Current Medicine Research and Practice xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Current Medicine Research and Practice

journal homepage: www.elsevier.com/locate/cmrp

Original article

Predicting mortality following gastrointestinal surgery

Ramesh Lal Sapra[a,*], Siddharth Mehrotra[b], Shailendra Lalwani[b], Vivek Mangla[b], Naimish Mehta[b], Samiran Nundy[b]

[a] Ganga Ram Institute for Postgraduate Medical Education and Research, Sir Ganga Ram Hospital, Old Rajinder Nagar, New Delhi 110060, India
[b] Department of Surgical Gastroenterology and Liver Transplant, Sir Ganga Ram Hospital, Old Rajinder Nagar, New Delhi 110060, India

ARTICLE INFO

Article history:
Received 1 November 2017
Accepted 22 December 2017
Available online xxx

Keywords:
Neural network
Gastrointestinal surgery
Imbalanced data
Predicting mortality

ABSTRACT

Introduction: Compared to survival data the accurate prediction of mortality following gastrointestinal surgery is a complex task. This is because mortality is rare and this makes the data unbalanced for classification. Consequently, the prediction accuracy tends to be biased towards the more abundant class i.e. survival. This essentially requires pre-processing of the data as well as an intelligent model to classify these cases. Our study aimed to identify the risk factors and develop an improved prediction model using such imbalanced data.
Methods: We examined retrospective data on 3500 patients admitted to our surgical gastroenterology unit and applied a linear logistic model to identify the risk factors for mortality as well as an artificial neural network(ANN) technique for predicting the mortality with greater accuracy.
Results: Logistic regression indicated that patients requiring inotropic support or having gastrointestinal haemorrhage were at an almost four times greater and patients admitted as emergencies were at almost two and a half times greater risk of dying. Among the eight ANN models, we identified two based on ten predictors one specifically for predicting survival with a high accuracy (93%) & sensitivity (98%); and the other for predicting mortality with a high accuracy (85%) & sensitivity (83%) using the synthetically modified oversampling technique(SMOTE) along with under-sampling of the majority class.
Conclusion: The ANN models with the SMOTE applied to the mortality class along with under-sampling of the survival class data provided a high prediction accuracy and sensitivity for mortality. However, the developed models need further testing on unseen cases.

1. Background

Predicting outcomes such as the mortality or survival of patients following gastrointestinal (GI) surgery depends upon a large number of factors which may directly or indirectly affect the outcome independently or jointly. Thus, predicting the mortality essentially requires an intelligent model that understands the pattern as well as the structure of the data. Over the past half century the logistic regression technique based on statistical principles has been the preferred choice for many researchers owing to its simplicity and easy interpretation of the predictors' role in assessing the outcome. However logistic regression, which is based on the cause and effect principle, may not be the right choice for understanding the complex nature of relationships between the outcome and a set of predictors including their interactions. Unlike conventional logistic models, ANNs are designed the way the human brain works – by incorporating additional hidden layers between the input and output layers in the models. ANNs are trained for classification through supervised learning on a subset of data to understand the pattern in the data. They are further tested and validated on the remaining data set.

One of the biggest challenges in the development of a prediction model for mortality is the imbalanced nature of classification data as mortality occurs much less frequently compared to survival. This nature of the data, consequently tends to bias the prediction towards the survival(majority) class[1] and hence requires the data to be subjected to some sort of pre-processing with special techniques before subjecting them to predictive modelling.[2]

The applicability of ANNs in the field of gastroenterology have been discussed at length and reviewed critically.[3–5] Their potential has also been explored in predicting the survival and mortality of patients in isolated groups(hepatocellular carcinoma, acute pancreatitis, liver cirrhosis, nonvariceal upper GI bleeding, colorectal and gastric cancer) but not in a generalized group of patients undergoing gastrointestinal surgery. Notably, very few

* Corresponding author.
E-mail address: saprarl@gmail.com (R.L. Sapra).

https://doi.org/10.1016/j.cmrp.2017.12.001
2352-0817/

studies have been done to predict mortality as compared to survival. ANN based mortality prediction has been made in patients for colorectal cancer,[6] cirrhosis of the liver,[7] end stage liver disease,[8] severe acute pancreatitis[9,10] and acute biliary pancreatitis.[11] Most of the ANN based studies conducted in the past for predicting outcomes (survival/mortality) have been designed for specific conditions with sample sizes ranging from 84 patients predicting survival from pancreatic cancer after radical surgery[12] to 37,500 patients predicting 5 year survival from colon carcinoma.[13] These networks have been designed having nodes from four[14] to 65 nodes.[15] To the best of our knowledge no general model has been developed which can predict the surgical outcome across different gastrointestinal operations giving due weightage to the accuracy of a minority (mortality) class prediction. We made an attempt to develop ANN prediction models with a minimal set of risk factors using retrospective data on a large number of cases. The developed models also take care of the imbalanced nature of the data for enhancing the predictive performance using synthetically modified oversampling technique (SMOTE).[16]

2. Methods

2.1. Risk factors studied

Our study included retrospective data on 37 factors (pre-, intra- & post-operative) in 3496 cases admitted to the gastrointestinal surgery department of Sir Ganga Ram Hospital, New Delhi, India between 2010 and 2014. The factors included in the analysis were: age, sex, operation category, platelet count (pre & postoperative) priority (emergency/elective), haemoglobin level (pre- & postoperative), serum albumin(pre & postoperative), serum bilirubin (pre- & postoperative), total leucocyte count, serum creatinine (pre & post), chronic renal failure, hypertension, diabetes mellitus, tuberculosis, coronary artery disease, hypothyroidism, chronic obstructive pulmonary disease, gastrointestinal haemorrhage (GIH), chronic liver disease, blood transfusion, blood loss, postoperative complications, ionotropic requirement and mortality.

