COVID-19 PATIENTS ANALYSIS AND RISK PREDICTION BASED ON LIFESTYLE DISEASES THROUGH INDIAN DATASET

Vijaykumar Patil
Research Scholar at Department of Computer Engineering,
Bharati Vidyapeeth College of Engineering
Navi Mumbai, Maharashtra, India

Dr. Dayanand Ingle
Professor and Head at Department of Computer Engineering,
Bharati Vidyapeeth College of Engineering
Navi Mumbai, Maharashtra, India

Abstract — A Novel Coronavirus disease (COVID-19) is a transferable virus triggered by a recently revealed coronavirus. World Health Organization (WHO) declared it as pandemic worldwide. COVID-19 was originated from Wuhan, a city of China and spared over the more than 190 countries over the world. The USA, Spain, Italy, France even India and every country suffered a lot by this epidemic. The indications of COVID-19 are Fever, Cough, Shortness of breath or trouble in breathing, Chills, Repeated shaking with chills, Muscle torment, Headache, Sore throat which is normal as any formal flu which each individual feel during season transition. In this article, the statistical analysis like chi-square analysis, age-wise and diseases-wise classification of recovered and deceased patients are performed and also the different types of Machine Learning models like Multiple linear regression, Naïve Bayes Classifier, and Multilayer Perceptron Classifier are proposed for formal analysis and risk prediction of patients with different age group and individuals having lifestyle-based diseases with COVID-19. The dataset used for this study downloaded from covid19india.org, available in .csv format which included travel history of patients, relation with any existing COVID-19 patient, and record of any lifestyle-based diseases like diabetes, hypertension, respiratory problem, etc.

Keywords — COVID-19, Machine Learning, Multilayer Perceptron, Naïve Bayes

I. INTRODUCTION

Now-a-days diabetes, hypertension and respiratory problem are the types of disease are become worldwide problem wherein it starts with any stage of life. These types of diseases are becoming day by day serious issue that among two, every second individual is suffering from it. Their always exists a situation where we find many uncertainties and differences in opinions in medical diagnosis process, so they exist a need to find ultimate solution to avoid risk factor in life [9, 10]. COVID illness 2019 (COVID-19), brought about by serious intense respiratory disorder COVID 2 (SARS-CoV-2), is a continuous pandemic. The figure of individuals tainted by the Coronavirus infection is expanding quickly. Up to May 12, 2020, 1,436,198 instances of COVID-19 have been accounted for in more than 200 nations and regions, bringing about around 85,521 passing's (with a lethal pace of 5.95%) [1]. This has coordinated to-wards the public health concern in the international local area, as the World Health Organization (WHO) recognized the episode to be a Public Health Emergency of International Concern (PHEIC) on 30th Jan, 2020, and remembered it as a pandemic on March 11, 2020 [2, 3]. COVID-19 has influenced everybody is day by day lives. Something like 316 million individuals from every country have been approached to remain at home to hinder the pandemic. In this perspective, organizations have been helpless to make considerable changes. Working environment tasks of numerous organizations went virtual. The impact of the advanced change on efficiency and corporate culture has been concentrated widely. In the interim, what COVID-19 has meant for purchasers, and the utilization culture has gotten generally restricted consideration.

Numerous representatives have rushed to change in accordance with the computerized change. Online video gathering software Zoom reports a 78% development in profits, and Google Meet reports a roughly 60% increment in client deals, where individuals go through 2 billion minutes in online gatherings each day [4-5]. Employees might not have any desire to return to actual offices even after the pandemic facilitates. Administrators foresee that a portion of the effects from COVID 19 on work environments may be enduring, where adaptable work hours and computerized gatherings are relied upon to progressively supplant the conventional
corporate culture [6]. Organizations are trying different things with decentralized dynamic and new software to make the new computerized work culture successful and as useful as working in actual offices [4]. Despite the fact that it is basic for organizations to make this change as successfully and quickly as could be expected, some different results that are basic to the reasonability of busi-nesses have often been forgotten about from the conversation – the effects of COVID 19 on purchasers and the market. Pandemic, even though it was viewed as an improbable occasion for quite a while before the COVID-19 episode, has been distinguished as one of the critical dangers to organizations, as indicated by a review led in 2007[12]. Virtual gatherings and task the board advancements have been progressively carried out even before the pandemic, and the computerized change has been moderately bothering free for some organizations [4]. The [7] pandemic as a gas pedal of the primary change in utilization and the computerized change in the commercial center. An audit coordinated in 1998 when e-straight introduced the better methodology for shopping on the web, tracked down that 46% of early adopters were using online business a significant part of the time, while simply 8% by and large adopters had an experience of web shopping [13, 14]. Nonetheless, the continuous examination on electronic shopping conduct exhibits that dis-semination may happen in light of framework impacts, yet the pandemic can moreover be a trigger. As shown by a continuous audit of 2,200 adults in the U.S., 37% of study respondents have considered moving to web shopping after COVID-19[15].

