Predictive model of cardiac arrest in smokers using machine learning technique based on Heart Rate Variability parameter

Shashikant R. and Chetankumar P.
Department of Instrumentation and Control Engineering,
College of Engineering Pune, Pune, India

Abstract
Cardiac arrest is a severe heart anomaly that results in billions of annual casualties. Smoking is a specific hazard factor for cardiovascular pathology, including coronary heart disease, but data on smoking and heart death not earlier reviewed. The Heart Rate Variability (HRV) parameters used to predict cardiac arrest in smokers using machine learning technique in this paper. Machine learning is a method of computing experience based on automatic learning and enhances performances to increase prognosis. This study intends to compare the performance of logistical regression, decision tree, and random forest model to predict cardiac arrest in smokers. In this paper, a machine learning technique implemented on the dataset received from the data science research group MITU Skillogies Pune, India. To know the patient has a chance of cardiac arrest or not, developed three predictive models as 19 input feature of HRV indices and two output classes. These model evaluated based on their accuracy, precision, sensitivity, specificity, F1 score, and Area under the curve (AUC). The model of logistic regression has achieved an accuracy of 88.50%, precision of 83.11%, the sensitivity of 91.79%, the specificity of 86.03%, F1 score of 0.87, and AUC of 0.88. The decision tree model has arrived with an accuracy of 92.59%, precision of 97.29%, the sensitivity of 90.11%, the specificity of 97.38%, F1 score of 0.93, and AUC of 0.94. The model of the random forest has achieved an accuracy of 93.61%, precision of 94.59%, the sensitivity of 92.11%, the specificity of 95.03%, F1 score of 0.93 and AUC of 0.95. The random forest model achieved the best accuracy classification, followed by the decision tree, and logistic regression shows the lowest classification accuracy.

Keywords Cardiac arrest, Heart Rate Variability, Machine learning, Accuracy, Precision, Area under the curve

Paper type Original Article
1. Introduction
Long-term smoking is a significant and self-governing risk factor of cardiovascular disease, cardiac arrest, and coronary artery disease. According to the World Health Organization (WHO), concerning 1.1 billion people are smokers worldwide, among them, 7 million people die every year, and nearly 15,500 people die every day from smoking. Smokers are likely to develop ischemic heart disease at a younger age and are most likely to die of sudden death. Smoking makes the heart work considerably harder, lowers its oxygen supply, increases the possibility of coagulation in blood vessels, and increases the risk of heartbeat alterations [1,2].

HRV is a representation of changes in normal heartbeat rhythms. HRV is a non-invasive measuring tool for the assessment of the autonomous nervous system for heartbeat regulation. SA node maintains the normal heart rhythm, controlled by the autonomous nervous system's (ANS) sympathetic and parasympathetic branches [2,4]. Sympathetic activity tends to increase heart rate and decrease heart rate through parasympathetic activity. The prevalence of sympathetic and parasympathetic activity affects the heart's rhythm. Researchers have found that HRV parameter decreased in the case of cardiac disease in smokers. HRV parameters are, therefore, crucial for predicting heart disease.

In the previous studies, the cardiac arrest predictive model proposed on the Cleveland Clinical Foundation Heart Disease dataset, which is a part of the UCI machine learning repository. The data set has 76 raw attributes. However, all of the predictive experiments used only 13 attributes. The inputs attributes are Age, Sex, Chest Pain, Resting blood pressure, Serum cholesterol, Fasting blood sugar, Resting electrocardiographic results, Maximum heart rate achieved, Exercise-induced angina, ST depression, Slope of the peak exercise ST segment, Number of significant vessels colored by fluoroscopy and Thal. However, in the past study, there is no predictive model which can predict cardiac arrest in the smoker. In these predictive model, the time domain, frequency domain, and non-linear parameter used as the input attribute. HRV parameters are more accurate to predict cardiac arrest in the smoker. HRV not only address the present health status but also indicate the future occurrence of disease.

To predict the cardiac arrest, three machine learning predictive model implemented. Techniques of machine learning widely used in clinical diagnosis. It is a broad discipline with statistical and computer science foundations that endorse a set of different algorithms for predictive model construction. Machine learning does not require an alternate algorithm for the different data set. The objective of this study was to develop three predictive models, Logistic Regression (LOR), Decision Tree (DT) and Random Forest (RF) based on the HRV parameter for cardiac arrest prediction [3]. Sklearn, pandas, numpu, matplotlib packages used in a python tool for data manipulation to implement an algorithm for machine learning. The predictive model was assessed based on accuracy, precision, sensitivity, specificity, F1, and AUC score.

