Developing a Clinical Decision Support System for Prediction Postoperative Coronary Artery Bypass Grafting Infection in Diabetic Patients

Marjan Ghazisaeedi1, Leila Shahmoradi1, Ali Garavand2, Masoumeh Maleki3*, Shahabeddin Abhari4, Marjan Ladan5, Sina Mehdizadeh6

ABSTRACT

Background: Postoperative infection in Coronary Artery Bypass Graft (CABG) is one of the most common complications for diabetic patients, due to an increase in the hospitalization and cost. To address these issues, it is necessary to apply some solutions.

Objective: The study aimed to develop a Clinical Decision Support System (CDSS) for predicting the CABG postoperative infection in diabetic patients.

Material and Methods: This developmental study is conducted on a private hospital in Tehran in 2016. From 1061 CABG surgery medical records, we selected 210 cases randomly. After data gathering, we used statistical tests for selecting related features. Then an Artificial Neural Network (ANN), which was a one-layer perceptron network model and a supervised training algorithm with gradient descent, was constructed using MATLAB software. The software was then developed and tested using the receiver operating characteristic (ROC) diagram and the confusion matrix.

Results: Based on the correlation analysis, from 28 variables in the data, 20 variables had a significant relationship with infection after CABG (P < 0.05). The results of the confusion matrix showed that the sensitivity of the system was 69%, and the specificity and the accuracy were 97% and 84%, respectively. The Receiver Operating Characteristic (ROC) diagram shows the appropriate performance of the CDSS.

Conclusion: The use of CDSS can play an important role in predicting infection after CABG in patients with diabetes. The designed software can be used as a supporting tool for physicians to predict infections caused by CABG in diabetic patients as a susceptible group. However, other factors affecting infection must also be considered for accurate prediction.

Keywords
Decision Support Systems; Clinical; Surgical Wound Infection; Coronary Artery Bypass; Diabetes

Introduction

Cardiovascular disease is one of the most important threats to human health [1], and Coronary Artery Disease (CAD) is one of the most dangerous types of heart diseases in developed countries [2]. Different treatment depends on amount of obstruction of coronary arteries. Those patients diagnosed with CAD may be treated by Coro-
Coronary Artery Bypass Graft (CABG), which is a surgical procedure for treating coronary heart disease to improve blood flow supply to the heart [3, 4].

Heart operation is one of the very important surgeries that need the patient to be hospitalized for about a week, even a couple of days in Coronary Care Unit (CCU) [5]. Postoperative infection is one of the most dangerous complications, which may increase hospitalization between 5 and 20 days and change dressings up to 2-3 times per day, i.e. the hospital cost will increase [6, 7]. Surgical wound infections may be caused by bacteria during operation or post operation, and it may also put the patients into serious issues like difficulty in rehabilitation process, sepsis, organ injury or even death [8].

Diabetes is one of the main common causes of postoperative infection in heart surgeries, and it is also a reason for coronary atherosclerosis, decreasing the chance of the treatment and increasing the side effect [9]. Identify the possibility of CABG in early-stage has a critical rule to control and decrease the chance of suffering from the possible disease in such a group of patients [10]. This group of patients need to be identified before the heart surgery operates to decrease any chance of post-surgery infection [11]. Clinical Decision Support System (CDSS) is a powerful tool to predict the possibility of postoperative infection, especially in CABG [12, 13].

CDSS helps to predict, identify, and diagnose the diseases in a short time especially in diseases, which its diagnose needs more time, cost and complicated tests [14, 15]. Decision making is a cognitive process, involves gathering information, evaluating situations, generating and selecting proposals, and implementing solutions. The decision support system is often used by decision-makers to reduce cognitive errors and maximize the efficiency of actions. A decision support system is a computer-based information system designed to support corporate and business decision-making activities [16]. The expansion of knowledge in the medical field and the complexity of diagnosis and treatment decisions have drawn the attention of experts to use decision support systems in various diseases, and therapies [17-19]. In recent years, the intelligence systems, such as CDSS, have been used to assist in the delivery of health care for various purposes in cardiovascular surgery, as well as the complications of diabetes [20]. The use of machine learning algorithms is very sensitive and specific to the diagnosis of coronary artery disease [21, 22]. Pimentel et al pointed out the important and effective role of machine learning algorithms in the early detection of coronary heart disease in patients with diabetes [23].

Gatti et al examined some statistical models to predict the CABG infection, and conducted the statistical models that can’t predict postoperative infection after CABG completely [24]. Due to high priority of CDSS and high possibility risk of infection after CABG especially in diabetic patients, this study aimed at developing a CDSS to predict postoperative infection after CABG diabetic patients.

**Material and Methods**

In this developmental study, the study population included CABG medical records in patients with diabetes from 2009 to 2013. From 1061 records, 210 were selected through simple random sampling. Data gathering was done using a data extraction form, divided into three main sections: preoperative, intraoperative, and postoperative. It should be noted that all stages of the study were carried out without regard to the names of the individuals.

