

## Chapter 4

# Big Data Analysis for Cardiovascular Diseases: Detection, Prevention, and Management

**Miguel A. Sánchez-Acevedo**

*Universidad de la Cañada, Mexico & Universidad Popular Autónoma del Estado de Puebla, Mexico*

**Zaydi A. Acosta-Chí**

*Universidad de la Cañada, Mexico*

**Beatriz A. Sabino-Moxo**

*Universidad de la Cañada, Mexico*

**José A. Márquez-Domínguez**

*Universidad de la Cañada, Mexico*

**Rosa M. Canton-Croda**

*Universidad Popular Autónoma del Estado de Puebla, Mexico*

### ABSTRACT

*In the healthcare field, plenty of clinical data is generated every day from patient records, surveys, research papers, medical devices, among others sources. These data can be exploited to discover new insights about health issues. For helping decision makers and healthcare data managers, a survey of research works and tools covering the process of handling big data in the healthcare field is included. A methodology for CVD prevention, detection and management through the use of tools for big data analysis is proposed. Also, it is important to maintain privacy of patients when handling healthcare data; therefore, a list of recommendations for maintaining privacy when handling healthcare data is presented. Specific clinical analysis are recommended on those regions where the incidence rate of CVD is high, but a weak relation with the common risk factors is observed according to historical data. Finally, challenges which need to be addressed are presented.*

DOI: 10.4018/978-1-5225-8185-7.ch004

## **INTRODUCTION**

Big Data term could be relative to the context, but quantities starting from petabytes are always considered as Big Data. Every second, a lot of data is generated from several sources: traditional databases, text files, remote data, mobile devices, sensors, etc. The way in which data is presented can be classified as structured, semi-structured, and unstructured data. In order to analyze the collected data, it is required to prepare it; starting with an exploration, followed by description, and then visualizing the data with graphs. Good results in the analysis phase can be guaranteed with quality of data. Quality of data is achieved by removing data with missing values, merging duplicate records, generating best estimates for invalid values, and removing outliers. After data preparation, a set of analysis techniques can be applied to discover relevant information; most used techniques are: classification, clustering, regression, graph analytics, and association analysis. All the stages involved in big data processing allow the generation of new insights that could be used to take actions for solving the problem under study.

According to the World Health Organization (WHO, 2016), cardiovascular disease (CVD) is the main cause of death in the world; the organization estimates that more than 17.5 million people died of CVD in 2012, and this number has increased in last years. Coronary heart disease, cerebrovascular disease, peripheral arterial disease, rheumatic heart disease, congenital heart disease, deep vein thrombosis, pulmonary embolism, heart attacks, and strokes are the most common CVDs, while coronary heart disease is the leading cause of premature death. The probability of suffering a coronary heart disease is increased by risk factors like unhealthy diet, physical inactivity, tobacco use, and harmful use of alcohol; also, those risk factors alter normal blood conditions of pressure, glucose, and lipids, and generate overweight and obesity. Solutions to this problematic seems easy: not smoking, eat healthily, and exercise; however, people is not motivated to perform those activities and it is necessary to identify the reasons of this behavior in order to establish public politics that contribute to reduce CVD. There are regions in the world where CVD has increased, but a lack of resources in healthcare institutions reduce the possibility of identifying the roots of the problem; nevertheless, some risk factors are easily observable (poverty, stress, age, gender, overweight, obesity, tobacco use, etc.) and can guide to a more deeper analysis in regions where insufficient information is available; on the other hand, health issues with known risk factors can be addressed with proposed solutions in regions with similar characteristics.

Nowadays, it is possible to access data related to CVD in several databases around the world, National Cardiovascular Disease Database, Coronary Artery Disease Gene Database, Ensanut, Centers for Disease Control and Prevention, and ONEdat, to name a few. With those data available, tools for big data analysis can be employed for discovering new correlations among data that could not be observed before. For those countries where a lack of resources limits the number of clinical analysis that can be performed in the population, big data analysis can contribute to identify common characteristics with people in other regions of the world and select more pertinent analysis to be performed. This chapter proposes a methodology for using healthcare databases to detect CVD risk factors, determine the high profitable actions for preventing CVD based on risks detected, and identify what are the best clinical analysis and drugs for treatment of CVD according to historical success cases. Although great advances have been achieved in big data field, there are issues that have not been addressed yet. A CVD can be originated by several factors and not always presenting the same symptoms, making its detection more difficult; therefore, challenges to be addressed are presented.

16 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/big-data-analysis-for-cardiovascular-diseases/225355](http://www.igi-global.com/chapter/big-data-analysis-for-cardiovascular-diseases/225355)

## Related Content

---

### A Review on IoT-Driven Technologies for Heart Disease Diagnosis and Prediction

Makarand Shahade and Mangesh M. Ghonge (2022). *Leveraging AI Technologies for Preventing and Detecting Sudden Cardiac Arrest and Death* (pp. 21-34).

[www.irma-international.org/chapter/a-review-on-iot-driven-technologies-for-heart-disease-diagnosis-and-prediction/308834](http://www.irma-international.org/chapter/a-review-on-iot-driven-technologies-for-heart-disease-diagnosis-and-prediction/308834)

### PPG-Based Cardiovascular Disease Predictor Using Artificial Intelligence

Dhanalakshmi S., Gayathiridevi B., Kiruthika S. and E Smily Jeya Jothi (2022). *Leveraging AI Technologies for Preventing and Detecting Sudden Cardiac Arrest and Death* (pp. 218-239).

[www.irma-international.org/chapter/ppg-based-cardiovascular-disease-predictor-using-artificial-intelligence/308842](http://www.irma-international.org/chapter/ppg-based-cardiovascular-disease-predictor-using-artificial-intelligence/308842)

### Palliative Care

May Hua (2015). *Modern Concepts and Practices in Cardiothoracic Critical Care* (pp. 105-123).

[www.irma-international.org/chapter/palliative-care/136905](http://www.irma-international.org/chapter/palliative-care/136905)

### Complementary and Alternative Medicine Use in Hypertension: The Good, the Bad, and the Ugly: Hypertension Treatment From Nature – Myth or Fact?

Aymen Shatnawi, Alison Shafer, Hytham Ahmed and Fawzy Elbarbry (2017). *Emerging Applications, Perspectives, and Discoveries in Cardiovascular Research* (pp. 255-287).

[www.irma-international.org/chapter/complementary-and-alternative-medicine-use-in-hypertension/176223](http://www.irma-international.org/chapter/complementary-and-alternative-medicine-use-in-hypertension/176223)

### Survey on Cardiac Rehabilitation

Rabindranath Sahu (2023). *Cognitive Cardiac Rehabilitation Using IoT and AI Tools* (pp. 113-121).

[www.irma-international.org/chapter/survey-on-cardiac-rehabilitation/325527](http://www.irma-international.org/chapter/survey-on-cardiac-rehabilitation/325527)