“Machine learning is getting popular in all industries with the main purpose of improving revenue and decreasing costs: by using Machine learning technique they automate and optimize their process to solve challenging tasks very efficiently [19,20]”. Due to the reputation of ML in all the variety of applications and knowledge extraction, this research aims to broadly debate the role of lifestyle of an individual, age and various disease analysis are helpful to fighting the COVID-19 as well as in such future pandemic, which will rouse future pragmatic applications and methodological exploration[21]. In the accompanying investigation, first download information from covid19india.org for COVID-19 examination, and afterward pre-measure it as indicated by singular patient’s outline measurement then, at that point sum up famous AI strategies in the expectation and information exporting. Finally, paper discuss several results of such statistical and machine learning analysis which helps to predict risk analysis of patients based on its lifestyle related diseases. Note that this study is mostly related dataset of COVID-19 studies up to April 29, 2020.

The rest of the paper is organized or structured as, section II is the Literature Review which is the key works dealing with COVID 19 disease assessment, in Sect. III is designed as Proposed Methodologies used for the analysis of disease with co-morbidity. The dataset preparation and it attributes are illustrated in sect.-IV, the Result and discussion expounded in Sect. V, and definitive last section is conclusions are drawn in Sect. VI.

II. LITERATURE REVIEW

COVID-19, including picture securing, division, finding, and development. They especially center around the incorporation of X-beam and CT processed using AI, the 2 of which are broadly utilized in the forefront clinics, to delineate the most recent advancement of clinical imaging and radiology battling against COVID-19. Contrasted with the customary imaging work process that intensely depends on human works, AI empowers progressively sheltered, exact and effective imaging arrangements. The applications those late AI-engaged in COVID-19 principally incorporate the committed imaging stage, the contamination and lung and area division, the clinical appraisal and analysis, just as the spearheading fundamental and clinical research [8]. Additionally, numerous business items have been created, which effectively incorporate AI to battle COVID-19 and obviously exhibit the capacity of the innovation. N. Zheng et al., in [5] cross breed man-made consciousness builds an AI model is projected for COVID-19 prediction. To start with, as conventional pestilence models treat all people with coronavirus as experiencing a similar disease rate, an improved powerless contaminated model is intended to appraise the assortment of the disease rates for the diffusion laws breaking down and advancement pattern. Next, bearing in mind the impacts of control measures, avoidance and the expansion of the open's anticipation mindfulness, the new branch of NLP segment and the extensive memory is used to organize are implanted into the IISI model to assemble the half breed. The impact of government control gauges to be considered, the media's forthright statements, and the expansion in open mindfulness involving plague avoidance, this article utilizes pretrained NLP models to remove highlights from applicable updates on different regions and urban areas. The LSTM is designed with extricated highlights are in this way joined with the system to address the variation of this disease rate evaluated by the IISI model, majorly which
could predict the quantity of contaminated cases dependent
different advancement pattern and the associating laws.
S. Park et al., in [10] proposed a motivated tactic by which the
factual assessment of the CXR radiographs which having
potential evidence for prediction. Exploratory outcomes of
these methods used clinically understandable strikingness
maps which show that this strategy achieves cutting edge
implementation which is beneficial for COVID-19 decision. In
this paper authors mean to additionally explore profound cnn
and assess its plausibility for COVID-19 analysis. Tragically,
under the present general wellbeing crisis, it is hard to gather
huge arrangement of well-curated information for preparing
neural network-based system. Hence, one of the fundamental
focal points of this paper is to build up a neural system
engineering that is reasonable for preparing with constrained
