

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Nonlinear Systems in Healthcare towards Intelligent Disease Prediction

*Parag Chatterjee, Leandro J. Cymberknop  
and Ricardo L. Armentano*

## Abstract

Healthcare is one of the key fields that works quite strongly with advanced analytical techniques for prediction of diseases and risks. Data being the most important asset in recent times, a huge amount of health data is being collected, thanks to the recent advancements of IoT, smart healthcare, etc. But the focal objective lies in making sense of that data and to obtain knowledge, using intelligent analytics. Nonlinear systems find use specifically in this field, working closely with health data. Using advanced methods of machine learning and computational intelligence, nonlinear analysis performs a key role in analyzing the enormous amount of data, aimed at finding important patterns and predicting diseases. Especially in the field of smart healthcare, this chapter explores some aspects of nonlinear systems in predictive analytics, providing a holistic view of the field as well as some examples to illustrate such intelligent systems toward disease prediction.

**Keywords:** nonlinear systems, healthcare, artificial intelligence, computational intelligence, machine learning, predictive analytics, chronic disease, cancer, cardiometabolic disease, Parkinson's disease

## 1. Introduction

*“Prevention is better than cure”*

*—Desiderius Erasmus*

A nonlinear system is a system in which the change of the output is not proportional to the change of the input [1–3]. Especially in the field of healthcare, most of the health systems being inherently nonlinear in nature, nonlinear systems are of special interest to researchers hailing from multidisciplinary areas. Nonlinear dynamical systems, describing changes in variables over time, may appear chaotic, unpredictable, or counterintuitive, contrasting with much simpler linear systems. Nonlinear modeling still has not been able to explain all of the complexity present in human systems, and further models still need to be refined and developed. However, nonlinear modeling is helping to explain some system behaviors that linear systems cannot and thus will augment our understanding of the nature of complex dynamic systems within the human body in health and in disease states [4].

The delivery of healthcare is a complex endeavor at both individual and population levels. At the clinical level, the tailored provision of care to individuals is guided, in part, by medical history, examination, vital signs and evidence. In the twenty-first century these traditional tenets have been supplemented by a focus on learning, metrics and quality improvement. The collection and analysis of data of good quality are critical to improvements in the effectiveness and efficiency of health care delivery [5]. This is also catalyzed by the boost in the field of eHealth across the world. eHealth is emerging as a promising vehicle to address the limited capacity of the health care system to provide health behavior change and chronic disease management interventions. The field of eHealth holds promise for supporting and enabling health behavior change and the prevention and management of chronic disease [6].

With a global increase in the adoption of Electronic Health Records (EHRs) [7–12], the volume and complexity of the data generated increases in all dimensions. In addition to the EHR-sourced patient data, the additional data available from other sources like the data about medical conditions, underlying genetics, medications, and treatment approaches is humongous. But human cognition to learn, understand, and process the data being finite [13], the traditional medical methods of analysis does not stand always to be the most efficient. Thus, computer-assisted methods to organize, interpret, and recognize patterns from these data are needed [14].

In the recent years, the underlying value of data is unfolding like never before and newer systems are being developed concentrating on the data analysis to make sense of the data. Especially in the field of healthcare, the aspect of intelligent data analytics is one of the most trending topics worldwide. One of the focal areas where such analyses have been applied is in the field of chronic diseases. By 2020, chronic diseases are expected to contribute to 73% of all deaths worldwide and 60% of the global burden of disease. Moreover, 79% of the deaths attributed to these diseases occur in the developing countries. Four of the most prominent chronic diseases—cardiovascular diseases (CVD), cancer, chronic obstructive pulmonary disease and type 2 diabetes are linked by common and preventable biological risk factors, notably high blood pressure, high blood cholesterol and overweight, and by related major behavioral risk factors. Action to prevent these major chronic diseases should focus on controlling these and other key risk factors in a well-integrated manner [15]. Apart from the chronic diseases, a key area where nonlinear models are applied from the perspective of intelligent prediction is human movement and locomotion. This leads to topics like fall detection, abnormal gait detection and diseases like Parkinson's. The outreach of intelligent prediction is spread to wider domains like transplantations [16], for example, to predict the success of a liver transplant by analyzing all the relevant health parameters.

Moreover, with the recent trends of smart sensors and eHealth devices powered by the Internet of Things (IoT), the data acquired is more comprehensive and detailed. Both prediction and prevention systems in this case usually use some fundamental steps in common, like collection of data from sensors and its analysis, followed by computing the risk and other possibilities [17]. The entire pool of data originating from this field is mostly nonlinear, invoking the need for the development of nonlinear analysis and predictive models.

Exploring the possible actions toward prevention of the chronic diseases, the key challenge lies in early detection of the diseases. Most of these diseases do not exhibit clearly identifiable signs at the early stage. This leads to harvesting the possibility of early detection of these diseases using artificial intelligence (AI). From the perspective of data science, the fundamental and most valuable resource in this aspect is the health data.