2. Method
HRV is analyzed using the time domain, the frequency domain, and the non-linear approach. The data set obtained from data science research group MITU Skillogies Pune, India (Available on- https://mitu.co.in). The data set includes a total of 1562 non-smoker and smoker instances belongs to the middle age group (40–60) from India, out of that 751 people are non-smokers, and 811 people are smokers. In the smoker group, cardiac arrest observed. The data set classified into cardiac arrest and non-cardiac arrest classes with 19 HRV input features (Attributes). The dataset verified by doctors (Table 1).

All of the above, indices are features of input to the predictive model of machine learning (Figure 1).

Machine learning by modeling makes predictions. Predictive modeling is the method of creating models that predict the final result. Machine learning intends to build computing
systems that can evolve to their knowledge and learn from them. Typically, machine learning functions categorized into three deep divisions. These are: 1) Supervised learning with a feature of a system that relies on categorized training data, 2) Unsupervised learning to which the learning model intends to indicate the unsorted data framework, and 3) Reinforcement learning is the system in which the complex environment cooperates.

In this paper, the supervised learning model implemented as the data set is categorized. The supervised model of learning aimed to predict the value of a variable called output variable from a set of variables called input variable. The set of input variable are characteristics called as feature/attributes. The set of input and output variable used as training and testing data. Training data is the known data, whereas testing data is the unknown data to be predicted. Logistic regression (LOR), Decision tree (DT), Random forest (RF), k-Nearest Neighbors (k-NN), Support vector machine (SVM), Naive Bayes (NB) and Artificial neural network (ANN) are some of the most common techniques [5–7]. Three machine learning predictive models used: Logistic regression, Decision tree, and Random forest. The details are below-

2.1 Logistic regression (LOR)
Logistic regression is effectively a linear classification model rather than the regression model. It is a standard method of categorization predicated on the data probabilistic statistics. This model describes variables of dichotomous output, which can be used to predict disease. Let us suppose our hypothesis is-

\[ h_{\gamma}(x) = g(\gamma^T x) = \frac{1}{1 + e^{-\gamma^T x}} \]  

(1)

based on this hypothesis, we get the sigmoid function or logistical function

\[
\begin{align*}
\text{Hemodynamic Parameter} & \quad 1. \text{SBP} \\
\text{Time Domain Parameter} & \quad 2. \text{DBP} \\
\text{Frequency Domain Parameter} & \quad 1. \text{Mean HR} \\
\text{Nonlinear Parameter} & \quad 2. \text{Mean RR} \\
\text{Class} & \quad 3. \text{SDNN} \\
\text{SBP-Systolic Blood Pressure, DBP-Diastolic blood pressure, HR-Heart Rate, RR-RR interval, SDNN-Standard deviation of normal to normal interval, RMSSD-Root mean square of standard deviation, TP-Total power, LF-Low frequency, HF-High frequency, ms²-Millimeter square, nu-Normalized unit, DFA-Detrended Fluctuation Analysis, AppEN-Approximate Entropy, SampEN-Sample Entropy.} \\
\end{align*}
\]

Table 1. HRV parameter/Number of Predictor.

| Hemodynamic Parameter | 1. SBP |
|-----------------------|--------|
| Time Domain Parameter | 2. DBP |
| Frequency Domain Parameter | 1. Mean HR |
| Nonlinear Parameter | 2. Mean RR |
| Class | 3. SDNN |
| | 4. RMSSD |
| | 5. TP |
| | 6. LF (ms²) |
| | 7. HF (ms²) |
| | 8. LF (nu) |
| | 9. HF (nu) |
| | 10. LF/HF |
| | 11. SD1 |
| | 12. SD2 |
| | 13. SD1/SD2 |
| | 14. DFA-α1 |
| | 15. DFA-α2 |
| | 16. AppEN |
| | 17. SampEN |
| SBP-Systolic Blood Pressure, DBP-Diastolic blood pressure, HR-Heart Rate, RR-RR interval, SDNN-Standard deviation of normal to normal interval, RMSSD-Root mean square of standard deviation, TP-Total power, LF-Low frequency, HF-High frequency, ms²-Millimeter square, nu-Normalized unit, DFA-Detrended Fluctuation Analysis, AppEN-Approximate Entropy, SampEN-Sample Entropy.
Figure 1. Partial View of the Data set displaying the data.

