ANN model used in the study

Performance evaluation

The performance of the classification was assessed using the following parameters: Sensitivity, Specificity, Precision, Negative Predictive Value (NPV), False Positive Rate (FPV), False Discovery Rate (FDR), False Negative Rate (FNR), Accuracy, and F score. Table 3 (16).

Table 3. Performance evaluation criteria

Measure	Formula
ACC	(TP + TN) / (TP + TN + FN + FP)
SN	TP / (TP + FN)
SP	TN / (TN + FP)
FPR	FP / (TN + FP)
PREC	TP / (TP + FP)
F1	2 * PREC * REC / (PREC + REC)

Acc: Accuracy, SN: Sensitivity, SP: Specificity, FPR: False positive rate; PREC: Precision: F: F score; TP: True positive: TN: True negative; FP: False positive; FN: False negative

Data Analysis

The models used in the study were tested in the Matlab R2018b (The MathWorks, Inc. Cambridge, United Kingdom) platform on a computer with an i7 9750 H CPU, 2.6GHz, 16 GB RAM and Geforce GTX 1050 gCPU. Laboratory data analysis (laboratory characteristics of patients) was performed on Jupyter Notebook using Python 3.0 (Python Software Foundation, Oregon, USA) program with Pandas library.

RESULTS

In the present study, a classification model using ANN, SVM, KNN, and DTL with the inputs obtained from biochemical parameters was proposed as an alternative to Gram staining for the prediction of UTI in ICU patients. Table 4 presents the performance of each classifier.

Table 4. Performance of the classifiers used in the study

	SN	SP	PREC	ACC	F1
ANN	0.8859	0.7358	0.9041	0.8465	0.8949
SVM	0.7684	0.8333	0.9865	0.7723	0.8639
KNN	0.7228	0.8333	0.9865	0.7290	0.8343
DTL	0.7593	0.4250	0.8425	0.6931	0.7987

SN: Sensitivity, SP: Specificity, PREC: Precision, Acc: Accuracy, F: F score

As seen in Table 3, ANN had the highest accuracy (84.6%) and sensitivity (88.5%) values. Additionally, although SVM and KNN had the highest precision value (98.7%), ANN had the highest F1 value (89.4%).

DISCUSSION

The present study investigated the effectiveness of an artificial intelligence algorithm of initial treatment selection which was created with biochemical parameters and was utilized as an alternative to Gram staining in ICU patients with UTI. This diagnostic algorithm, which was created based on urinalysis, CRP and CBC results will guide empirical treatment and the findings of the study will also provide clinicians a different alternative in the selection of broad-spectrum antibiotics without increasing the risk of treatment failure.

Antibiotic resistance remains a serious public health problem (17). Accordingly, rapid and cost-effective procedures developed in line with technological advancements are needed to decrease redundant use of broad-spectrum antibiotics and thereby to reduce the development of antibiotic resistance. The diagnostic algorithm developed in the present study can be used in ICUs as a useful high-potential tool for the selection of broad-spectrum antibiotics.

Culture analysis performed for detecting the causative agent within several days after hospitalization is the golden standard for the treatment of patients suspected with infection. However, in such patients, considering the benefits of prompt treatment, empirical antibiotic therapy is the first-line treatment particularly in patients hospitalized in ICU. Additionally, empirical antibiotic therapy has also been shown to reduce mortality in such patients (21-24).

An infection caused by multi-resistant Gram-negative bacilli results in significantly higher mortality than an infection caused by other pathogens; therefore, knowledge of the morphology of the organism in the initial treatment plan is of paramount importance. In contrast to studies reporting on a moderate correlation between Gram staining and culture (19,20), Yoshimura et al. (18) examined the effect of Gram staining on initial therapy in patients with ventilator-associated pneumonia and provided the first evidence that Gram staining could reduce the use of broad-spectrum antibiotics.

Although enteric bacteria (especially E. coli, Gram-negative) are responsible for most UTIs, there are a wide range of pathogens causing UTI. Of note, Staphylococcus saprophyticus, Enterococcus faecalis, and Streptococcus agalactiae are the most common causes of Gram-positive UTI. On the other hand, antibiotic choices of Gram-positive and Gram-negative bacteria are different from each other (25,26). Accordingly, it is highly important to develop diagnostic options that can show these differences and start treatment promptly to narrow down the treatment options. In such patients, considering the benefits of prompt treatment, empirical antibiotic therapy should be the first-line treatment particularly in patients hospitalized in ICU. Additionally, empirical antibiotic therapy has also been shown to reduce mortality in such patients (22,27,28).