preparing informational collection, which can at present
produce radiologically interpretable outcomes. Since most as
often as possible watched watched circulation examples of COVID-19
in CXR are respective inclusion, fringe dissemination and
ground-glass opacification (GGO), an appropriately planned
neural system ought to reflect such radiological discoveries.
Besides, by broadening the possibility of the angle weighted
class enactment map (Grad-CAM), which is designed based
on novel probabilistic theory which is one more significant
commitment of this paper that considers of fix insightful
ailment likelihood in producing worldwide strikingness map.
The subsequent class initiation map unmistakably shows the
easily understandable outcomes that are all around connected
with NN discoveries.
J. Zhang et al., in [11] proposed the Leverage advances, for
example, built polymerases, isothermal intensification, and
direct enhancement from complex grids may improve the
profitability of current foundation, while rising advances like
CRISPR diagnostics, pictorial end point recognition, and
ePCR strategies for corrosive but nucleic detecting that may
take as prompt o the at-home tests. The exercises scholarly,
and advancements prodded from the COVID-19 pandemic
could update our worldwide general wellbeing framework to
more likely battle possible flare-ups later. The capacity to
perform indicative examines inside each family unit can
possibly upset in this current situation administration. They
additionally checked on advances that can possibly bring at-
home sub-atomic tests to the real world. Bottlenecks in every
innovation are recognized to empower future designing
endeavors.
Theo Wibisono et al., in [12] introduced models to a great
extent depend on homogeneous blending models, which treat
everyone from the populace as having indistinguishable
contamination hazard. Instinctively, such a supposition that is
unreasonable. Certain segment gatherings (e.g., social
insurance laborers, kids and the older), have higher disease
dangers. Furthermore, personal conduct standards, for
example, utilization of open transportation sway disease
dangers. Utilizing contact systems to speak to the degree of
contact between populace individuals and statistics
information to rough geographic area and travel designs, they
reproduce the movement of a bead spread sickness from the
different areas under the Greter Toront. The outcomes are
occasionally shown on maps of territory utilizing GIS for
representation and arranging purposes.

III. MATERIAL AND METHODOLOGY

The proposed methodology shown in Fig 1, which divided in
three stages, first is to collecting dataset and perform pre-
processing where only required patient’s data is kept and other
data (or attributes or tuples having in complete entries) which
are not considered in proposed research are discarded. In
proposed methodology patient referred as P and relative of
referred as R. After finalization of data divided in three cases
1. Patients having travel history (Considered International/Domestic) including relative also, 2. Age wise
categorization of data and 3. The patients having pre-disease
history like diabetes, heart and hypertension etc. Second
statistical analysis performed on prepared dataset and then
finally employed the Machine Learning technique to predict
risk of COVID – 19. In proposed methodology the Chi square
analysis, Other Statistical Measures (like age and disease wise
classification), Multiple Regression Analysis, Multilayer
Perceptron Analysis and Naive Bayes classification are
applied to verify different factors which contribute in risk
prediction of COVID-19 patients.
**Multiple-linear regression (MLR):** It also identified purely as statistical technique; a multiple regression forecast the outcome using numerous clarifying variables to of a response variable. The MLR designed by setting a only goal that is to model the association between the descriptive (independent) variables and reply to response (dependent) variable. The Formula for Multiple Linear Regression is

\[ y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_p x_{ip} + \epsilon \]

where, for \( i = n \) observations:
- \( y_i \) = dependent variable
- \( x_{i1}, x_{i2}, ..., x_{ip} \) = explanatory variables
- \( \beta_0 \) = y-intercept (constant term)
- \( \beta_1, \beta_2, ..., \beta_p \) = slope coefficients for each explanatory variable
- \( \epsilon \) = the model’s error term (also known as the residuals)

The variables of dataset considered as follows to fit Multiple Linear Regression model

