International Journal of Big Data and Analytics in Healthcare
Volume 5 • Issue 2 • July-December 2020

Cardiovascular Risk Detection Through Big Data Analysis

Miguel A. Sánchez-Acevedo, Universidad de la Cañada, Mexico & Universidad Popular Autónoma del Estado de Puebla, Mexico
https://orcid.org/0000-0002-0996-0038

Zaydi Anaí Acosta-Chi, Universidad de la Cañada, Mexico

Ma. del Rocío Morales-Salgado, Universidad Popular Autónoma del Estado de Puebla, Mexico

ABSTRACT

Cardiovascular diseases are the main cause of mortality in the world. As more people suffer from diabetes and hypertension, the risk of cardiovascular disease (CVD) increases. A sedentary lifestyle, an unhealthy diet, and stressful activities are behaviors that can be changed to prevent CVD. Taking measures to prevent CVD lowers the cost of treatments and reduces mortality. Data-driven plans generate more effective results and can be applied to groups with similar characteristics. Currently, there are several databases that can be used to extract information in real time and improve decision making. This article proposes a methodology for the detection of CVD and a web tool to analyze the data more effectively. The methodology for extracting, describing, and visualizing data from a state-level case study of CVD in Mexico is presented. The data is obtained from the databases of the National Institute of Statistics and Geography (INEGI) and the National Survey of Health and Nutrition (ENSANUT). A k-nearest neighbor (KNN) algorithm is proposed to predict missing data.

KEYWORDS

Big Data, Cardiovascular Disease, Data Science, KNN, Machine Learning, Pool Cohort Equation, Structured Data, Web Tool

INTRODUCTION

According to the World Health Organization (WHO, 2018), ischaemic heart disease and stroke, are the main cause of death in the world; the organization estimates that 15.2 million people died due to those diseases in 2016. The main factors of cardiovascular disease (CVD) are hypertension, dyslipidemia, abnormal glucose levels and diabetes (Anstiss & Passmore, 2020). Despite its high mortality, CVD can be prevented through behavioral changes as physical activity, healthy diet, limiting tobacco exposition, restricting alcohol consumption, and reducing stress (Hooker, 2013). Changing the lifestyle of people is a big challenge due to the increase in sedentary work and the consumption of high calorie foods, so dedicated attention and training for CVD prevention is required (Saeed et al., 2018). Plans and actions to reduce the risk of CVD can be improved by analyzing the data generated every day from the treatment of people with CVD cases, including their origin and evolution.

DOI: 10.4018/IJBDAH.2020070101

Multiple databases are fed with CVD data every day around the world. Centers for Disease Control and Prevention (CDC, 2020), National Survey for Health and Nutrition (ENSANUT, 2020), National Cardiovascular Disease Database (NCVD, 2020), and World Health Organization (WHO, 2020) are some organizations that keep databases available online. Data stored in public databases can be supplemented with patient records, clinical analyzes, diagnoses, and data collected from mobile devices that people use to track their treatments. New correlations between data can be discovered through big data analysis tools. Big Data analysis can help identify common characteristics in people with CVD, and those results could help select actions of greater impact that could be carried out in countries where the lack of resources restricts the amount of clinical analysis that can be performed in the population.

The term Big Data was coined by Gartner Group in 2008 (Emmanuel & Stanier, 2016) for referring to the large amount (volume) of data that is generated continuously (velocity) from several sources (variety) every day (Edward & Sabharwal, 2015). To analyze the collected data, an exploration is performed, followed by a description, and then visualized to interactively explore the content; in this way, users can identify patterns and infer correlations to support decisions (Bikakis, Papastefanatos, & Papaemmanouil, 2019). Good results in the analysis phase can be guaranteed with quality of data. Quality is ensured by evaluating the accuracy, completeness, consistency, distinctness, precision, timeliness and volume of data (Cappiello, Samá, & Vitali, 2018) . After data preparation, a set of analysis techniques can be applied to discover relevant information; most used techniques are: pattern matching, classification, clustering, and regression (Mogha, Ahlawat, & Singh, 2018). 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 Hong et al. (2018) Big Data in healthcare can be divided in four categories: medicine and clinics, public health and behavior, medical experiment, and medical literature. In the first class the data is extracted from electronic health records (Hoffman, 2016), electronic medical records (Setiawan, Utami, Mengko, & Indrayanto, 2014), personal health records (Okore, Bakyarani, & Phil, 2015) and medical images (Tahmassebi et al., 2019). Data of second class is collected from electrocardiograms (Cipresso, Rundo, Conoci, & Parenti, 2019) and vital signs (Mohammad Forkan, Khalil, & Atiquzzaman, 2017). The third class obtains data from molecular biology (Cannataro, 2019), human body samples, and clinical trials (Mayo et al., 2017). Finally, medical literature is obtained from structured knowledge, journal and conference articles (Wang et al., 2018).

