Using artificial neural networks for pattern recognition of post-surgical infections

Uso de redes neurais artificiais para reconhecimento de padrões de infecções pós-cirúrgicas

DOI:10.34119/bjhrv6n1-260

Recebimento dos originais: 09/01/2023
Aceitação para publicação: 06/02/2023

Fernanda Tomé
Bachelor in Production Engineering
Institution: Universidade de Caxias do Sul (UCS)
Address: Rua Francisco Getúlio Vargas, 1130, Bloco G
E-mail: ftome1@ucs.br

Lessandra Michelin
PhD in Biotechnology
Institution: Universidade de Caxias do Sul (UCS)
Address: Rua Francisco Getúlio Vargas, 1130, Bloco S
E-mail: lessandra@gmail.com

Rodrigo Schrage Lins
Master of Clinical Research in Infectious Diseases
Institution: Universidade de Caxias do Sul (UCS)
Address: Rua Francisco Getúlio Vargas, 1130, Bloco S
E-mail: rodrigolinsmd@gmail.com

Deise Renata Bringmann
Master in Environmental Sciences and Engineering
Institution:Universidade de Caxias do Sul (UCS)
Address: Rua Francisco Getúlio Vargas, 1130, Bloco S
E-mail: deise.renata@ucs.br

Leandro Luis Corso
PhD in Mechanical Engineering
Institution:Universidade de Caxias do Sul (UCS)
Address: Rua Francisco Getúlio Vargas, 1130, Bloco G
E-mail: leandro.lcorso@gmail.com

ABSTRACT
The objective is to use Artificial Intelligence (AI) for identifying which surgical patients have a likelihood ratio of developing an infection. We included in the study all the patients who underwent surgeries with wound class considered clean at a regional public hospital in Brazil. The first step was a retrospective analysis of risk factors and a correlation test for identifying which clinical variables are best related to post-discharge infections. Then, we developed an Artificial Neural Network (ANN) for pattern recognition to detect incidence of infections. The ANN can make accurate predictions in 77.3% of the cases in which an infection will occur, and

the AUROC of the model is 0.9050. Thus, it is possible to take actions before the patients develop it, improving the quality of life and mental health as well as avoid increasing costs.

Keywords: surgical site infections, infection prevention, post-operative care, biomedical engineering, statistics.

RESUMO
O objetivo é usar a Inteligência Artificial (IA) para identificar quais pacientes cirúrgicos têm uma razão de probabilidade de desenvolver uma infecção. Incluímos no estudo todos os pacientes que se submeteram a cirurgias com classe de ferida considerada limpa em um hospital público regional no Brasil. O primeiro passo foi uma análise retrospectiva dos fatores de risco e um teste de correlação para identificar quais variáveis clínicas estão mais bem relacionadas às infecções pós-descarga. Em seguida, desenvolvemos uma Rede Neural Artificial (ANN) para o reconhecimento de padrões para detectar a incidência de infecções. A ANN pode fazer previsões precisas em 77,3% dos casos em que uma infecção ocorrerá, e a AUROC do modelo é de 0,9050. Assim, é possível tomar medidas antes que os pacientes a desenvolvam, melhorando a qualidade de vida e a saúde mental, bem como evitar o aumento dos custos.

Palavras-chave: infecções do local cirúrgico, prevenção de infecções, cuidados pós-operatórios, engenharia biomédica, estatística.

1 INTRODUCTION

Over time, medicine has shown progress in many areas, including technology and the use of mathematics to solve healthcare issues (TUNC, ALAGOZ, e BURNSIDE, 2014) The provision of health services is considered an extremely challenging field because of its dynamic essence, as well as the inherent complexity and uncertainty. Multidisciplinary teams play an important role in making it possible to understand and organize large amounts of data in order to help with decision making (ARINGHIERI TÀNFANI e TESTI, 2013; SILVA FILHO;DE OLIVEIRA, 2022).