The health data holds immense potential for detailed analyses towards the early detection and prediction of diseases. The prediction of diseases using computational intelligence is multilevel. Most of the health systems being inherently nonlinear in nature, it provides an enormous opportunity to analyze those intricate details of the health systems while searching for the traits or early signals of diseases. On the other hand, given that the importance of health data (mostly the superficially and non-invasively obtained behavioral, physiological and metabolic health data) is quite crucial, a huge opportunity lies on the aspect of analysis of this health data toward the prediction of diseases. The computational aspect of disease prediction is also multifold, including aspects of data analysis, signal and image processing and other fields. However, this chapter is focused to the aspect of data analytics and computational intelligence, highlighting the key aspects of the health data of people pertaining to nonlinear systems, discussing the field of machine learning and intelligence toward the prediction of diseases.

## **2. Data modeling in healthcare toward predictive analysis**

*“Data is the new oil. It’s valuable, but if unrefined it cannot really be used.”*

*—Clive Humby*

The backbone of the intelligent prediction systems in healthcare is the data. Thanks to the skyrocketing advancements in data collection strategies and tools, it estimates a yearly growth of 48% [18] with projected growth rate to more than 2000 exabytes by 2020 [19]. On one hand, this poses an enormous challenge to handle this data. But on the other hand, this also keeps the potential in performing detailed analysis toward interesting insights. To understand the health data from a holistic view, the attributes like volume, variety, velocity, and veracity need to be considered, to decipher its value. This implies that on a larger scale, handling the health data as big data is inevitable. The first part is making this data suitable for analysis. About 80% of the world’s healthcare data being unstructured [20], it poses a huge challenge for the data preprocessing. Health data obtained from multiple sources often lack seriously in the aspect of interoperability or uniformity to model and process together. Even with the rise of EHRs, a major challenge lies in normalizing the data and making it suitable for modeling. With a lot of heterogeneous smart eHealth devices and components of IoT, tying the data together stands extremely difficult. For this reason, several EHR management system are being designed recently considering the data models and the aspect of preprocessing; thus, the EHRs are collected in such a way that fits the data models or facilitates the same. An inclusive data preprocessing system holds immense potential to support the aspect of health data modeling from a comprehensive perspective. Especially in the case of nonlinear systems, the aspect of data modeling is critical for processing enormous health data. Badly constructed data models not only skew the results but may also produce erroneous insights appearing to be correct.

The traditional way modeling the health data includes the field of data mining, with the aspects of knowledge discovery in databases (KDD) and exploratory data analysis (EDA). Most of these old KDD techniques visualize data from a perspective of database, looking for interesting knowledge in the data from a high-level point of view. These techniques provide better view to the data and especially in the field of healthcare, is often quite useful in analyzing patterns in large health datasets. However, specifically in the area of prediction of diseases and risks, recent systems are focused toward predictive models to perform a more formal, scalable and more efficient analysis and prediction. Comparing

with the traditional data mining models, these predictive models are made in a tailored-approach, mostly dedicated to a specific goal (for example, prediction of cardiovascular diseases). Often, such models also involve machine learning and computational intelligence for the aspect of analysis and prediction.

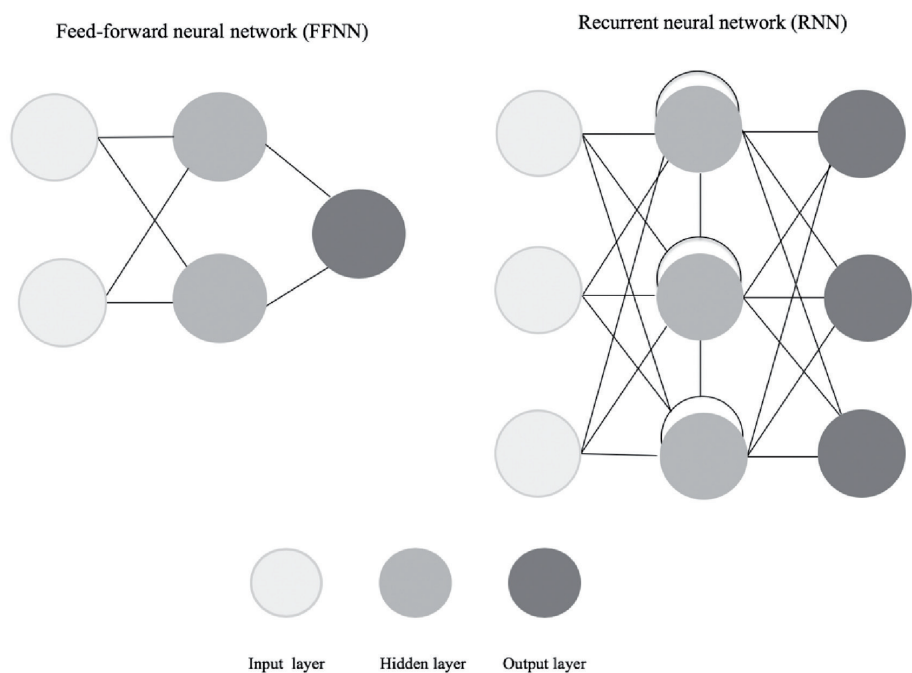
One of the most crucial tasks in this respect is to identify the mathematical model of a system from measurements of the system inputs and outputs. Especially in the field of disease and risk prediction, the data handled is mostly multidimensional. Keeping the focus toward nonlinear modeling, the first check is of course the identification of the data model, if it is a linear model or a nonlinear one. This is usually done using the superposition principle (properties of additivity and homogeneity). However, in some cases for finding the pattern in the health data, linear models do count useful. A more crude but common approach is to start with a linear model. After the initial tests to check the suitability of the linear model, which if turns out not good enough, leads to the replacement by a nonlinear approach to model the system.