This work proposes a methodology for CVD risk detection at state-level making use of a web tool for processing data obtained from public databases. The paper is organized as follows: first, the related work is presented; then, a methodology for the detection of CVD through a big data tool is described; the methodology to extract, describe, and visualize data from a CVD case study at the state level in Mexico is presented below; then, a K-Nearest Neighbor (KNN) algorithm is proposed for the classification of missing data; finally, future work and challenges to be addressed are presented.

BACKGROUND

Alshraideh et al. (2014) developed an algorithm to detect arrhythmias in individuals by identifying 18 attributes of personal information and electrocardiogram (ECG) signals. The system is made up of an Android application that obtains the ECG signal wirelessly and sends the information to a web server where the algorithm is stored. If CVD is detected, an SMS is sent to the user to view their information through a web interface. Roses et al. (2017) conducted a review to identify the effectiveness of positron emission tomography supplemented with computed tomography (PET/CT) for the detection of CVD. PET/CT is used to determine the presence of ischemia in myocardial tissue. They conclude that there is currently a limitation of the indicators, especially in asymptomatic patients, which is why new protocols are required to improve the effectiveness of PET/CT for the detection of CVD. Hernesniemi et al. (2018) created a database to integrate data from hospital records and national registries, disease phenotype, details of invasive operations, patient clinical information, biometric data physiological, anatomical measurements, and laboratory results obtained from 73,000 individuals treated in TAYS Heart Hospital; the next step will be to analyze the data with a convolutional neural network. Pfohl et al. (2019) developed an adversarial network to consider under-represented minorities in the risk model of atherosclerothic cardiovascular disease (ASCVD) by considering race, gender, and age. Manonmani and Balakrishnan (2020) proposed a semantic annotation process on healthcare data collected from sensors, handheld devices, and datasets to establish relationships between the characteristics identified in the data through fuzzy rule mining. Once the data is annotated, classification algorithms can be used to determine whether or not a patient has a specific disease. Leone et al. (2020) identified that the risk of CVD in patients with systemic lupus erythematosus (SLE) is underestimated using standard risk calculators; For this reason, they propose the use of echocardiography as a non-invasive detection tool to accurately assess the risk of CVD in patients with SLE.

Detecting CVD risk is a process which requires to analyze several factors on patients or healthy persons. The risks assessment can be performed through the identification of risk factors, imaging methods or methods like fibrinogen, homocysteine, and lipoprotein-associated phospholipase 2. According to the European Guidelines on cardiovascular disease prevention in clinical practice (Piepoli et al., 2016), for patients 20 to 79 years of age who are free from clinical atherosclertic cardiovascualar disease (ASCVD), the first step is to assess ASCVD risk factors: Age, gender, race/ethnicity, current smoker (yes/no), any family history of heart attack in first degree relative (parent/sibling/child) (yes/no), diabetes, total and HDL cholesterol, use of lipid lowering medication (yes/no), systolic blood pressure (mmHg), and use of antihypertensive medication. For patients outside this age range, providers should refer to applicable clinical practice guidelines and adult primary prevention guidelines.