The research variables are in three main categories:

A) Preoperative research variables:

- Gender, type 2 diabetes, over-weight, age, Chronic Obstructive Pulmonary Disease (COPD), history of infection in other parts of the body, wounds elsewhere in the body, positive nose culture staphylococcus carrier, history of diseases, causing immune deficiency,
Developing a CDSS for Prediction

chronic renal failure, and dialysis, duration of corticosteroid medication (like a patient that becomes chemotherapy), test result, history of heart disease, being a smoker, blood pressure, blood sugar, and Hemoglobin A1c (HBA1C) test result.

B) Intraoperative research variables:
Operation time, pumping time, number of grafts, and the use of mammary artery.

C) Postoperative research variables:
The number and the type of antibiotics after surgery, post-operative events (including respiratory failure), intubation time, length of stay in ICU.

To exclude ineffective variables on infection from CABG in the data extraction form, an initial analysis was performed using the One-Way-ANOVA test.

Given the important role of Artificial Neural Network (ANN) and different therapeutic applications than other artificial intelligence algorithms [25, 26], and its ability to learn easier, in this study, ANN was used. The ANN can detect the hidden aspects of science and hidden relationships between actual inputs and outputs [27].

For designing and implementing a neural network, the first step is to choose the right neural network for this task. To implement this network, the functions, available in MATLAB software, were used. The patients are supposed to be divided into infections and non-infections categories according to the data using an ANN.

One of the most suitable networks to solve the categorization problem is the perceptron neural network. In this study, a perceptron network with one hidden layer was used. In this network, there is a hidden layer and output layer, as shown in Figure 1.

A gradient descent algorithm, used in this research, guarantees that error function does not go through the incremental process of learning, and it also generally results in faster convergence by comparing this algorithm with batch gradient descent techniques and conjugating gradient algorithms.

The designed network consists of a layer with 10 nodes and a gradient descent algorithm for training. Because the system can be used for those who do not have specialized programming knowledge, the appropriate interface was defined, designed and implemented based on selective neural network and topological structure by the user interface based on the MATLAB graphics module. Eventually, the trained network was saved and used in Graphical User Interface (GUI) designed for the program. To test the CDSS, system performance analysis was performed using the confusion matrix (sensitivity, specificity, and accuracy) (Equations 1 to 3) and the ROC chart.

Figure 1: The neural network model used in the study
1) Sensitivity = \( \frac{TP}{TP + FN} \)

2) Specificity = \( \frac{TN}{FP + TN} \)

3) Accuracy = \( \frac{TP + TN}{TP + TN + FP + FN} \)

**Results**

From 1061 records of CABG with diabetes, 210 cases were selected. From the 210 cases reviewed, 94 cases were women and the rest were men, and the mean age of patients was 61.07.

Table 1 shows the distribution of postoperative infection status by gender.

For developing the CDSS, we used most important variables which are effective in postoperative CABG infection by the statistical tests include over-weight, age, COPD, positive nose culture (staphylococcus carrier), test results, history of hospitalization in intensive care units, history of surgery, history of cardiac surgery, history of heart disease, blood sugar, operation time (duration of operation), pump time and length of stay in the ICU (p<0.05).

To design software based on the network analysis, the variables were used as study variables.

Descending gradient methods were widely used to train this ANN. After executing the program and completing the training and testing process, the results were obtained as seen in Table 2, comparing two different values of the number of hidden layers for 500 repetitions, and its effect on the accuracy of the algorithm. The results also showed that increasing the repetition caused the model overfit.

After creating and approving the final model, the interface was designed for CDSS to predict postoperative infection CABG in diabetic patients.

Table 3 shows the confusion matrix of the model, calculated by applying the test data set. After designing, the system evaluation was done, based on the results of software testing using Equations 1, 2 and 3, and the sensitivity of software, specificity and detection accuracy were calculated 69%, 97%, 84%, respectively. Figure 2 shows the ROC (Receiver Operating Characteristic) diagram.

<table>
<thead>
<tr>
<th>Number of hidden layers neurons</th>
<th>Percentage of training accuracy</th>
<th>Percentage accuracy of test phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>99</td>
<td>92.9</td>
</tr>
<tr>
<td>8</td>
<td>97.5</td>
<td>91.3</td>
</tr>
</tbody>
</table>

**Table 1: post Coronary Artery Bypass Graft (CABG) infection status by gender and diabetes status**

<table>
<thead>
<tr>
<th></th>
<th>Infected and diabetic</th>
<th>Infected and non-diabetic</th>
<th>All infections</th>
<th>Disinfected and diabetic</th>
<th>Disinfected and non-diabetic</th>
<th>All non-infections</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>percent</td>
<td>percent</td>
<td>percent</td>
<td>percent</td>
<td>percent</td>
<td>percent</td>
<td>percent</td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>94</td>
<td>21.78</td>
<td>44</td>
<td>17.69</td>
<td>23</td>
<td>29.17</td>
<td>210</td>
</tr>
<tr>
<td>Men</td>
<td>116</td>
<td>78.22</td>
<td>158</td>
<td>82.31</td>
<td>107</td>
<td>70.83</td>
<td>202</td>
</tr>
<tr>
<td>Total</td>
<td>210</td>
<td>202</td>
<td>130</td>
<td>72</td>
<td>8</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

**Table 2: Comparison of the effect of different amounts of hidden layer neurons on program accuracy for a maximum of 500 iterations.**
Developing a CDSS for Prediction

The paper functioning of the diagnostic system for patients with postoperative CABG infection has demonstrated curved behaviour and given the appropriate line spacing of the graph to the central line.