Among several nonlinear models used in analyzing health data especially aimed at intelligent disease prediction, methods like nonlinear regression, clustering, decision trees, nonlinear support vector machines (SVMs), and artificial neural networks (ANNs) are quite profuse in recent times within the field of predictive medical analytics. For example, in nonlinear regression, the observational health data is modeled by a function which is a nonlinear combination of the model parameters and depends on one or more independent variables. The data is fitted by a method of successive approximations. If the health data available is labeled, supervised learning models like SVMs are used often, to analyze the data using classification and regression analysis. Given a set of training examples, each marked as belonging to one or the other of two categories, a SVM training algorithm builds a model that assigns new examples to one category or the other. Training data is usually divided into a training data (70%) and test data (30%). To map nonlinear functions, kernels can be used in SVMs. A kernel is a function that maps the data into a higher dimensional space where the linear mapping is possible. One of the main advantages of SVM with respect to modeling nonlinear systems is the possibility to use kernels, making it possible to represent very complex functions. However, when compared to linear regression, the main drawback is the need of more training and prediction times [21].

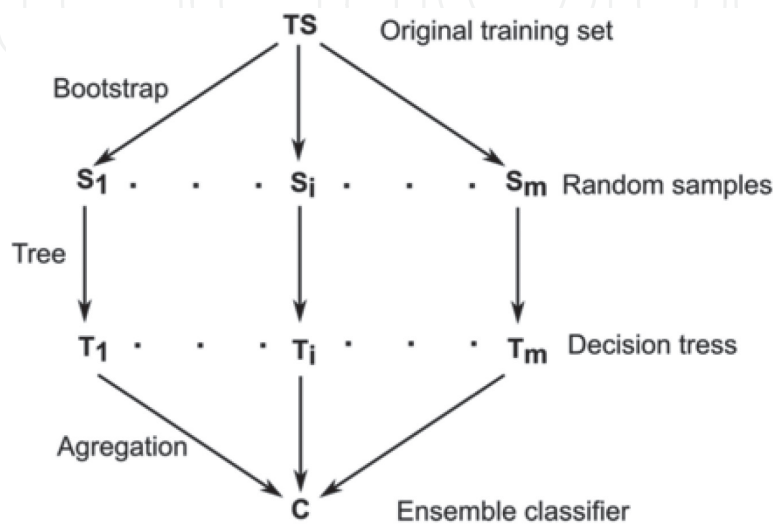
Neural networks on the other hand encompass a large class of models and learning methods and are nonlinear statistical models [22]. A recent survey of AI applications in healthcare reported uses in major disease areas such as cancer or cardiology and artificial neural networks as a common machine learning technique [23]. Such networks are organized in layers made of a number of interconnected nodes which contain an activation function. Data is provided to the network via the input layer, following which, the processing is performed in one or more hidden layers using a system of weighted connections. The last hidden layer is linked to the output layer where the result is given [21]. The healthcare domain of intelligent risk prediction is largely governed by the aspect of pattern recognition or finding relationships among several health and behavioral parameters and to study their impacts. One of the principal advantages of ANN (**Figure 1**) [24] is that it can model different types of relationships; systems which otherwise may have been very difficult to represent correctly could be modeled quickly and relatively easily using ANNs. However, compared to other types of networks, ANNs tend to be slower in training. Despite being a system of parallel computation, the slowness of the training step is due to the fact that individual artificial neurons are usually processed sequentially [21].

Ensemble classifiers are constructed from a given training data set and predict the class of a previously unseen object by combining the predictions obtained from

these basic classifiers. The importance of different ensemble classifiers has also been at rise attributing to the possibility of determining the risk groups among patient population. For example, the family of simple probabilistic classifiers like naive Bayes classifiers discover application in programmed medicinal analysis [25]. Performing regular analysis of healthcare for a large population makes it possible to act early in the case of health hazards and risks [26]. Clinical decision support systems often count useful in this aspect to assist the medical personnel in designing treatment strategies [27]. Such systems are mostly constructed using decision trees. Decision trees are flowchart structures in which each internal node denotes a test on a characteristic, each branch signifies the result of the test, and each leaf node denotes a class label. The paths from the root to leaf denotes classification rules [25]. Random forest algorithms (**Figure 2**) [21] are specifically suited for decision tree classifiers. In this technique, the basic classifiers are decision trees obtained by manipulating the input features. Basically, random forest builds multiple decision



**Figure 1.**  
*Artificial neural networks.*



**Figure 2.**  
*Working principle of random forest.*

trees and merges them together to get a more accurate and stable prediction. A key advantage of random forest is, that it can be used for both classification and regression problems; a huge part of current machine learning systems in health-care is related to such problems. However, the major limitation of random forest algorithms is that a large number of trees can make the algorithm slow down and ineffective for real-time predictions. In general, these algorithms are fast to train, but quite slow to create predictions once they are trained. A more accurate prediction requires more trees, which results in a slower model [28].