The assessment of cardiovascular risk is not necessary for specific subgroups of asymptomatic individuals at unusually high risk, such as those with genetically determined extreme values of traditional risk factors, hypercholesterolemia, hypertension and others diseases like diabetes, chronic kidney disease, obstructive sleep apnoea, and erectile dysfunction (Arnett et al., 2019). These parameters are presented below:

1. Diabetes and age > 60 years;
2. Diabetes with microalbuminuria (>20 mcg/min or urinary albumin:creatinine ratio [UACR] >2.5 mg/mmol for males, >3.5 mg/mmol for females);
3. Moderate or severe chronic kidney disease (CKD) (persistent proteinuria or estimated glomerular filtration rate [eGFR] <45 mL/min/1.73 m2);
4. A previous diagnosis of familial hypercholesterolaemia;
5. Systolic blood pressure (SBP) ≥180 mmHg or diastolic blood pressure (DBP) ≥110 mmHg;
6. Serum total cholesterol (TC) >7.5 mmol/L.

Others methods for diagnosis that are used as a complement to the risk assessment are:

1. **Imaging Studies:** Measurement of carotid intima-media thickness (Ravani et al., 2015) and/or screening for atherosclerotic plaques by carotid artery scanning (Romanens, Sudano, Adams & Schober, 2019), measurement of ankle–brachial index (Bendermacher et al., 2012), computed tomography for coronary calcium (AlJaroudi et al., 2019), and exercise electrocardiography with microalbuminuria (Al-Saffar, Nassir, Mitchell & Philipp, 2015);
2. **Measurement of Inflammatory and Thrombotic Biomarkers:** High-sensitivity C-reactive protein (CRP) and homocysteine (Cheng, Tu, Shen, Zhang & Ji, 2018), fibrinogen (Kaur, Singh, Indu, & Singla, 2012), homocysteine and lipoprotein-associated phospholipase 2 (Lp-PLA$_2$) (Nicholls, 2018);

3. **Genetic Testing:** Deoxyribonucleic acid (DNA)-based tests for common genetic polymorphisms (Dinter et al., 2019);
4. **Assessment of Psychosocial Risk Factors:** Low socioeconomic status, lack of social support, stress at work and in family life, depression, anxiety, hostility, and the type D personality (Sood & Gidding, 2016).

The application of the methods described above will depend of the level of risk:

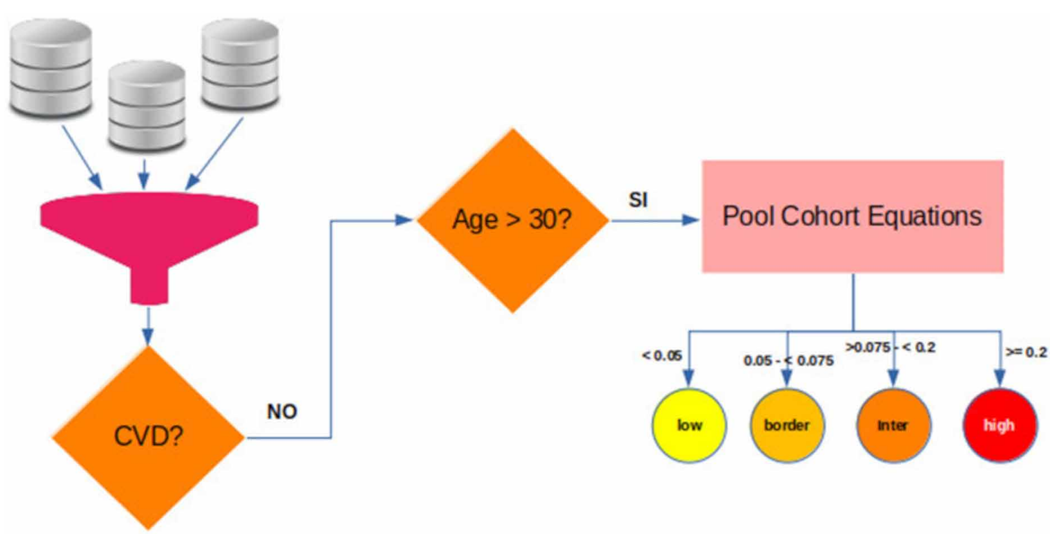
- **Moderate Risk:** High-sensitivity-C-reactive protein (CPR) and homocysteine (Deebukkhum, Pingmuangkaew, Tangvarasittichai & Tangvarasittichai, 2012), measurement of ankle–brachial index (Nishimura et al., 2017), computed tomography for coronary calcium (AlJaroudi et al., 2019), and exercise electrocardiography (Van de Sande, Breuer & Kemps, 2016);
- **High Risk of a Recurrent Acute Atherothrombotic Event:** Lipoprotein-associated phospholipase 2 (Lp-PLA_2) (Rallidis et al., 2012).