There are several examples where the union of medicine and mathematics brought countless benefits to patients (RESINO et al., 2011; XIE et al., 2014; SUN et al., 2015; CORSO et al.,2016; ESTEVA et al., 2017; ROFFMAN et al., 2018). One tool that has been widely used is Artificial Intelligence (AI), which attempts to simulate the behavior of the nerve cells in the biological central nervous system via computational networks (HAYKIN 2004; GRAUPE, 2013;).

The technological advancement has modified not only the patient care needs, but also the diagnostic methods of Healthcare-associated Infections (HAIs). Therefore, the identification and prevention of HAIs is necessary before the patients are infected. In Brazil, the standards, criteria and methods to control infections are established by the National Agency of Sanitary Surveillance (ANVISA)

The World Health Organization (WHO) has launched initiatives to improve safety issues in surgeries. One of them is the development of the Safe Surgery List that can be used in any operating room with the goal of improving team communication, avoiding inappropriate anesthetic practices, reducing the number of incidents and complications, reducing the surgical mortality rate, as well as avoiding surgical infection (WHO,2023) . However, while health institutions are pursuing several infection prevention and control initiatives, there is a need for continuous improvement globally (ABBAS e PITTET, 2016; TROUGHTON et al., 2018).

According to the National Health Surveillance Agency of Brazil, “the surgical site infections (SSI) are infections related to surgical procedures, with or without implant placement, in hospitalized or outpatient care patients” (ANVISA). That kind of infection can extend the length of hospital stays, and incur significant cost increases (PERENCEVICH et al., 2003; DE LISSOVOY et al., 2009; PEREIRA et al., 2022). In addition, infections detected from post-discharge patients also present innumerous consequences. Aside from the financial costs, there is a significant decrease in mental health component of health surveys (PERENCEVICH et al., 2003). Thus, the infections represent not only high financial costs, but also negatively impact on patient quality of life. The development of methods to prevent the occurrence of SSI is necessary to avoid such consequences (PERENCEVICH et al., 2003; BADIA et al., 2017).

This study aims to verify the correlation between risk factors to the incidence of infection, and identify a pattern of post-surgical infections on clean surgeries conducted at the University Hospital of Caxias do Sul (UHCS). This way, it is possible to take action before the patients develop post-surgical infections.

2 METHODS

The research was divided into five main activities in order to achieve the objective established previously. Figure 1 presents those crucial stages for the development of the study. The first step was the creation of a database containing the medical records of patients who underwent clean surgeries, from November 2017 until March 2018, at the University Hospital of Caxias do Sul (UHCS). The UHCS is a university hospital located in southern Brazil that assists patients of the country’s public healthcare system.

The Service of Hospital Infections Control (SHIC) imposes a set of actions in order to avoid and identify infections. There are two checklists, which are adaptations of the Surgical Checklist proposed by WHO, used for clean surgeries. Table 1 shows which variables were included in the study, as well as where that data was gathered. The changeable risk factors, found at the medical records, are pertaining to the characteristics of the preoperative patients. Those attributes were analyzed retrospectively to design the profile of the population that was included in the study.

After that, we made correlation tests using both the coefficients of Pearson and Spearman in order to analyze the relation of risk factors and the occurrence of infection. In a correlation analysis, the relation between the variables is verified in order to know how the alteration of one of them can influence on the result of another (RON e BETSY,2010; MATTOS, KONRATH e AZAMBUJA, 2017). The last step of the research was the development of an Artificial Neural Network (ANN) that has induction supervised learning for pattern recognition and feedforward topology (RICH, KNIGHT E NAIR 2009). The data was divided in three sections: 70% for training. 15% for validation and 15% for tests of the model. The idea is that, after the system learns the patterns, we can input the risk factors of a patient and have the chances of SSI as outcome.

On Figure 2, it can be seen a model of how an ANN works. The variables x_n at the input layer are the risk factors, cited on Table 1, and they were the data given to the ANN to learn the patterns of infections. In the hidden layers, a weight w is attributed to the relation of each risk factor to the incidence of infection. Thus, h_n represents a function of the inputs (clinical data) and associated weights in each passing neuron. There is also an error dependent function, the backpropagation algorithm (RICH, KNIGHT E NAIR 2009), that adjust the weights and minimizes any divergence.