- \( y_i \) = dependent variable: risk ratio
- \( x_{i1} \) = Travel history
- \( x_{i2} \) = Relative travel history
- \( x_{i3} \) = Relative found COVID +ve
- \( x_{i4} \) = Diabetes
- \( x_{i5} \) = Heart disease
- \( x_{i6} \) = Brain issue
- \( x_{i7} \) = TB
- \( x_{i8} \) = Thyroid
- \( x_{i9} \) = Mental Disorder
- \( x_{i10} \) = Respiratory problem
- \( x_{i11} \) = Hypertension
- \( x_{i12} \) = Kidney
- \( x_{i13} \) = External injury
- \( \beta_0 \) = age of an individual
- \( \beta_1 \) = coefficient value of dependent variable when \( x_{i1} \) changes
- \( \beta_2 \) = coefficient value of dependent variable when \( x_{i2} \) changes
- \( \beta_{13} \) = coefficient value of dependent variable when \( x_{i13} \) changes
- \( \epsilon \) = the model error term

**Multilayer Perceptron Classification (MLP):** It is a model designed based on ANN model which apply feed-forward maps to all possible sets of input data attributes onto a set of fitting output. The MLP designed with nodes having directed graph and multiple layers, with each layer from MLP is completely attached to the next layer. Aside from the info node, every node is a neuron (or handling component) with a non-direct actuation work [16]. The multilayer perceptron classifier architecture shown in Fig. 2. The fit the network model i.e. MLP with back propagation. The during design of current network model a Softmax function is applied as an activation function and Cross-entropy as an error function.

**Backpropagation** (Age, Travel History, Pre-decease details, n, nin, nout, nhidden)

The i to j is \( x_{ij} \) considered as input and the i to unit j is unit denoted wij called as weight from input unit.

1. create a feed-forward network with \( n_{in}, n_{hidden}, n_{out} \) units
2. Initialize all network weights to small random numbers
3. Until the termination condition is met, Do
   - For each \( \langle x, t \rangle \) in training examples, Do
     - Propagate the input forward through the network:
       1. Input \( \langle x, t \rangle \) to the network and compute \( o_k \) of every unit \( u \)
       - Propagate the errors back through the network:
         2. For each network output unit \( k \), calculate its error term \( \delta_k \)
            \( \delta_k \leftarrow o_k (1 - o_k) (t_k - o_k) \)
         3. For each hidden unit \( h \), calculate its error term \( \delta_h \)
            \( \delta_h \leftarrow o_h (1 - o_h) \sum_{j \in \text{outputs}} w_{hk} \delta_k \)
         4. Update each weight \( w_{ij} \)
            \( w_{ij} \leftarrow w_{ij} + \Delta w_{ij} \) where \( \Delta w_{ij} = \eta \frac{\partial E}{\partial w_{ij}} \)
4. Minimize the errors through Back propagation algorithm [17].

The proposed method in his paper considered the subsequent closure circumstances in the MLP classifier:
• Fixed number of epochs or repetition.
• Threshold for error.
• Discrete validation set for error on a fall up to certain threshold.

**Naive Bayes classifier:** A NBC is a based on probabilistic approach which is popularly used for classification; it is implemented based on Bayes theorem shown as below.

\[
P(A | B) = \frac{P(B | A)P(A)}{P(B)}
\]

The hypothesis made here is that the predictors or features are shown in Table 1 all are impartial. The age, travel history and predeceases data of patients used to training the Naive Bayes classifier model and corresponding target variable ‘Patient_Status’ (either Deceased or Recovered). To classify risk analysis of patient whether Deceased or Recovered based on age, travel history and predeceases. The above theorem rewritten as

\[
P(D | R) = \frac{(R_1 | D) P(D)}{P(R)} + \frac{(R_2 | D) P(D)}{P(R)} + \ldots + \frac{(R_n | D) P(D)}{P(R)} \ldots (2)
\]

Were,
- \(P(D|R)\) - probability (posterior) of deceased patients
- \(P(D)\) - prior probability of deceased patients,
- \(P(R|D)\) - likelihood or probability of recovered patients or predictor
- \(P(R)\) – recovered predictors prior probability.

The measure the performance of different measured used evaluated using confusion matrix which is one of very well-known measured in ML. The terms used here are TP, TN, FP and FN labeled below.

- **TP** (True positive): number of patients which actually deceased predicted by model.
- **TN** (True Negative): number of false prediction of patients which actually deceased predicted by model.
- **FP** (False positive): number of patients which actually deceased but predicted as recovered.
- **FN** (False Negative): number of false patients which actually deceased but predicted as recovered.