Predictive model of cardiac arrest in smokers.
The variable \( z \) represents the prominence to the set of the \( g(z) \) input variable. The variable \( z \) is an indicator of the contribution of all input variable used in the model. It is given as:

\[
z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \ldots \beta_n x_n
\]  

(3)

where \( \beta_0 \) is the intercept and \( \beta_1, \beta_2 \ldots \beta_n \) are regression coefficient. Logistic regression is a practical way to define the association between one or more variables of input and output, described as a probability that only has two possible values such as disease (‘YES’ or ‘NO’/‘1’ or ‘0’). We used ten-fold cross-validation on the training data set in our logistics model. LOR model gives 87–89% test data accuracy and a correct F1 score [5,7].

As the number of predictors is more, to create a less complicated model, regularization techniques used to address over-fitting. A regression model that uses the L1 regularization technique is called Lasso Regression and model which uses L2 is called Ridge Regression.

### 2.2 L1 regulation on least square

Least Absolute Shrinkage and Selection Operator combines the coefficient’s “absolute magnitude value” to the loss function as a penalty term.

\[
\sum_{i=1}^{n} \left( Y_i - \sum_{j=1}^{p} X_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^{p} | \beta_j |
\]  

(4)

The first term is the sum of square error term, and the second term is the penalty term. If lambda is zero, then we will get back square error term whereas immense value will make coefficients zero; hence, it will under-fit.

### 2.3 L2 regulation on least square

Ridge regression adds a coefficient of “squared magnitude” to the loss function as a penalty term.

\[
\sum_{i=1}^{n} \left( y_i - \sum_{j=1}^{p} x_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^{p} \beta_j^2
\]  

(5)

If lambda is zero, then we will get a square error term back here. If lambda is very large, however, it will append too much weight and result in under-fitting. Having said that how lambda is selected is essential. To avoid over-fitting issues, this technique works very well.

The critical difference among these techniques is that Lasso shrivels the coefficient of less significant feature to zero, so some features entirely removed. In case a large number of features is considered, this regularization technique fit for the selection of features. In this model, the L1 regularization technique used because it minimizes the unpredictability of the learned model by completely ignoring certain features, known as sparsity. L2 regularization is not valid for a selection of features but preferably seeks to reduce the model’s unpredictability by avoiding huge weighting of features.

### 2.4 Decision tree (DT)

A decision tree is a tree-like flowchart, building a binary tree. In the classification problem, the decision tree algorithm is most useful. A decision tree is an algorithm using supervised learning, data that already know the responses used to build the tree. Its performance is
mostly associated with the accuracy of the classification achieved on the training data set and the tree size. Decision tree algorithm is a strategic approach to developing models of classification from a collection of the training dataset. Decision tree structures constructed in a top-down nested form of dividing and conquering strategy.

Its framework involves training data modeling of nodes and branches. The first node is called the root node, separating each data until a termination criterion fulfilled. The decision tree consists of three structural features, which are (i) The root node (parent node) is an attribute selected as the base on which to build the tree, (ii) The internal node (child node) is the attributes that reside within the tree, (iii) The leaf node (terminating node) is the end node and the decision tree completed. The decision tree stopping criteria is that all samples belong to the same kind of class for a specified node; there are no residual attributes for more splitting [8]. There are many types of decision trees, but most commonly known are Information Gain (IG), Gini Index (GI) and Gain Ratio (GR) types. A decision tree can be produced using ID3, J-48, C4.5, C5.0 algorithms. Best accepted among, is C5.0 algorithms. Making the decision tree more compact and lowering the decision rule, pruning method used.

2.5 Random forest
Random forest is a classification method, a part of the ensemble learning model that integrates weak classifier predictions. It develops an indicator ensemble with a collection of decision trees growing in randomly chosen data subspace where each tree grew according to a discrete parameter in the ensemble [9]. It is quick and easy to implement, produces predictions that are highly accurate, and can handle a vast number of variables input without over-fitting. The algorithm starts with forming a combination of trees that will help each vote for a class; voting includes splitting the training data into smaller equal subsets and constructing a decision tree. The tree is built using the Random Forest algorithm as –

Let X be the number of classes, and Y be the number of variable in the data set.