2.2. Imputation of the data

We began with retrospective data of some 3500 cases. For ANN building, we excluded those cases where data on multiple clinical attributes were found to be missing. Missing numeric values were replaced either with the median or mean values depending upon the distribution. Missing nominal data, if any was substituted with the abundant class value. However, for logistic regression analysis case wise deletion was followed wherever a missing value occurred. A total of 3117 cases data was processed for building ANNs.

2.3. Conventional statistical and ANN models

We applied a conventional logistic model to identify the significant factors for predicting the mortality using forward likelihood function with the help of SPSS 17.0 software. The Pearson-Chi square test was used to study the association with the nominal attributes. Multilayer perceptron artificial neural network (ANN) models were designed with varying sets of predictors using WEKA 3.6 software.

2.4. Pre-processing of data and model development

To overcome the limitations of the imbalanced data, as discussed earlier, the data was pre-processed using SMOTE or combination of SMOTE plus under-sampling. Under-sampling was done by simply removing the majority (survival) class cases randomly until the minority (mortality) class became a specified percentage of the majority class. Whereas, SMOTE is a special technique in which the minority class is oversampled by creating "synthetic" examples along the line segments joining any/all of the k minority class nearest neighbours. We tried different levels of SMOTE i.e., 100, 200, 300 and 400%. However, we could not find an advantage in prediction accuracy beyond 400% of SMOTE. Each ANN model was validated by 10-fold cross validation. Here it is worth mentioning that k-fold cross validation is an improved method over the holdout method where generally some 70% of the data is used in training the model and rest 30% for its testing. The data set in k-fold cross validation is divided into *k* subsets, and the holdout method is repeated *k* times. Each time, one of the *k* subsets is used as the test set and the other *k-1* subsets are put together to form a training set. Consequently k-fold cross validation gives highly precise estimate of the prediction parameters. In addition to this, we further chose 10 random seeds for each model. Thus, our testing of the model was based on 100 random subsets where each subset consists of untrained data (10% of entire data set).

2.5. ANN models and attributes/factors selection

For the purpose of selecting an optimal set of attributes, we tried a number of ANN models (8) with multilayer perceptron architecture and back propagation algorithm having a single hidden layer with varying number of nodes selected automatically by the software. Since the data was highly imbalanced, the geometric mean (GM) of the sensitivities of the two classes along with the area under the ROC, were used for judging the performance of the models. Fig. 1 shows the flow chart of ANN modelling. Following were the eight models designed with varying number of predictors/attributes:

1. **M18**: ANN model with Pre-operative attributes (18)

Age, sex, priority (emergency/elective), platelet count(pre-) haemoglobin level (pre-), serum albumin(pre-), serum bilirubin (pre-), total leucocyte count (pre-), serum creatinine (pre-), chronic renal failure, hypertension, diabetes mellitus, tuberculosis, coronary artery disease, hypothyroidism, chronic obstructive pulmonary disease, gastrointestinal haemorrhage(GIH), chronic liver disease

2. **M36**: ANN model with all (Pre-,intra- and postoperative 36 attributes)

3. **M10**: ANN model with minimal set of attributes (10):

Serum albumin (pre- & postoperative), serum creatinine (pre- & postoperative), total leucocyte count (postoperative), serum bilirubin (postoperative) platelet count(postoperative), priority (emergency/elective), gastrointestinal haemorrhage (GIH) and ionotropic requirement. These ten attributes were selected through 'CfSubsetEval' attribute evaluator and 'BestFirst' search method of WEKA software. The 'CfSubsetEval' evaluates the worth of a subset of attributes by considering the individual predictive ability of each feature along with the degree of redundancy between them. The 'BestFirst' search method searches the space of attribute subsets by greedy hill climbing augmented with a backtracking facility.

4. **M10 + SM100**: Model developed on 100% SMOTE with ten attributes
5. **M10 + SM200**: Model developed on 200% SMOTE with ten attributes
6. **M10 + SM300**: Model developed on 300% SMOTE with ten attributes
7. **M10 + SM400**: Model developed on 400% SMOTE with ten attributes
8. **M10 + SM400 + US**: Model developed on 400% SMOTE with ten attributes plus under- sampling (US) of survival cases.

Please cite this article in press as: R.L. Sapra, et al., Predicting mortality following gastrointestinal surgery, Curr Med Res Pract. (2017), https://doi.org/10.1016/j.cmrp.2017.12.001