The genetic testing is not recommended, the testing do not add significantly to diagnosis, risk prediction, or patient management but the psychosocial risk factors can be assessed through data mining to improve treatment adherence.

METHODOLOGY FOR CARDIOVASCULAR RISK DETECTION (MCRD)

There exist several databases where individual characteristics are stored and can be used to perform and automated analysis in order to identify accurately CVD risk levels. For example, the Dartnet Institute provides a dataset containing records of blood pressure, LDL, Total Cholesterol, HDL, among others, obtained from patients of several institutions in the United States. In the United Kingdom the same data can be obtained from the Consumer Data Research Centre (CDRC). In Europe data is collected by the European Genome-phenome Archive and the European Society of Cardiology. Others datasets can be obtained from diverse institutions around the world, which can be used for identifying CVD risks levels by regions following the next methodology (Figure 1):

Figure 1. CVD detection from big data



1. **Datasets Merging:** Collect all data available which include most factors as possible of those described above. A more accurate assessment is obtained when multiple risk factors are analyzed. More data available increases accuracy;
2. **Identify People With Established CVD and Exclude:** When data is collected from several sources, data about persons which are currently under CVD treatment could be included. So, exclude them and continue working with the remaining registries;
3. **Age classification:** Separate men and women older than 30. Include those registries in the high level risk and continue working with the remaining registries;
4. **Risk Assessment and Risk Stratification:** For classifying the remaining registries apply the Pooled Cohort Equations (Yadlowsky et al., 2018) and classify according to the Table 1.

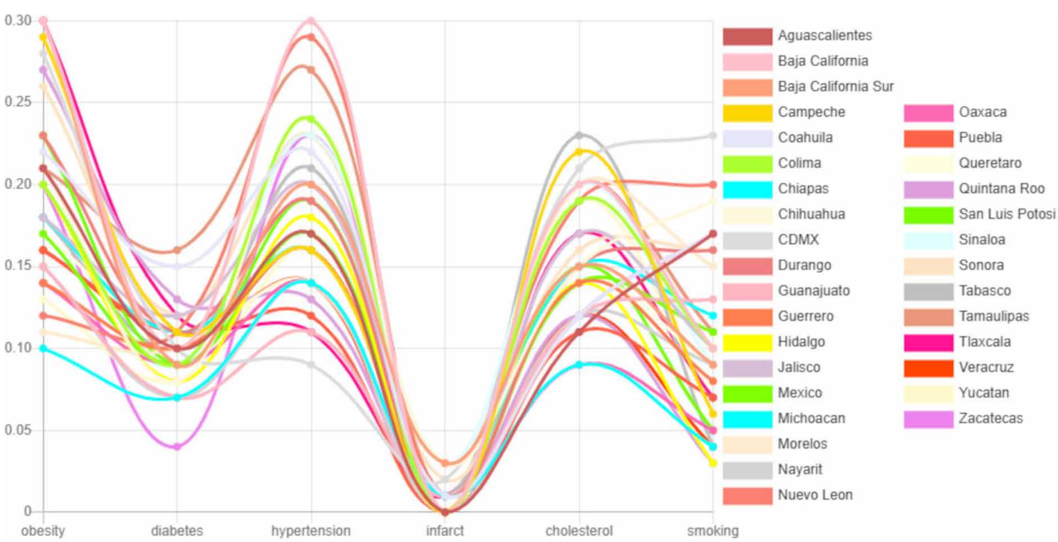
Table 1. Absolute cardiovascular risk (Arnett et al., 2019)