On the output layer, the system shows if there is high or low probability of developing an infection. The evaluation of the results was done using parameters to measure the performance of the application: the confusion matrix and the area under the curve ROC, also known as AUROC (DEPERLİĞLU, 2018; WICHARD, CAMMANN e STEPHAN; TOLXDORFL, 2008; KUNHIMANGALAM , OVALLATH e JOSEPH, 2012). Figure 3 shows the equations used to calculate each one of the parameters of the confusion matrix.

Table 1. Risk factors from patients who underwent clean surgeries at UHCS

Electronic medical records	Surgical Center Checklist	Post-discharge follow-up Checklist
Age	Type of surgery	Incidence of infection
Immunosuppression	Administration of the surgical prophylaxis	
Obesity	Surgery duration > 3 hours or with CEC	
Alcoholism	Prophylaxis during a procedure	
Smoking	UHCS sterilized surgical instruments	
Cardiopathy	Surgical technologist	
Pneumopathy	Surgical attire	

ISSN: 2595-6825

Nephropathy	Hands antisepsis (surgeon, technologist)
Hepatopathy	Trichotomy location
Neuropathy	Intubation
Diabetes Mellitus	Use of O_2
	Glycemic control
	Operating room doors closed during surgery
	Skin washing and antisepsis
	Standard operating protocol
	Exchange of gloves in the procedure
	Verification of sterilization records

Source: authors (2023)

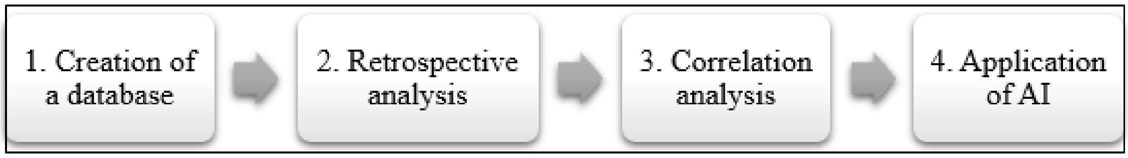
Table 2. Baseline characteristics

Variable	No SSI (n=325) No. (%)*	SSI (n=24) No. (%)*
Age (IQR)	54 (35-73)	51 (28-74)
Obesity	8 (2.46)	2 (8.33)
Imunossupression	1 (0.31)	0 (0.00)
Alcoholism	9 (2.77)	1 (4.17)
Smoking	50 (15.38)	4 (16.67)
Cardiopathy	124 (38.15)	14 (58.33)
Pneumopathy	18 (5.54)	2 (8.33)
Nefropathy	2 (0.62)	1 (4.17)
Neuropathy	12 (3.69)	2 (8.33)
Diabete Mellitus	56 (17.23)	6 (25.00)

Source: authors (2023)

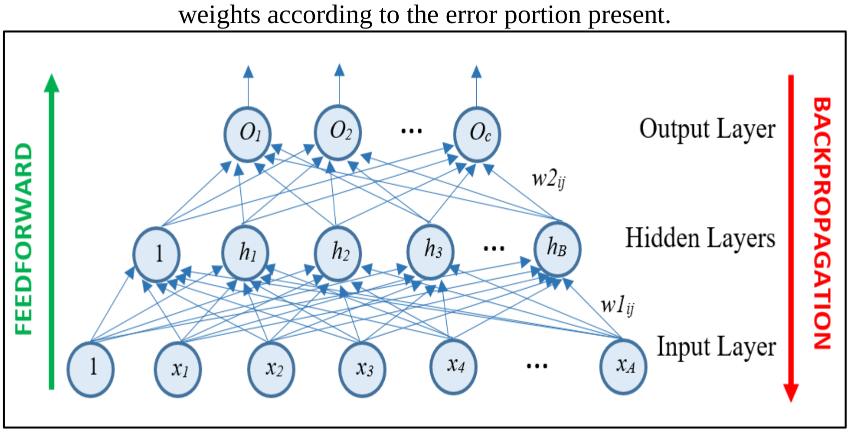
Table footnotes: *Data are presented as no. (%) or median (IQR).