- The input variable y is used to assess the node of the tree.
- Choose y variable randomly and calculate the best split for each tree node.
- The tree is finally fully grown and not pruned. A new sample to predict, the tree is pulled down. At the end of the terminal node, the training sample ascribed to the label. This procedure is repeated several times across all trees and observed as a prediction of Random Forests [10].

3. Predictive model
In our predictive model, Dataset collection block contains patient details of smokers suffered from heart disease. Feature/Attribute selection process selects the critical features for the prediction of cardiac disease. After feature selection, preprocessing involved to remove the outlier and make dataset normalized. Min-max normalization most often referred to as feature scaling in which the numerical range values of a data feature, i.e., a property, are lowered to a scale between 0 and 1. The following formula used to calculate z, i.e., the normalized value of a member of the set of observed values of x-

\[
z = \frac{x - \text{min}(x)}{\text{max}(x) - \text{min}(x)}
\]  

where min and max are in x given their range, the minimum, and maximum values.

Various classification techniques applied to preprocessed data. Finally, model evaluation is performed based on different measures (Figure 2).
4. Result and discussion

Evaluation of the model is the processes for calculating the effectiveness of the data set results. Data manipulation is carried out using a python tool. The dataset divided into two parts for training and testing purpose. We trained our model with 80% training data and tested the remaining 20% data. In this study, we used 10-fold validation method to measure the performance of the entire classification technique. Various statistical measurement aspects such as accuracy, precision, sensitivity, specificity, F1 score, AUC evaluate the performance of all classification algorithms.

Accuracy is the measure of the model’s correct predictions. Precision is used to determine the classifier’s ability to deliver accurate positive predictions. Sensitivity measures the positive instances that the classifier identifies as having heart disease [9]. Specificity is used to assess the classifier’s potential to examine cases of negative cardiac arrest. F1 score measures a weighted precision and sensitivity average. For the classification algorithm excellent performance, F1 score must be 1 and 0 for the bad performance. The classifier AUC value ranges from 0.5 to 1. The AUC value below 0.5 implies that the classifier could not differentiate between true and false; an appropriate classifier is worth close to 1 [10]. ROC is an accuracy measure. It has two dimensions, the x-axis represents specificity (False positive rate), and the y-axis represents sensitivity (True positive rate) [11,12].

The detailed predictions generated from the training and testing data set described in the form of confusion matrices. A confusion matrix is a matrix of classification results. Tables 2 and 3 shows the result in tabular form.

**Table 2.**

| Sr. No. | Predictive Model     | Accuracy | Precision | Sensitivity | Specificity | F1 Score | AUC |
|---------|----------------------|----------|-----------|-------------|-------------|----------|-----|
| 1       | Logistic Regression  | 89.67%   | 90.04%    | 88.72%      | 90.58%      | 0.89     | 0.91|
| 2       | Decision Tree        | 91.27%   | 98.23%    | 85.26%      | 98.90%      | 0.91     | 0.94|
| 3       | Random Forest        | 98.64%   | 99.67%    | 97.56%      | 99.68%      | 0.99     | 1   |

**Table 3.**

| Sr. No. | Predictive Model     | Accuracy | Precision | Sensitivity | Specificity | F1 Score | AUC |
|---------|----------------------|----------|-----------|-------------|-------------|----------|-----|
| 1       | Logistic Regression  | 88.50%   | 83.11%    | 91.79%      | 86.03%      | 0.87     | 0.88|
| 2       | Decision Tree        | 92.59%   | 97.29%    | 90.11%      | 97.38%      | 0.93     | 0.94|
| 3       | Random Forest        | 93.61%   | 94.59%    | 92.11%      | 95.03%      | 0.93     | 0.95|
The current study found that, the logistic regression model achieved a classification accuracy of 88.50% with a precision of 83.11%, sensitivity of 91.79%, specificity of 86.03%, F1 score of 0.87 and AUC of 0.88; the decision tree (C5.0) reached to an accuracy of 92.59% with precision of 97.29%, sensitivity of 90.11%, specificity of 97.38% F1 score of 0.93 and AUC of 0.94. However, among the three models assessed, random forest performed best. The random forest had a classification accuracy of 93.61% with a precision of 94.59%, sensitivity of 92.11%, the specificity of 95.03%, F1 score of 0.93, and AUC of 0.95. The ROC curve of all three models is given in the following figure. The random forest model showed better performance than the decision tree model, and the decision tree model reported better than the logistic regression. The study result showed that the best predictor is the random forest model (Figures 3–5).