Categories	Score
Low Risk	< 5%
Borderline Risk	5% - < 7.5%
Intermediate Risk	>= 7.5% - < 20%
High risk	>=20%

RESULTS AT STATE LEVEL IN MEXICO

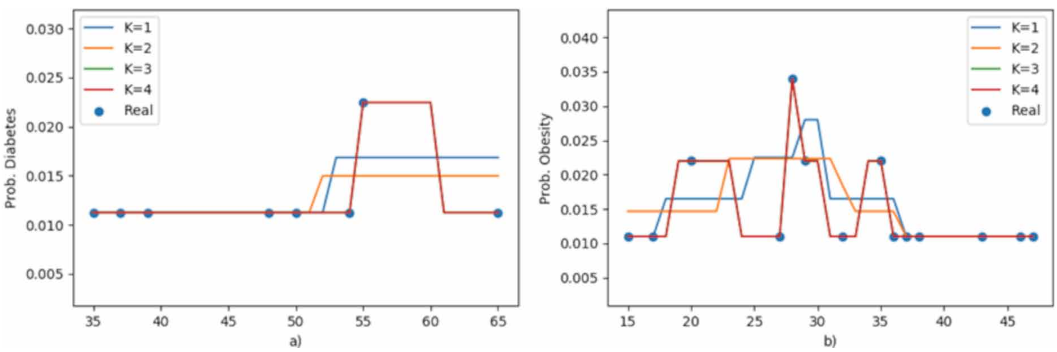
A web tool was designed to evaluate the proposed methodology. The first version of the tool allows the analysis of information at the state level. First, data is collected from public databases of ENSANUT and INEGI in Mexico; then, the variables to consider (obesity, diabetes, hypertension, infarct, cholesterol, and smoking) are extracted from the data (Figure 2).

Figure 2. Data collected from ENSANUT and INEGI in 2018



Once the variables are extracted from the databases, a classification algorithm (k-closest neighbors) is trained to fill in the missing values from the databases for each variable and each state. Figure 3a illustrates the probability of diabetes in the state of Aguascalientes, Mexico; the data available in the databases is represented by blue dots and the predictions obtained after training the algorithm with 1, 2, 3 and 4 neighbors are illustrated with different colored lines. It is possible to observe that good results are obtained with 3 neighbors. The precision of the trained set is 0.24. For better accuracy more data is required. These data can be complemented with the new databases published by the National Institutes in the years in which the population evaluation is carried out. Figure 3b illustrates the results on the obesity variable obtaining a precision of 0.30.

Figure 3. a) KNN Prediction Algorithm for Diabetes in Aguascalientes, Mexico; b) KNN Prediction Algorithm for Obesity in Aguascalientes, Mexico



After completing all the data with the KNN algorithm for each state, the grouping cohort equation (Arnett et al., 2019) is applied to obtain the risk level category. Percentage of population older than30 years (Figure 4), with obesity (Figure 5), and diabetes (Figure 6) are visualized in an interactive map.

CONCLUSION

The availability of more data sources opens up the possibility of finding patterns that improve decision-making; however, analyzing all the data generated is a challenging activity. Big Data tools help process large amounts of data and discover patterns that can be visualized through interactive interfaces that allow the analysis of information. In this work, a state-level analysis of CVD in Mexico was carried out through a web tool following the proposed methodology. This information can be used to establish policies that encourage people to change their habits to reduce the risk of CVD. As future work, the application will be improved to allow the analysis of information at the municipal level and to visualize the behavior of several years in real time. To make better predictions more data have to be collected and the machine learning algorithm improved.