Figure 1. Research Development



Source: authors (2023)

Figure 2. Artificial Neural Network model - Feedforward: the input data is propagated forward the network, and weights are given to evaluate the relation of the variables to the incidence of infection. Backpropagation: the error signals are propagated back through each of the hidden layers, and a computation is made to adjust the weights according to the error portion present.



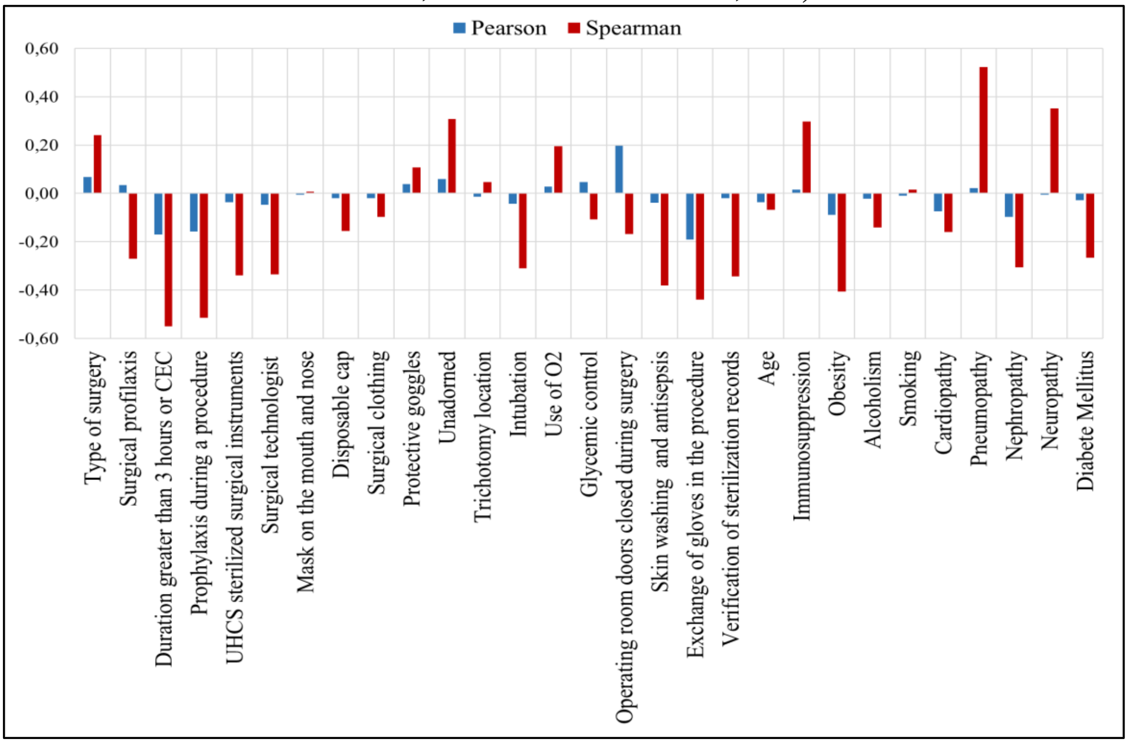
Source: Adapted from (LUGER,2009; RICH, KNIGHT e NAIR 2009)

Figure 3. Confusion matrix.

		TARGET CLASS		
		From rule: 0	From rule: 1	
OUTPUT CLASS	From data: 0	TRUE POSITIVES (TP)	FALSE POSITIVES (FP)	Positive predictive value $\frac{TP}{TP+FP}$
	From data: 1	FALSE NEGATIVES (FN)	TRUE NEGATIVES (TN)	Negative predictive value $\frac{TN}{TN+FN}$
		Sensitivity $\frac{TP}{TP+FN}$	Specificity $\frac{TN}{TN+FP}$	Accuracy $\frac{TP+TN}{TP+TN+FP+FN}$