![Figure 3](image1.png)
**Figure 3.** ROC curve for the Logistic Regression Model.

![Figure 4](image2.png)
**Figure 4.** ROC curve for Decision Tree Model.
4.1 Hyperparameter optimization

Hyperparameter optimization or tuning is the issue in machine learning to determine a set of ideal hyperparameters for an algorithm of learning. A hyperparameter is a parameter that measures the process of learning using its value. Hyperparameters are meta parameters which are associated with the learning algorithm. Finding the best values for hyperparameters that generalizes the model for better accuracy is Hyperparameter tuning/optimization. Performance of the machine learning model is dependent on the various hyperparameter such as hidden layers, several units per layer, activation function, regularizer, learning rate.

The value of the hyperparameter can be changed manually by machine learning engineer before training the model explicitly. In this study, the machine learning algorithm is Logistic Regression, Decision Tree, and Random forest. Hyperparameter of these models are (Table 4).

The logistic regression model requires actual inputs and predicts the likelihood of the input corresponding to the preferred class. If the probability is >0.5 the output taken as the preferred class, otherwise the other class predicts. The logistic regression has coefficients observed in Eq. (3). The learning algorithm’s task to find the highest values based on the training data for the coefficients ($\beta_0, \beta_1$ and so on). Using stochastic gradient descent, we can estimate the coefficient values. We can use a straightforward update equation to calculate the current coefficient values.

$$\beta_0 = \beta_0 + \alpha \times (y - \text{prediction}) \times \text{prediction} \times (1 - \text{prediction}) \times \beta_0$$

where $\beta_0$ is the coefficient for the update, and the performance of predicting using the model is the prediction. Alpha is the parameter need to define before the training. This is the learning

| Algorithm          | Hyperparameters                  |
|--------------------|----------------------------------|
| Logistic Regression| - Learning Rate                  |
|                    | - Regularizer                    |
|                    | - Depth of Trees                 |
|                    | - Number of Decision Trees       |
| Decision Tree      |                                  |
| Random Forest      |                                  |

Figure 5.
ROC curve for Random Forest Model.
rate and regulates how much the coefficients change or learn every time the model is updated. In Eq. (7), the x term represents input value for the coefficient and \( \beta_0 \) represents the value of intercept, which considered to be 1. The learning rate alpha returns how rapidly we updated the parameters. We updated the model by the different learning rate. If the value of alpha is more, it will overshoot the optimal value; it is too small, it requires too many iterations to get the optimal value. Hence it is crucial to the used well-tuned learning rate. We updated the model by the different learning rate. At 0.001 learning rate, we got the optimal accuracy value (Table 5).

In the Decision Tree model, depth of tree model decides the accuracy of the algorithm. Initially, the training, testing accuracy of the decision tree model was 100% and 88.10% respectively by keeping the default values of hyperparameter, which results in overfitting of the decision tree. In the real world scenario, the model must perform well on testing data not just on training data (Figure 6 and Table 6).

| Hyperparameter of Logistic Regression | Tuned to |
|--------------------------------------|----------|
| Penalty                              | ‘L1’ Regularizer |
| Alpha (learning rate)                | 0.001    |

Table 5. Hyperparameter of Logistic Regression Model.

| Hyperparameter of Decision Tree model | Tuned to |
|--------------------------------------|----------|
| Criterion                            | ‘gini’, ‘entropy’ |
| Depth of Trees                       | 2        |

Table 6. Hyperparameter of Decision Tree Model.

| Hyperparameter of Random Forest model | Tuned to |
|--------------------------------------|----------|
| Criterion                            | ‘gini’, ‘entropy’ |
| No. of Decision Trees                | 10       |
| Maximum Features                     | Auto     |