**IV. DATASET PREPARATION**

The covid19.org provide different well-prepared dataset of COVID-19 patients from India, they also allow to use different forms of API to access it online and integrate in our application or for analysis. In this paper the file labeled as death_and_recovered.csv from covid-19 is used, it included attributes as follows

<table>
<thead>
<tr>
<th>Row Dataset Attributes</th>
<th>Attributes of Processed Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date of admission</td>
<td>Age Bracket</td>
</tr>
<tr>
<td>Age of Patient</td>
<td>Travel History of Patient</td>
</tr>
<tr>
<td>Gender</td>
<td>Domestic travel History</td>
</tr>
<tr>
<td>Current Status</td>
<td>Reason for travel</td>
</tr>
<tr>
<td>City of Patient</td>
<td>Is relative +ve</td>
</tr>
<tr>
<td>District</td>
<td>Is Patient Diabetes</td>
</tr>
<tr>
<td>State</td>
<td>Any Heart disease</td>
</tr>
<tr>
<td>State code</td>
<td>If Brain issues</td>
</tr>
<tr>
<td>Notes (descriptive data)</td>
<td>Is Patient having TB</td>
</tr>
<tr>
<td>Nationality of Patient</td>
<td>Is Patient having Thyroid</td>
</tr>
<tr>
<td>Is Patient having Respiratory problem</td>
<td>Is Mental disorder</td>
</tr>
<tr>
<td>Is Patient having Hypertension</td>
<td></td>
</tr>
</tbody>
</table>

The row dataset which is downloaded from covid-19.org contains “Notes” attribute. It contains data relevant to the covid patients in descriptive form which is about patients history which converted on categorical attributes as shown in Table 1. The processed dataset having 7408 rows in which only 433 patient’s status is records and having a historical medical data is described as Note. The Status is categorical variable or predictor which having two binary values as first recovered which denoted as 0 and second one is deceased represented as 0.

**V. RESULT AND DISCUSSION**

**A. Chi-square analysis**

The analysis result of Chi square for finalized dataset shown as follows:

- **Ha:** The patients age is less than 55, having greater percentage of Recovery
- **H0:** The patients age is less than 55, having less percentage of Recovery

<table>
<thead>
<tr>
<th>Age-wise Analysis</th>
<th>Age is &gt;=55</th>
<th>Age is &lt; 55</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deceased</td>
<td>211</td>
<td>109</td>
</tr>
<tr>
<td>Recovered</td>
<td>32</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>243</td>
<td>189</td>
</tr>
</tbody>
</table>
The Table 2 and Fig 3 shows Age-wise Analysis of COVID – 19 patients which are categories as Deceased and Recovered with age as threshold value 55. This table shows 243 patients out of 432 having age is greater or equal to 55 and 189 patients age having less than 55. The Table 3 and Table 4 shows p-value 9.68152E-13 and 0.034672162 respectively. The p-value from Table 3 is far less than significance level 0.5 so it clearly indicates the null hypothesis are rejected strongly.

**B. Other Statistical Measures**

The age wise risk analysis concluded in Fig 4; it shows the no of patients (Frequency) vs age bracket Histogram with mean age 56 and std. dev 16, but it clearly indicates that the patients having age is greater than 55 having high risk of COVID -19. Fig 5 shows Nonlinear Regression: Exponential Model, here $y$ is taking log of both sides shows relative intensity of age to the number of COVID-19 patients is $7.537e^{0.036x}$ and R-Square indicate that 75.45 % of variability of number of patients explained by variable age. Similarly, Fig 6 and Fig 7 shows age wise classification deceased and recovered patients, the relative intensity deceased patients is $2.203e^{0.05x}$ and R-Square indicate that 80.40 % of variability of number of patients those are deceased explained by variable age.
The Recovered Vs Deceased age wise risk analysis shown in Fig 8. It shows the patients average age between 28 to 30 having greater percentage of chance to recovery whereas age more than 56 may be the less. The Fig 9 shows the analysis of lifestyle-based diseases analysis of COVID-19 patients with different diseases like diabetics, heart, hypertension, kidney, respiratory problem, TB, thyroid, mental disorder etc. The individuals having diabetics, heart, hypertension and respiratory problem having high risk from COVID-19 diseases than all other types of disease. These included diseases in risk analysis in Figure 8 are lifestyle disease which arises with age.