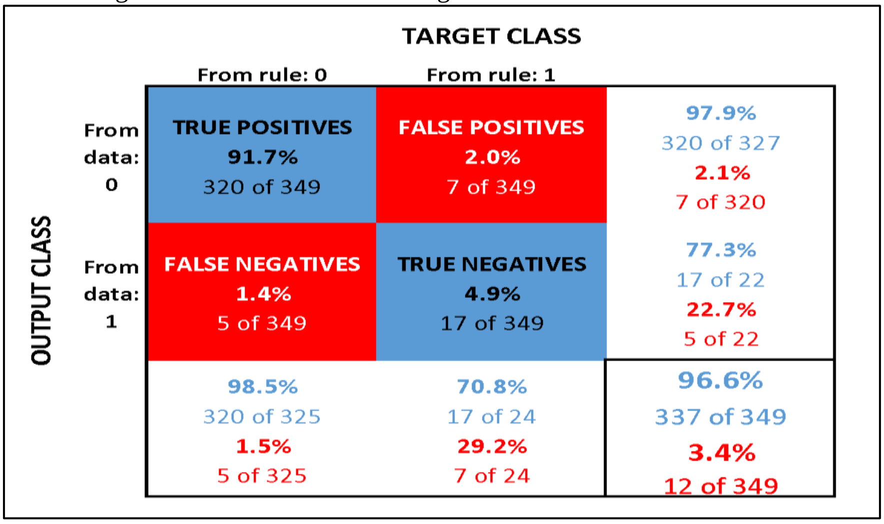
Source: authors (2023)

Figure 4. Pearson and Spearman correlation - The coefficient amplitude can range from -1 to 1. The values close to 1 represent a strong positive correlation, and those close to -1 characterize a strong negative correlation. In those cases where there is no correlation or it is very weak, the value is close to zero (RON e BETSY,2010; MATTOS, KONRATH e AZAMBUJA, 2017).



Source: authors (2023)

Figure 5. Results of the final diagnostic model: Confusion matrix



Source: authors (2023)

3 RESULTS

The database contains ninety-five kinds of procedures, and it was possible to verify that infections had happened in just nineteen of them. Out of 349 surgeries, 24 presented post-operative infections; it represents 6.89% of all the procedures in the database. On Table 2, it is shown the profile of the patients included in this research, regarding some changeable risk factors.

In the correlation tests using both coefficients of Pearson and Spearman, it was possible to notice that the results differ in the use of one coefficient and another. This may be justified by the fact that the Spearman coefficient is used for both linear and nonlinear data. Figure 4 shows a comparison of the results of the correlation of each item with the incidence of infection. The blue bars present the results obtained using Pearson's coefficient, and the red bars using Spearman.

Using Pearson, the linear correlation ranges from -0.19 to 0.20, while for Spearman the rank correlation is in the range of -0.55 to 0.52. In both cases, there is a weak correlation between risk factors and incidence of infection. This demonstrates the complexity of predicting infection after clean surgeries based directly on correlation data. Thus, AI presents itself as a tool to assist in identifying the relationship between information.

After analyzing the comportment of the data, AI was applied. The architecture of the network was reshaped many times aiming to get the best results, in other words, looking for a network that could achieve at least 95% accuracy identifying both cases when had or did not have postsurgical infections. During the test stage, the number of hidden layers of the network was changed until it was possible to find the best range. It was possible to notice an increase in the accuracy of the results when the network had between 35 and 45 hidden layers. The ANN containing 42 was the one that had the highest accuracy; it recognized when the patients developed infection or not.

Figure 5 depicts the confusion matrix generated using the ANN described earlier and the database used for this research. The sensitivity and specificity of the codes for identifying infections after clean surgeries were 98.5% and 70.8%, respectively. The PPV was 97.9% and the PNV was 77.3%. It is also informed both the number of observations classified and the percentage that they represent out of the input data. This way, it is possible to state that the system classified correctly:

a) 320 of the cases that did not have infection, representing 91.7% of the input data;

b) 17 of the cases that had infection, representing 4.9% of the input data.

Among the incorrect classifications, it was obtained the following results:

c) 7 of the cases that did not have infection were considered like they had, representing 2.0% of the data;

d) 5 of the cases that had infection were considered like they did not have, representing 1.4% of the data;

In general, the ANN of this study had an accuracy of 96.9% and, consequently, an error rate of 3.4%. It recognizes 77.3% of the cases when there is incidence of infection. The AUROC of the model was 0.9050, which corroborates the high performance of the ANN model.