Figure 3 showed the user interface (UI) of developed CDSS.

Discussion

CABG infection can be predicted with good accuracy using the accurate data in CDSS. In the present study, variables affecting postoperative CABG infection in diabetic patients were identified based on the previous studies [28, 29].

The use of artificial intelligence tools has an important role in predicting the complications of diabetes and can be also used as a powerful supporting tool [30]. Due to the high importance of using ANN algorithms in classifying data, we used it to construct the CDSS for predicting infection after CABG in diabetic patients.

In a study, Esmaeily et al, compared the performance of data mining algorithms to identify risk factors for type 2 diabetes and found that the ANN algorithms perform better than Support Vector Machine (SVM) and Multiple Logistic Regression (MLR) algorithms (Specificity=81.2%, sensitivity=63%, and accuracy=78.7%) [31]. Also, Adawi et al, compared the application of ANN and bivariate logistic regression in their studies in predicting the diagnosis of patients with hypertension and diabetes and selected the ANN algorithm as the most appropriate algorithm [32]. Fisher et al, used the neural network in a study of a web-based system to detect and differentiate strabismus species in Liverpool in 2009 [33]. This system is web-based and the physician can easily enter information after a patient examination. In this study, a multilayer perceptron algorithm was used to train the data [33], which was similar to the neural network structure used in the present study.

The problem in this study (predicting of postoperative infection in CABG in patients with diabetes) is resolved with an ANN, developed by a Multi-Layer Perceptron (MLP) because it is easier to implement and train against other methods. The results of software testing were calculated for software sensitivity 69%, specificity 97%, and detection accuracy of 84%.

A study was performed by Ghaderzadeh et al, to design an ANN-based CDSS for early detection of prostate cancer from benign enlargement of the prostate that system performance analysis was performed by the confocal matrix and the ROC graph. In this study, the sensitivity of the decision system is 92.11 and its specificity was 97.06 that has a high percentage, overall, for all cases, the system could accurately detect 94.44 all cases of can-

<table>
<thead>
<tr>
<th>Table 3: Confusion matrix by applying the test set data.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creating an infection</td>
</tr>
<tr>
<td>-----------------------</td>
</tr>
<tr>
<td>Creating an infection</td>
</tr>
<tr>
<td>No infection</td>
</tr>
<tr>
<td>Total columns</td>
</tr>
</tbody>
</table>

The selected network receiver operating characteristic (ROC) curve associated with the test data set.
Shahmoradi et al., used MLP as an ANN algorithm as it was the most efficient algorithm for predicting acute appendicitis with accuracy 92.90%, sensitivity=80%, and specificity=97.5% [34], which in terms of the obtained indices corresponds to the present results.

It should be noted that the reason for the sensitivity amount (80%) of software against other two indices can be due to the lack of the data on patients with infection; that can be helpful in identifying people who are infected. We used of high-level coding to developing the CDSS by MATLAB, which is a functional and multifunctional language with many capabilities.

This advantage makes it possible to integrate the designed system across different user environments and operating systems by the expert system interface program that the designer is limited and depends on its environment. The high performance of the system is also remarkable, considering the small number of records. In future studies, a better prediction pattern and ultimately a higher performance CSSD whit a higher volume of data will likely be obtained.

Considering that more information is needed to design this system, including the use of High-Efficiency Particulate Air (HEPA) filters and non-control of dressing set date, proper handwashing and disinfection, sterile operating room control, and sterile quality control of the operating room and medical equipment, it is recommended that these cases be recorded in the patient medical records. Many patients get information after surgery, but the problem is solved outpatient or even referred to other centres; thus, the number of patients with post-operative infection may not be accurate. It is recommended to follow up on the complications of high-risk surgery in hospitals for recording in the patient medical records.

**Conclusion**

The designed CDSS in this study can be used for all patients, who have undergone CABG surgery, in addition to diabetic patients to
make a timely decision on the right opportunity to deal with nosocomial infections. The use of a large statistical community can increase the reliability of this system. The increase in performance indicators, including the accuracy of diagnoses, it has been widely used and encouraged other physicians to use it.

The software designed for this study can be provided as a supporting tool for physicians to prevent CABG infections in the diabetic patient as a sensitive and susceptible group to infection. However, other factors mentioned in the system should be considered as they can increase the sensitivity of the software to make more accurate diagnoses. Given the limited information about patients undergoing CABG surgery, it is recommended that authorities make the necessary decisions to create a specialized registry for this surgical procedure to provide the basis for extensive research in this field.

Acknowledgment
This study is extracted from a MSc dissertation in Health Information Technology by support of the Tehran University of Medical Sciences with Number 3/280/f/67.

Conflict of Interest
None

References