Table 7. Hyperparameter of Random Forest Model.
Hyperparameters for a random forest include the number of decision trees in the forest and the number of characteristics that each tree considers when dividing a node. The variables and thresholds used to divide each node learned during practice are the parameters of a random forest. In this model, we optimized the value of the number of decision tree and the number of featured considered by each tree (Table 7).

n_estimators:
The number of trees constructed before taking the maximum vote or prediction averages. The more significant number of trees will offer higher performance but slow down the process. The value of decision tree chosen based on the capability of the processor, which makes predictions more stable.

max_features:
These are the highest amount of features that can be tried in an individual tree by Random Forest. There are numerous choices for assigning maximum features in Python. Here are some of them:

Auto/None: This will take all the features that make sense in each tree.
Sqrt: Square root choice will take the total quantity of features in a single run. For example, if the total number of variables is 25, the algorithm takes only 5 of them in the individual tree (Table 7).

In the previous study, cardiac arrest prediction based on the input attribute like blood pressure, cholesterol, blood sugar, chest pain, blood sample parameter, ECG results. In this study, the prediction is based on the HRV parameter and more accurate than existing method, this is the uniqueness of the study.

5. Conclusion
In summary, we compared three predictive models used 19 attributes of HRV to predict cardiac arrest in smokers. The result indicated that the random forest model performed best on the accuracy, precision, sensitivity, specificity, F1 score, and AUC. This study can help future researchers to choose the model of deep learning to obtain more accurate results.

References
[1] L.N. Coughlin, A.N. Tegge, C.E. Sheffer, W.K. Bickel, A machine-learning approach to predicting smoking cessation treatment outcomes, Nicotine Tob. Res. (2018).
[2] F. Lombardi, T.H. Mäikikallio, R.J. Myerburg, H.V. Huikuri, Sudden cardiac death: role of heart rate variability to identify patients at risk, Cardiovasc. Res. 50 (2) (2001) 210–217.
[3] R. Devi, H.K. Tyagi, D. Kumar, Heart rate variability analysis for early stage prediction of sudden cardiac death, World Acad. Sci. Eng. Technol. Int. J. Electr. Comput. Energy. Electron. Commun. Eng. 10 (3) (2016).
[4] U.R. Acharya, K.P. Joseph, N. Kannathal, C.M. Lim, J.S. Suri, Heart rate variability: a review, Med. Biol. Eng. Comput. 44 (12) (2006) 1031–1051.
[5] M. Hassan, M.A. Butt, M.Z. Baba, Logistic regression versus neural networks: the best accuracy in prediction of diabetes disease, Asian J. Comp. Sci. Technol. 6 (2) (2017) 33–42.
[6] D. Khanna, R. Sahu, V. Baths, Deshpande, Comparative study of classification techniques (SVM, logistic regression, and neural networks) to predict the prevalence of heart disease, Int. J. Machine Learn. Comput. 5 (3) (2015) 414.
[7] K. Balasubramanian, R.N. Kumar, Improvising heart attack prediction system using feature selection and data mining methods, Int. J. Adv. Res. Comp. Sci. 1 (4) (2010).
[8] M.M. Kirmani, S.I. Ansarullah, Prediction of heart disease using decision tree a data mining technique, IJCSN Int. J. Comp. Sci. Network. 5 (6) (2016) 885–892.

[9] D. Ramesh, Y.S. Katheria, Ensemble method based predictive model for analyzing disease datasets: a predictive analysis approach, Health Technol. (2019) 1–13.

[10] J. Patel, D. TejalUpadhyay, S. Patel, Heart disease prediction using machine learning and data mining technique, Heart Disease 7 (1) (2015) 129–137.

[11] H. Kaur, V. Kumari, Predictive modelling and analytics for diabetes using a machine learning approach, Appl. Comp. Inf. (2018).

[12] R. Kumar, A. Indrayan, Receiver operating characteristic (ROC) curve for medical researchers, Indian Pediatr. 48 (4) (2011) 277–287.

**Corresponding author**
Shashikant R. can be contacted at: shashikantrathod.bme@gmail.com

For instructions on how to order reprints of this article, please visit our website: [www.emeraldgrouppublishing.com/licensing/reprints.htm](http://www.emeraldgrouppublishing.com/licensing/reprints.htm)

Or contact us for further details: permissions@emeraldinsight.com