Unsupervised learning techniques are also profuse in health analytics especially in the field of analyzing health risks. Since a considerable part of health-data often arrive unlabeled, unsupervised learning methods help in finding patterns in the data or to analyze the health scenario over a big population. In this aspect, clustering techniques like k-means, gaussian distribution models, and mean-shift clustering often stand very useful in separating a group of patients into different clusters and then to analyze in detail the salient features and distinct characteristics. Among all the unsupervised learning algorithms, clustering via k-means might be one of the simplest and most widely used algorithms. Briefly, k-means clustering aims to find the set of k clusters such that every data point is assigned to the closest center, and the sum of the distances of all such assignments is minimized [29]. Especially when the relationships among different health parameters and their respective impacts are not known well, clustering techniques are used to separate a patient population in order to study the distinguishing features and influencing factors. In this aspect, nonlinear techniques of clustering also count useful.

Another crucial aspect with respect to machine learning algorithms is the aspect of bias and variance. The bias is an error from erroneous assumptions in the learning algorithm. High bias can cause an algorithm to miss the relevant relations between features and target outputs (underfitting). The variance is an error from sensitivity to small fluctuations in the training set. To visualize the degree to which a machine learning algorithm is suffering from bias or variance with respect to a data problem, learning curves are important tools. Learning curves are displays in which the performance of the machine learning algorithms are plotted with respect to the quantity of data used for training where the plotted values are the prediction error measurements [30].

Nevertheless, the choice of algorithm is dependent on multiple factors, the most important being the type of the dataset. Apart from that the aspect of prior knowledge about the data, computational complexity and expected results are also deciding factors, and the correct use of the model is extremely crucial in this regard [31]. Recent research has delved into uniting different techniques to provide hybrid machine learning algorithms [32]. Nevertheless, it is clear that the use of machine learning and computational intelligence takes an active role in predicting the health risks and the probability of diseases using the intelligence hidden in the health data.

### **3. Applications of computational intelligence toward prediction of diseases**

The use of computational intelligence with an objective to predict the health risks and diseases is an extensive and multi-step process. In this section, different scenarios are explained with respect to predictive analytics, aimed at disease and risk prediction. The first case is based on cardiometabolic diseases, and therefore, the most important component of the predictive system stands to be the detailed physiological and health data.

For example, in the field of cardiometabolic diseases, several parameters (age, gender, systolic and diastolic blood pressure, cholesterol, diabetes and smoking

habits) are considered for patients registered in the database and their risk scores was calculated. The risk scores give a general idea to classify the entire population into high and low risk groups [33]. Nevertheless, alternative analyses are performed to identify the underlying risk groups for each health parameter in the entire population. But this analysis is scalable for a further detailed technique for prediction of risks and influence of different health parameters on the cardiometabolic disease of a population. For example, Framingham Risk Score is used to predict the hard-coronary heart diseases (myocardial infarction or coronary death) and is calculated based on predictors like age, total cholesterol, high-density lipoproteins (HDL), systolic blood pressure, treatment for hypertension and smoking status [34]. In this case, the Framingham Risk can be expressed as a linear equation considering all the parameters. However, considering sample sets of people which are smaller and more specific, there exists the possibility of nonlinear relationships of several other parameters pertaining to cardiovascular risks, not usually considered in classical risk prediction models.

The traditional approach based on the identification and treatment of risk factors has proven to be insufficient and ignores that the detection of its subclinical stage is valid to define cardiovascular risk strategies. Taking into account that the artery is the main protagonist in this disease, it is necessary to evaluate it directly through a morpho-structural and functional analysis with non-invasive, reliable, reproducible procedures that are applicable in the youngest population. The detection of subclinical disease and the precocity with which it is done defines a safe framework to derive the real individual cardiovascular risk. Because coronary calcifications are an early marker of atherosclerosis detectable non-invasively; a model of cardiovascular risk that incorporates them along with the classic risk factors could have a remarkable interest in clinical practice, having the potential to change the field of preventive cardiology. The traditional approach guides the prevention and treatment of arterial disease, atherosclerotic in particular, based on the detection of cardiovascular risk factors (e.g., hypertension, smoking, dyslipidemia etc.). This approach quantifies the probability (risk) that the subject has a cardiovascular event (accident) in the next 10 years of life. Thus, based on information from large populations and global cardiovascular risk tables, and information obtained regarding the risk factors of each subject (for example, blood pressure, blood lipids (LDL, HDL), etc.), this can be classified in one of three possible categories: low risk, intermediate risk or high risk. However, this method has limitations since it does not take into account the individual cardiovascular risk and to detect early atherosclerosis and other alterations of the arterial wall. It has been demonstrated that a significant number of people considered to be at intermediate risk with the traditional approach, in fact, have a high risk of presenting a cardiovascular event (for example, they have significant atheroma plaques in the coronary arteries). Moreover, quantitatively most deaths due to cardiovascular causes occur in subjects who present low or intermediate risk, evaluated by the traditional approach. This underestimation of individual risk, which shows the traditional risk quantification approach, determines that millions of people do not receive adequate medical treatment every day to reduce their cardiovascular risk. In other words, asymptomatic subjects, but vulnerable to having a cardiovascular or cerebrovascular accident in the short term, are not offered the benefits of available prophylactic therapies, because they have underestimated their real cardiovascular risk. For example, hypertension is considered an asymptomatic disease and is easy to detect; however, it has serious and lethal complications if it is not treated in time.