4 DISCUSSION

There are several examples when the use of coding for surveillance of infections and different cultures showed to be helpful tools for the improvement of the efficiency at hospitals. Previous studies has used algorithms for the surveillance of deep surgical site infections after colorectal surgery (MULDER et al., 2019) and for detecting infectious disease high-risk patients (HAYKIN 2004)

For the development of this study, it was essential to understand which information is collected and held at the hospital as well as the method and time of collection. In addition, interpreting the data was essential to figure out what was actually relevant to include in the database. For example, some data was deemed irrelevant to the input matrix used to apply AI because they are collected just one month after the surgical procedure.

The number of surgeries has increased every year, and countries with resource-limited settings present higher rates of infections compared to more developed nations. This is a very challenging area, where it is necessary to analyze a wide range of parameters and take multidisciplinary actions (ABBAS e PITTET, 2016). Through the union of expertise from medicine and mathematics, it is possible to develop cutting-edge technologies that are available to all citizens.

While the literature presents some intrinsic and modifiable risk factors, it can be determined at the correlation analysis that the risk factors are not correlated. Thus, there is a difficulty in identifying what are possible combinations of clinical variables that can influence on the development of a post-surgical infection. The ANNs are an effective option to help on this process because they consider many different parameters at the same time.

Overall, the application of AI in order to identify patterns of post-surgical infections had great performance, achieving 96.6% accuracy and AUROC 0.9050. The present study shows an opportunity to identify more than seventy percent of the patients who are likely develop an infection after a clean surgery

As a result of the practical application of the developed network, it may now be possible to take actions to avoid the occurrence of infections following clean surgeries. The medical team will be able to adjust post-surgical procedures for those patients who the system flags as being at increased risk. This way, the UHCS can reduce expenses related to the patients who have to return to the hospital due to infections. It will result not only in the decrease of tangible costs, but mainly on the intangible costs which are associated to the health and psychological well-being of patients.

Once the network learns how to map inputs and outputs, it can be trained to adapt the synapse weights. Therefore, as more data is inserted, the network can learn to become even more intelligent as time goes by. The most significant advantage of the applied methodology is that the use of the checklists proposed by WHO facilitates the process of replication in other hospitals, considering that those checklists are known worldwide. Thus, it is possible to apply the same tools and algorithms used in this research.

5 CONCLUSION

The ANN can efficiently tell which patients have a high probability of having infections after clean surgeries before they actually develop it. The main benefits for the UHCS, with the implementation of this system, are innovation and consequent decrease of the post clean surgical infection rate. An effective program of infections prevention must have an epidemiologic surveillance in continuous improvement of the methods that are used. This way, it is possible to ensure the safety of both patients and healthcare professionals.

REFERENCES

ABBAS, M.; PITTET, D. Surgical site infection prevention: a global priority. **Journal of Hospital Infection**, v. 93, n. 4, p. 319-322, 2016.

ANVISA - Agência Nacional de Vigilância Sanitária. Critérios diagnósticos de infecção relacionada à assistência à saúde: segurança do paciente e qualidade em serviços de saúde. Disponívelem:<https://www20.anvisa.gov.br/segurancadopaciente/index.php/publicacoes/item/criterios-diagnosticos-das-infeccoes-relacionadas-a-assistencia-a-saude>. Acesso em: 12 jan.23.

ARINGHIERI, R.; TÀNFANI, E.; TESTI, A. Operations research for health care delivery. **Computers & Operations Research**, v.40, n. 9, p. 2165-2250, 2013.

BADIA, J, M. et al. Impact of surgical site infection on healthcare costs and patient outcomes: a systematic review in six European countries. **Journal of Hospital Infection**, v. 96, n. 1, p. 1-15, 2017.

CORSO, L,L, et al. Optimization of a cemented femoral prosthesis considering bone remodeling. **Journal of Biomechanical Engineering**, v. 138, n. 1, p. 011002, 2016.

DE LISSOVOY, G, et al. Surgical site infection: incidence and impact on hospital utilization and treatment costs. **American journal of infection control**, v. 37, n. 5, p. 387-397, 2009.