Fig 8: Recovered Vs Deceased

The Recovered Vs Deceased age wise risk analysis shown in Fig 8. It shows the patients average age between 28 to 30 having greater percentage of chance to recovery whereas age more than 56 may be the less. The Fig 9 shows the analysis of lifestyle-based diseases analysis of COVID-19 patients with different diseases like diabetics, heart, hypertension, kidney, respiratory problem, TB, thyroid, mental disorder etc. The individuals having diabetics, heart, hypertension and respiratory problem having high risk from COVID-19 diseases than all other types of disease. These included diseases in risk analysis in Figure 8 are lifestyle disease which arises with age.

Fig 9: Disease wise Risk Classification of COVID-19 Patients

C. Machine Learning Techniques

In proposed methodology, Multiple Linear Regression, Naïve Bayes and Multilayer Perceptron (MLP) Classification are the types of Machine Learning techniques are used. Multiple Linear Regression and MLP are build using IBM SPSS Statistics tool. The Fig 10 shows Model Summary of Multiple Linear Regression; Here the Adjusted R Square values illustrate that the 42% variability of Patient_Status explained by Dependent variables (listed in Table 1). Fig 11 shows ANOVA test Table, it shows F-Value and P – Values as 14.329 and 0.00 respective which indicate strongly the null hypothesis ‘The patients age is less than 55, having less percentage of Recovery’ is rejected.

Fig 10: Summary of MLR

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.672a</td>
<td>.452</td>
<td>420</td>
<td>373</td>
</tr>
</tbody>
</table>

a. Dependent Variable: Patient_Status

Fig 11: ANOVA Table

The Model Summary of Multilayer Perceptron Classifier (Network) shown in Fig 12 and 13, which shows Cross Entropy Error and Percentage of Incorrect Predictions as 17.387 and 2.1% for training and 8.917 and 2.8% for testing respectively. The 73% of dataset used to train MLP and 27% records are to be used to train the model. The following Figure 12 shows table of Model Summary of MLP Classifier, during data preparation for MLP; the value 0 (zero) used to indicate patient status as Deceased and 1 (one) for Recovered patients. MLP result shows 97.2% of overall accuracy in the testing phase of model.

The machine learning classification model Naive Bayes build using python programming model where the 75% of dataset used to train and 25% are used test the model. The Table 4 shows Confusion Matrix generated through Naive Bayes classifier which shows statistics of actual vs predicted trade-off, where 32 observations are predicted as True – Positive and 29 are False – Negative and Type-I error indicate 5 observations and 3 are from Type- II error. The accuracy of Naive Bayes classification model given as 88.41% and other statistical measures are shown in Table 5.
VI. CONCLUSION

In examining the proposed research so far, it has been seen that the risk of COVID-19 patients is in more aged people, having age greater than 55 as well as the individuals those are already suffered from lifestyle-based diseases like diabetics, heart, hypertension, kidney, respiratory problem, mental disorder, etc. The proposed mechanism has been tested using statistical analysis (chi-square) and Machine learning algorithms like Linear and non-linear, Naive Bayes classification, and Multilayer Perceptron Classifier or Network. The statistical approach indicates that the variability of risk for COVID-19 disease explained by factors considered is in between 42% (min) to 80% (max). The prediction models also predict the risk with 88% (min) to 97% (max) accuracy level. The proposed study indicated the aged individual with already suffered from lifestyle-based diseases having a high risk of COVID-19. In future, accuracy level may be improved by adding more samples of data items and applying more powerful algorithms of machine learning or by increasing a greater number of hidden layers on MLP classifier. It will also improve by considering a greater number of diseases that individuals having because of their daily lifestyle. In the proposed study, only the Indian dataset is considered. In future, it may extend world-wide with a greater number of factors.

VII. REFERENCE


[17] Leo Dencelin X1, Ramkumar Analysis of multilayer perceptron machine learning approach in classifying protein secondary structures, Biomedical Research 2016; Special Issue: S166-S173