To demonstrate how coronary artery calcium (CAC) can be incorporated into the risk of traditional was calculated in 618 male patients, the Framingham model

and the probability that the CAC of each patient falls in every four categories of CAC (0, 1–100, 101–400 and >400) using linear and nonlinear regression models. Then they were adjusted based on a relative risk (RR) that weighted the risk of coronary heart disease in individuals and that are  $RR = 1.7$  (for a CAC of 1–100),  $RR = 3.0$  (for CAC 101–400), and  $RR = 4.3$  (for CAC > 400) obtained from a meta-analysis published by Fletcher. The predictive power was evaluated using ROC curves (receiver operating characteristic). The model included in the CAC has a remarkable predictive value of atherosclerosis of 0.74, which is the area of the ROC curve as a function of the number of sites with extracoronary plates including carotid, femoral and abdominal aorta (coded as 0–1 sites = 0; 2–3 sites = 1). The predictive scale indicated 0.90–1 = excellent, 0.80–0.90 = good, 0.70–0.80 = median, 0.60–0.70 = weak, 0.50–0.60 = zero. The calcium score is a numerical information that allows quantifying the magnitude of coronary atherosclerotic lesions and provides independent predictive information of risk factors in general mortality. The combination of modeling of the CAC with the modeling of conventional risk factors leads to a remarkable improvement in the predictive value of the overall risk assessment of Framingham through the reclassification of the risk of atherosclerosis to a degree that may be clinically important. Adding to this approach, the other indices of subclinical atherosclerosis such as arterial rigidity, intima media thickness, endothelial function, and the presence of plaques will generate an integrative risk that will determine and classify the subjects in relation to short-term risk of suffering a cardiovascular or cerebrovascular accident. It will allow to know in a more precise way the cardiovascular risk of a particular individual. It allows early detection (subclinical stage) of vascular alterations and offer the best current prophylactic therapies available [35–37].

All standard risk assessment models to predict cardiovascular diseases make an implicit assumption that each risk factor is related in a linear fashion to CVD outcomes [38]. Such models may thus oversimplify complex relationships which include large numbers of risk factors with nonlinear interactions [39]. The aspect of computational intelligence comes into play specifically in this situation to decipher the inherent patterns and relationships among the parameters apart from the known and formally specified set, thereby determining more nuanced relationships between risk factors and outcomes. Current approaches to predict cardiovascular risk fail to identify many people who would benefit from preventive treatment, while others receive unnecessary intervention. Machine learning offers the opportunity to improve accuracy by exploiting complex interactions between the risk factors [39]. The established ACC/AHA 10-year risk prediction model used eight core baseline variables (gender, age, smoking status, systolic blood pressure, blood pressure treatment, total cholesterol, HDL cholesterol, and diabetes). However, in [39], additional 22 variables (like Body Mass Index, Triglycerides, C-reactive protein, Serum fibrinogen, etc.) were included in the machine learning algorithms, aimed at finding the influence on cardiovascular diseases, thereby designing the predictive algorithm for the same. Among other machine learning algorithms, neural networks had a 3.6% improved prediction than the existing algorithm. The system of cardiovascular risk prediction varies widely based on geographical factors. Therefore, several risk scores have been developed in different parts of the world like the SCORE by the European Society of Cardiology or the HellenicSCORE in Greece to address more accurately a specific group of population for calculating the cardiovascular risks. The majority of these scores use a common set of the ‘classical’ CVD risk factors, e.g., age, sex, smoking, blood pressure and lipids levels, whereas others have also incorporated more advanced markers of CVDs.

Most of these risk-prediction tools are based on stochastic models, incorporating variables, based on cohort studies [40]. However, the alternative approaches of

machine learning like k-nearest neighbors, random forests and decision trees also generate results quite comparable to the classical risk prediction scores [41], thus demonstrating its possibility as alternative methods of CVD risk prediction along with its added advantages.

With the rise in the prevalence of hypertension globally and its associativity with other parameters of cardiovascular risks, computational techniques like feed-forward ANNs are used to model systolic blood pressure, diastolic blood pressure and pulse pressure variations with biological parameters like age, pulse rate, alcohol addiction, and physical activity level. In this aspect, ANN approaches provided more flexible and nonlinear models for prognosis and prediction of the blood pressure parameters than classical statistical algorithms [42]. Even with the increase of complex cardiovascular diseases, using machine learning models like random forest, ANNs, SVMs and Bayesian Networks to predict the in-hospital length of stay provides a positive impact on healthcare metrics [43]. Nonlinear models of unsupervised learning like clustering are commonly used in stratification of patient population and knowledge extraction from different groups. This is highly relevant in the prediction of risks because individuals with similar characteristics often present a similar risk profile [44].