DEPERLİĞLU, Ö. Classification of segmented heart sounds with Artificial Neural Networks. **International Journal of Applied Mathematics Electronics and Computers**, v. 6, n. 4, p. 39-44, 2018.

ESTEVA, A, et al. Dermatologist-level classification of skin cancer with deep neural networks. **Nature**, v. 542, n. 7639, p. 115-118, 2017.

GRAUPE, D. **Principles of artificial neural networks**. World Scientific, 2013.

HAYKIN, S.; NETWORK, N. A comprehensive foundation. **Neural networks**, v. 2, n. 2004, p. 41, 2004.

KUNHIMANGALAM, R.; OVALLATH, S.; JOSEPH, P, K. Artificial Neural Networks in the identification of peripheral nerve disorders. **Journal of Mechanics in Medicine and Biology**, v.12, n. 4, p.1-14, 2012.

LUGER, G,F. Artificial intelligence: structures and strategies for complex problem solving. 6. ed. Boston: **Pearson Education**, 2009.

MATTOS, V, L,D, de.; KONRATH, A,C.; AZAMBUJA, A,M,V. Introdução à Estatística com aplicações em Ciências Exatas. **Rio de Janeiro: LTC**, v. 1, 2017.

MULDER, T. et al. A diagnostic algorithm for the surveillance of deep surgical site infections after colorectal surgery. **Infection Control & Hospital Epidemiology**, v. 40, n. 5, p. 574-578, 2019.

PEREIRA, C, F, C, et al. Patient safety: vulnerabilities associated with the management of extra-hospital transfer in the framework of SUS. **Brazilian Journal of Health Review**, v. 5, n. 4, p. 17156-17170, 2022.

PERENCEVICH, E, N, et al. Health and economic impact of surgical site infections diagnosed after hospital discharge. **Emerging infectious diseases**, v. 9, n. 2, p. 196, 2003.

RESINO, S, et al. An artificial neural network improves the non-invasive diagnosis of significant fibrosis in hiv/hcv coinfected patients. **Journal of Infection**, v. 62, n. 1, p. 77-86, 2011.

ROFFMAN, D, et al. Predicting non-melanoma skin cancer via a multi-parameterized artificial neural network. **Scientific reports**, v. 8, n. 1, p. 1-7, 2018.

RON, L.; BETSY, F. **Estatística Aplicada**. 4 ed. São Paulo: Pearson Prentice Hall, 2010.

SILVA FILHO, W, S; DE OLIVEIRA, R, M, Sousa. The importance of artificial intelligence in radiology: a systematic literature review Importance of artificial intelligence in radiology: a systematic literature. **Brazilian Journal of Health Review** , v. 5, no. 4, p. 12638-12649, 2022.

SUN, G, et al. An infectious disease/fever screening radar system which stratifies higher-risk patients within ten seconds using a neural network and the fuzzy grouping method. **Journal of Infection**, v. 70, n. 3, p. 230-236, 2015.

TROUGHTON, R. et al. Mapping national surveillance of surgical site infections in England: needs and priorities. **Journal of Hospital Infection**, v. 100, n. 4, p. 378-385, 2018.

TUNC, S.; ALAGOZ, O.; BURNSIDE, E. Opportunities for operations research in medical decision making. **IEEE Intelligent Systems**, v. 29, n. 3, p. 59, maio/jun.2014.

WHO - World Health Organization. Implementation manual surgical safety checklist: safe surgery saves lives. Disponível em: http://www.who.int/patientsafety/safesurgery/ss_checklist/en/ Acesso em: 12.jan.2023.

WICHARD J, D.; CAMMANN, H, STEPHAN C.; TOLXDORFL, T. Classification Models for Early Detection of Prostate Cancer. **Journal of Biomedicine and Biotechnology**, v.8, p.1-7, 2008.

XIE, BO-QIA, et al. Ventilation/perfusion scintigraphy in children with post-infectious bronchiolitis obliterans: a pilot study. **PLoS One**, v. 9, n. 5, p. e98381, 2014.