The algorithm developed in the present study indicated that UTI can be predicted in ICU patients based on their routine CBC, urinalysis, and biochemical parameters with no need for Gram staining. Accordingly, this algorithm is appears to be a viable diagnostic option since it is highly cost-effective and was developed with a remarkably smaller number of datasets.

In our study, ANN, SVM, KNN, and DTL were used for classification. Since the system parameters in models such as ANN are problem-dependent, it cannot be predicted as to which of the parameters (e.g. number of layers in Multilayer Perceptrons [MLP], number of neural processors in hidden layers, learning coefficient) will provide an optimal result. Therefore, a good learning rate to be provided by these parameters must be discovered via trial and error. On the other hand, a comparison between/among classifiers are not recommendable, although an algorithm is likely to have a propensity for a particular problem.

Accuracy is calculated as the percentage of correctly classified instances to the total dataset; therefore, accuracy alone may not be sufficient in imbalanced classes. Sensitivity measures how well a test can identify true positives for people with the condition being tested, while specificity measures the ability of a test to produce an accurate negative result for people who do not have the condition being tested. Accordingly, all these three parameters should be evaluated together to obtain the best classification performance. High specificity helps prevent misunderstanding and avoidable unnecessary interventions (True Negative), while a highly sensitive test is needed particularly in ambiguous diagnosis or in cases of early disease (True Positive). The F1-score uses the harmonic mean instead of the arithmetic mean so as not to ignore extreme cases. For this reason, the F1 score was included in the evaluation metrics.

Our results indicated that the algorithm developed in the present study based on routine CBC, urinalysis, and biochemical parameters could be a useful alternative to Gram staining for the prediction of UTI in ICU patients.

CONCLUSION

Although the Gram staining technique was developed by Hans Christian Joachim Gram in the 1880s (29), it is commonly used in bacteriology laboratories. However, the time from the admission of the specimens to the staining and interpretation processes is remarkably long and depends on the technician's/laboratory specialist subjective interpretation even when there is only one specimen to be analyzed. Our findings indicated that Gram staining could be conducted with different alternatives and these alternatives may help ameliorate the laboratory workload. Accordingly, we consider that our study provides advanced and technological opportunities for the bacterial classification performed with Gram staining by combining biochemical tests with machine learning methods. We also believe that the study will shed light on future studies on this subject.

Ethical Consideration

The study protocol was approved by Fırat University Ethics Committee (Approval No.: 2020/08-41-486, Date: 29 April, 2020).

Conflict of Interest

The authors declare no conflict of interest.