The aspect of computational intelligence through nonlinear machine learning model even applies to other fields like survival prediction in transplantations and early detection of chronic diseases like cancer. Another interesting example of using computational intelligence and predictive analysis is the prediction of neurodegenerative diseases like Parkinson's disease. Disorders like Parkinson's disease and essential tremor which affect the normal movements of a person share some symptoms or manifestations that make the process of discrimination between them a difficult task. Clinical experience of the medical doctor is crucial at the moment of giving an accurate diagnosis. And still in such a case, that diagnosis is subjective and could be contaminated by several factors beyond the usual capacity of a medical personnel to analyze efficiently [45]. Especially with the use of wearable IoT-based sensors, data obtained about a patient's movement is extensive and complex. Nevertheless, it provides huge scope for using computational intelligence toward the prediction or early detection of such diseases. A major challenge in this aspect is the early detection of such disorders based on the patient's data obtained over a period of time, tracking its changes or finding patterns exhibiting similar trends of having the disease. In this case also, linear models do not count useful always since the parameters are quite dynamic and it needs the provision to continuously analyze other non-formalized parameters to find interesting traits leading to prediction. Thus, the aspect of computational intelligence is not only helpful in designing a better model for analysis but is also useful in prediction of diseases with higher accuracy.

## **4. Conclusion**

Nonlinear systems constitute an important part of the area of predictive analytics aimed at diseases and risks for people. In the new age of data and eHealth, the inherent knowledge of data has turned out to be of immense importance, which needs specific methods with computational intelligence. Especially for chronic diseases, long-term behavioral data stands quite crucial. Data modeling and predictive analytics open a huge avenue toward clinical decision support systems, which is a fundamental tool now-a-days for preventive and personalized healthcare and supports healthcare providers to have deeper insights into patients' data [46] and take clinical decisions [33]. Therefore, the use of intelligent prediction is primarily based

in two parts—modeling of the health data and analysis of the knowledge obtained. Computational intelligence and predictive analytics not only help in predicting the risks of diseases, but also supports largely in visualizing the holistic picture of health in a large population, aimed at designing more efficient and robust health-care strategies across the world. Of course, it deals with growing challenges like the complexity of health data obtained, lack of interoperable systems for extended and unified analysis, intrinsic bias of some machine learning algorithms, and the implementational difficulties. Nonetheless, the application of machine learning and nonlinear methods using computation intelligence have already demonstrated its potential in predicting health risks and diseases, and is expected to reshape the field of health analytics, early detection and prediction of diseases in a global perspective.

## Acknowledgements

The authors acknowledge the research facilities and technical assistance of the National Technological University (*Universidad Tecnológica Nacional*), Buenos Aires, Argentina.

The work is supported partly by the National Agency for Research and Innovation (*Agencia Nacional de Investigación e Innovación*) of Uruguay through its grant FSDA\_1\_2017\_1\_143653 (*Inteligencia Computacional en Salud. Creando Herramientas para Predecir la Sobrevida en Trasplante Hepático*).

Also it is supported partly by the National Interuniversity Council (CIN) of Argentina through the program '*Programa Estratégico de Formación de Recursos Humanos para la Investigación y Desarrollo (PERHID)*'.

The authors sincerely acknowledge the scientific view and insights of Dr. Josemaría Menéndez Dutra, Dr. Ofelia María Noceti Penza, and Dr. Solange Gerona Sangiovanni, from the Central Hospital of the Armed Forces (*Dirección Nacional de Sanidad de las Fuerzas Armadas*) in Montevideo, Uruguay.

## Author details

Parag Chatterjee<sup>1,2\*</sup>, Leandro J. Cymberknop<sup>1</sup> and Ricardo L. Armentano<sup>1,2</sup>

<sup>1</sup> Universidad Tecnológica Nacional, Buenos Aires, Argentina

<sup>2</sup> Universidad de la República, Uruguay

\*Address all correspondence to: [paragc@ieee.org](mailto:paragc@ieee.org)

## IntechOpen

© 2019 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Boeing G. Visual analysis of nonlinear dynamical systems: Chaos, fractals, self-similarity and the limits of prediction. *Systems*. 2016;**4**(4):37. DOI: 10.3390/systems4040037
- [2] Explained: Linear and Nonlinear Systems. MIT News [Retrieved: 2018-06-30]
- [3] Nonlinear Systems, Applied Mathematics. University of Birmingham. Available from: [www.birmingham.ac.uk](http://www.birmingham.ac.uk) [Retrieved: 2018-06-30]
- [4] Higgins JP. Nonlinear systems in medicine. *The Yale Journal of Biology and Medicine*. 2002;**75**(5-6):247-260
- [5] Wyber R, Vaillancourt S, Perry W, Mannava P, Folaranmi T, Celi LA. Big data in global health: Improving health in low- and middle-income countries. *Bulletin of the World Health Organization*. 2015;**93**:203-208. DOI: 10.2471/BLT.14.139022
- [6] Ahern DK, Kreslake JM, Phalen JM. What is ehealth (6): Perspectives on the evolution of ehealth research. *Journal of Medical Internet Research*. 2006;**8**(1):e4. DOI: 10.2196/jmir.8.1.e4
- [7] Charles D, King J Patel V, Furukawa M. Adoption of Electronic Health Record Systems among U.S. Non-Federal Acute Care Hospitals: 2008-2012; 2013. Available from: <http://www.healthit.gov/sites/default/files/oncdatabrief9final.pdf> [Accessed: 20 April 2014]
- [8] Shah NH. Translational bioinformatics embraces big data. *Yearbook of Medical Informatics*. 2012;**7**(1):130-134
- [9] Heinze O, Birkle M, Koster L, Bergh B. Architecture of a consent management suite and integration into IHE-based regional health information networks. *BMC Medical Informatics and Decision Making*. 2011;**11**:58
- [10] Tejero A, de la Torre I. Advances and current state of the security and privacy in electronic health records: Survey from a social perspective. *Journal of Medical Systems*. 2012;**36**(5):3019-3027
- [11] Mense A, Hoheiser-Pfortner F, Schmid M, Wahl H. Concepts for a standard based cross-organisational information security management system in the context of a nationwide EHR. *Studies in Health Technology and Informatics*. 2013;**192**:548-552
- [12] Faxvaag A, Johansen TS, Heimly V, Melby L, Grimsmo A. Healthcare professionals' experiences with EHR-system access control mechanisms. *Studies in Health Technology and Informatics*. 2011;**169**:601-605
- [13] Ross MK, Wei W, Ohno-Machado L. "Big data" and the electronic health record. *Yearbook of Medical Informatics*. 2014;**9**(1):97-104. DOI: 10.15265/IY-2014-0003
- [14] Waghlikar KB, Sundararajan V, Deshpande AW. Modeling paradigms for medical diagnostic decision support: A survey and future directions. *Journal of Medical Systems*. 2012;**36**(5):3029-3049
- [15] WHO. Integrated Chronic Disease Prevention and Control [Internet]. 2019. Available from: [https://www.who.int/chp/about/integrated\\_cd/en/](https://www.who.int/chp/about/integrated_cd/en/)
- [16] VanWagner LB, Ning H, Whitsett M, Levitsky J, Uttal S, Wilkins JT, et al. A point-based prediction model for cardiovascular risk in orthotopic liver transplantation: The CAR-OLT score. *Hepatology*. 2017;**66**:1968-1979. DOI: 10.1002/hep.29329

- [17] Hemmatpour M, Ferrero R, Gandino F, Montrucchio B, Rebaudengo M. Nonlinear predictive threshold model for real-time abnormal gait detection. *Journal of Healthcare Engineering*. 2018;**2018**. Article ID 4750104. 9 p. DOI: 10.1155/2018/4750104
- [18] Stanford Medicine. Stanford Medicine 2017 Health Trends Report Harnessing the Power of Data in Health [Internet]. 2017. Available from: [med.stanford.edu/content/dam/sm/sm-news/documents/StanfordMedicineHealthTrendsWhitePaper2017.pdf](http://med.stanford.edu/content/dam/sm/sm-news/documents/StanfordMedicineHealthTrendsWhitePaper2017.pdf)
- [19] Corbin, Kenneth. How CIOs Can Prepare for Healthcare ‘Data Tsunami’. CIO [Internet]. 2014. Available from: [www.cio.com/article/2860072/how-cios-can-prepare-for-healthcare-data-tsunami.html](http://www.cio.com/article/2860072/how-cios-can-prepare-for-healthcare-data-tsunami.html)
- [20] Health Data Archiver. Health Data Volumes Skyrocket, Legacy Data Archives On The Rise [Internet]. 2017. Available from: <https://www.healthdataarchiver.com/health-data-volumes-skyrocket-legacy-data-archives-rise-hie/>
- [21] Hippolyte T, Adamou M, Blaise N, Pierre C, Olivier M. Linear vs non-linear learning methods—A comparative study for forest above ground biomass, estimation from texture analysis of satellite images. *ARIMA Journal*. 2014;**18**:114-131
- [22] Hastie T et al. *The Elements of Statistical Learning*. Vol. 2. Heidelberg: Springer; 2009
- [23] Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology*. 2017;**2**(4):230-243
- [24] Shahid N, Rappon T, Berta W. Applications of artificial neural networks in health care organizational decision-making: A scoping review. *PLoS One*. 2019;**14**(2):e0212356. DOI: 10.1371/journal.pone.0212356
- [25] Deepa et al. Health care analysis using random Forest algorithm. *Journal of Chemical and Pharmaceutical Sciences*. 2017;**10**(3):1359-1361. Available from: [www.jchps.com/issues/Volume%2010\\_Issue%203/20171025\\_075052\\_0180417.pdf](http://www.jchps.com/issues/Volume%2010_Issue%203/20171025_075052_0180417.pdf)
- [26] Chatterjee P, Cymberknop L, Armentano R. IoT-based ehealth toward decision support system for CBRNE events. In: Malizia A, D’Arienzo M, editors. *Enhancing CBRNE Safety & Security: Proceedings of the SICC 2017 Conference*. Cham: Springer; 2018. DOI: 10.1007/978-3-319-91791-7\_21
- [27] Chatterjee P, Cymberknop L, Armentano R. IoT-Based Decision Support System Towards Cardiovascular Diseases. Córdoba, Argentina: SABI; 2017
- [28] Donges N. The Random Forest Algorithm. *Towards Data Science*. 2018. Available from: [towardsdatascience.com/the-random-forest-algorithm-d457d499ffcd](http://towardsdatascience.com/the-random-forest-algorithm-d457d499ffcd)
- [29] Healthcare.ai. Step by Step to K-Means Clustering. *Data Science Blog*. 2017. Available from: [healthcare.ai/step-step-k-means-clustering/](http://healthcare.ai/step-step-k-means-clustering/)
- [30] Mueller J, Massaron L. *Machine Learning for Dummies*. Hoboken: John Wiley & Sons, Inc.; 2016
- [31] Raschka S. Model evaluation, model selection, and algorithm selection in machine learning. *ArXiv*. 2018. Available from: [arxiv.org/abs/1811.12808v2](http://arxiv.org/abs/1811.12808v2)
- [32] Dinesh KG, Arumugaraj K, Santhosh KD, Mareeswari V. Prediction of cardiovascular disease using machine learning algorithms. In: *International Conference on Current Trends towards*