REFERENCES

1. Saint S, Chenoweth CE. Biofilms and catheter-associated urinary tract infections. Infect Dis Clin North Am. 2003; 17: 411-32.
2. Warren JW. The catheter and urinary tract infection. Med Clin North Am. 1991; 75: 481-93
3. Maki DG, Tambyah PA. Engineering out the risk for infection with urinary catheters. Emerg Infect Dis. 2001; 7: 342-7.
4. Liu Y, Xiao D, Shi XH. Urinary tract infection control in intensive care patients. Medicine (Baltimore). 2018; 97(38): e12195.
5. Richards MJ, Edwards JR, Culver DH, et al. Nosocomial infections in combined medical-surgical care units in the United States. Infect Control Hosp Epidemiol. 2000; 21: 510-15.
6. Liu VX, Fielding-Singh V, Greene JD, et al. The Timing of Early Antibiotics and Hospital Mortality in Sepsis. Am J Respir Crit Care Med. 2017; 196(7): 856-63.
7. Khilnani GC, Zirpe K, Hadda V, et al. Guidelines for Antibiotic Prescription in Intensive Care Unit. Indian J Crit Care Med. 2019; 23(Suppl 1): 1-63
8. Shlaes DM, Gerding DN, John Jr JF, et al. Society for Healthcare Epidemiology of America and Infectious Diseases Society of America Joint Committee on the Prevention of Antimicrobial Resistance: guidelines for the prevention of antimicrobial resistance in hospitals. Infect Control Hosp Epidemiol. 1997; 18:275-91.
9. Kaki R, Elligsen M, Walker S, et al. Impact of antimicrobial stewardship in critical care: a systematic review. J Antimicrob Chemother. 2011; 66:1223–30.
10. WHO. Global action plan on antimicrobial resistance. 2015. http://apps.who.int/iris/bitstream/10665/193736/1/9789241509763_eng.pdf. Acsessed 1 Nov 2018.
11. Kavvas ES, Catoiu E, Mih N, et al. Machine learning and structural analysis of Mycobacterium tuberculosis pan-genome identifies genetic signatures of antibiotic resistance. Nat Commun. 2018; 9(1): 4306.
12. Richardson A, Signor BM, Lidbury BA, et al. Clinical chemistry in higher dimensions: Machine-learning and enhanced prediction from routine clinical chemistry data. Clin. Biochem. 2016; 49; 1213-20.
13. Wildenhain J, Spitzer M, Dolma S, et al. Prediction of Synergism from Chemical-Genetic Interactions by Machine Learning. Cell Syst. 2015;1(6):383-95
14. Congjie He, Meng Ma, Ping Wang. Extract Interpretability-Accuracy balanced Rules from Artificial Neural Networks: A Review. Neurocomputing. 2020; 387: 346-58.
15. Xu M, Papageorgiou DP, Abidi SZ, et al. A deep convolutional neural network for classification of red blood cells in sickle cell anemia. PLoS Comput Biol. 2017; 13(10): e1005746.
16. H Ayyıldız, SA Tuncer. Determination of the effect of red blood cell parameters in the discrimination of iron deficiency anemia and beta thalassemia via Neighborhood Component Analysis Feature Selection-Based machine learning Chemometrics and Intelligent Laboratory Systems 196, 103886,2020
17. Munita JM, Arias CA. Mechanisms of Antibiotic Resistance. Microbiol Spectr. 2016 Apr;4(2):10.1128/microbiolspec.VMBF-0016-2015.
18. Yoshimura J, Kinoshita T, Yamakawa K, et al. Impact of Gram stain results on initial treatment selection in patients with ventilator-associated pneumonia: a retrospective analysis of two treatment algorithms. Crit Care. 2017; 21(1): 156.
19. O'Horo JC, Thompson D, Safdar N. Is the Gram stain useful in the microbiologic diagnosis of VAP? A meta-analysis. Clin Infect Dis. 2012; 55:551–61.
20. Tetenta S, Metersky ML. Tracheal aspirate Gram stain has limited sensitivity and specificity for detecting Staphylococcus aureus. Respirology. 2011; 16:86–9.

21. Sartika IN, Suarta K, Ardhani P. Diagnostic value of urine Gram staining for urinary tract infection in children. Paediatrica Indonesiana 2009; 49(4): 205-8.
22. Díaz-Martín A, Martínez-González ML, et al. Antibiotic prescription patterns in the empiric therapy of severe sepsis: combination of antimicrobials with different mechanisms of action reduces mortality. Crit Care. 2012; 16(6): R223.
23. Tseng CC, Huang KT, Chen YC, et al. Factors predicting ventilator dependence in patients with ventilator-associated pneumonia. Scientific World Journal. 2012; 2012: 547241
24. Çelikel T. Sepsis: Genel Bakış. Yoğun Bakım Dergisi 2005;5(2):73-74.
25. Wilson ML, Gaido L. Laboratory diagnosis of urinary tract infections in adult patients. Clin Infect Dis. 2004; 38(8): 1150-8
26. Urmi U L, Jahan N, Nahar S, et al. Gram-positive uropathogens: Empirical treatment and emerging antimicrobial resistance. Biomed Res Clin Prac 2019; 4: 1-4.
27. Sever C, Kulahci Y, Duman H. Prediction of mortality and causes of death in a burn centre: a retrospective clinical study. J Clin Anal Med 2011;2(3): p. 24-6
28. Ak O, Batirel A, Ozer S, Çolakoğlu S. Nosocomial infections and risk factors in the intensive care unit of a teaching and research hospital: a prospective cohort study. Med Sci Monit. 2011; 17(5): PH29-34
29. https://www.newscientist.com/article/2216418-hans-christian-gram-the-biologist-who-helped-investigate-bacteria/