- Converging Technologies (ICCTCT). 2018. DOI: 10.1109/icctct.2018.8550857
- [33] Chatterjee P, Armentano RL, Cymberknop LJ. Internet of things and decision support system for eHealth-applied to cardiometabolic diseases. In: International Conference on Machine Learning and Data Science (MLDS); Noida. Piscataway: IEEE; 2017. pp. 75-79. DOI: 10.1109/MLDS.2017.22
- [34] Framingham Heart Study. The Adult Treatment Panel III, JAMA. 2001 [Internet]. 2019. Available from: <https://www.framinghamheartstudy.org/fhs-risk-functions/hard-coronary-heart-disease-10-year-risk/>
- [35] Chironi G, Simon A, Megnien JL, Sirieix ME, Mousseaux E, Pessana F, et al. Impact of coronary artery calcium on cardiovascular risk categorization and lipid-lowering drug eligibility in asymptomatic hypercholesterolemic men. *International Journal of Cardiology*. 2011;**151**(2):200-204. DOI: 10.1016/j.ijcard.2010.05.024
- [36] Pessana F, Armentano R, Chironi G, Megnien JL, Mousseaux E, Simon A. Subclinical atherosclerosis modeling: Integration of coronary artery calcium score to Framingham equation. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2009. DOI: 10.1109/iembs.2009.5334049
- [37] Bucci CM, Legnani WE, Armentano RL. Clustering of cardiovascular risk factors highlighted the coronary artery calcium as a strong clinical discriminator. *Health and Technology*. 2016;**6**(3):159-165. DOI: 10.1007/s12553-016-0139-1
- [38] Obermeyer Z, Emanuel EJ. Predicting the future—Big data, machine learning, and clinical medicine. *The New England Journal of Medicine*. 2016;**375**(13):1216-1219. DOI: 10.1056/NEJMp1606181
- [39] Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS One*. 2017;**12**(4):e0174944. DOI: 10.1371/journal.pone.0174944
- [40] Panagiotakos D. Health measurement scales: Methodological issues. *Open Cardiovascular Medicine Journal*. 2009;**3**:160
- [41] Dimopoulos AC et al. Machine learning methodologies versus cardiovascular risk scores, in predicting disease risk. *BMC Medical Research Methodology*. 2018;**18**(1):1-11. DOI: 10.1186/s12874-018-0644-1
- [42] Bhaduri A, Bhaduri A, Bhaduri A, Mohapatra PK. Blood pressure modeling using statistical and computational intelligence approaches. In: IEEE International Advance Computing Conference. 2009. DOI: 10.1109/iadcc.2009.4809156
- [43] Daghistani TA, Elshawy R, Sakr S, Ahmed AM, Al-Thwayee A, Al-Mallah MH. Predictors of in-hospital length of stay among cardiac patients: A machine learning approach. *International Journal of Cardiology*. 2019;**288**:140-147. DOI: 10.1016/j.ijcard.2019.01.046
- [44] Paredes S, Henriques J, Rochar T, Mendes D, Carvalho P, Morais J, et al. A clinical interpretable approach applied to cardiovascular risk assessment. In: 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2018. DOI: 10.1109/embc.2018.8512956
- [45] Romero L, Chatterjee P, Armentano R. An IoT approach for integration of computational intelligence and wearable sensors for Parkinson's disease diagnosis and monitoring. *Health and Technology*. 2016;**6**(3):167-172. DOI: 10.1007/s12553-016-0148-0

[46] Chatterjee P, Cymberknop LJ, Armentano RL. IoT-based decision support system for intelligent healthcare—Applied to cardiovascular diseases. In: 7th International Conference on Communication Systems and Network Technologies (CSNT); Nagpur. 2017. pp. 362-366. DOI: 10.1109/CSNT.2017.8418567