Figure 4. Population older than 30 years

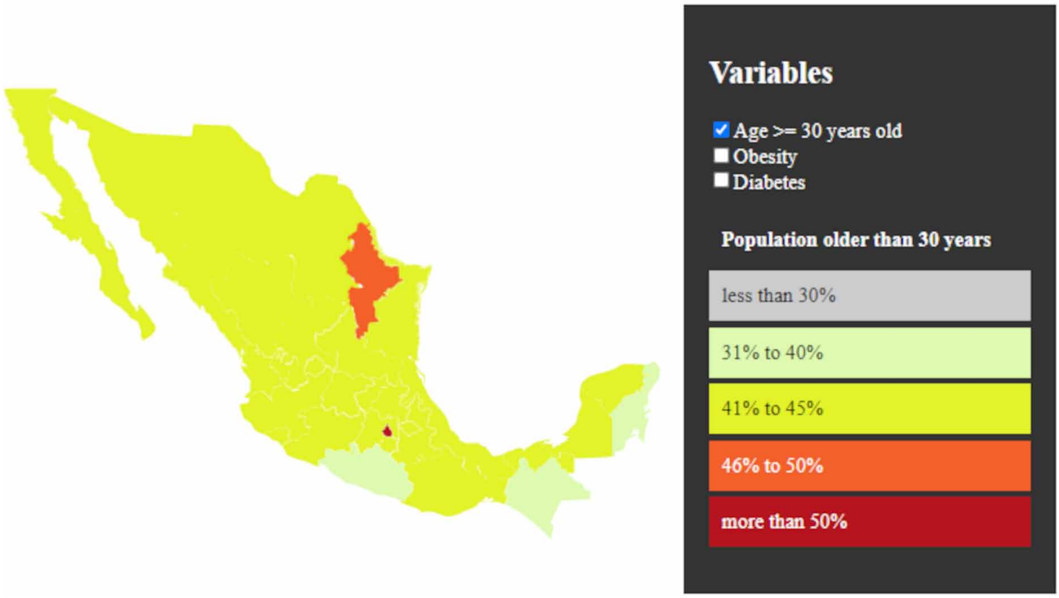


Figure 5. Percentage of the sample with obesity

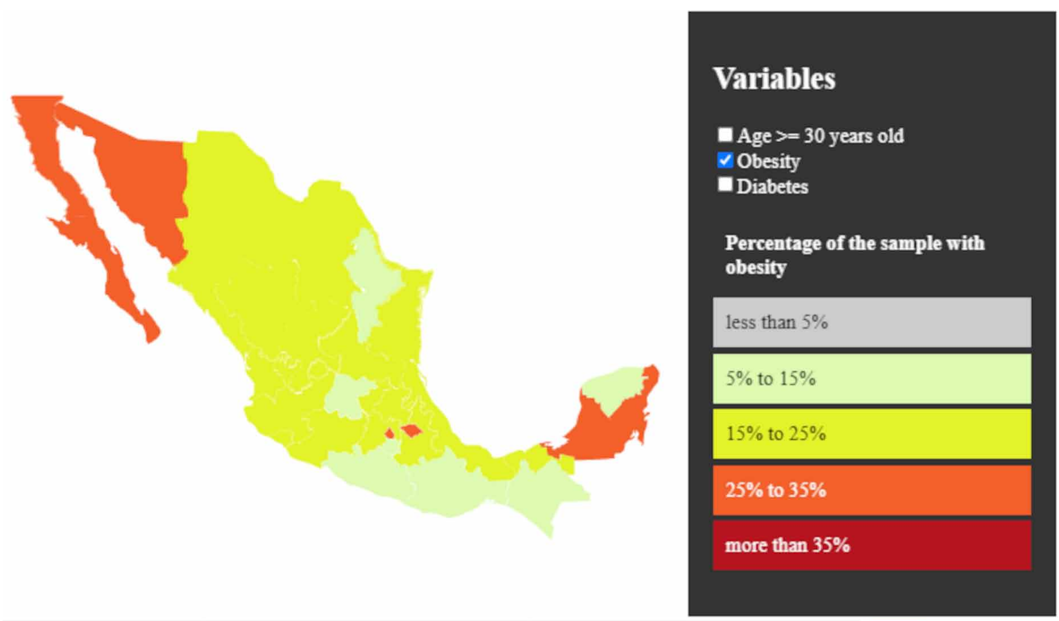
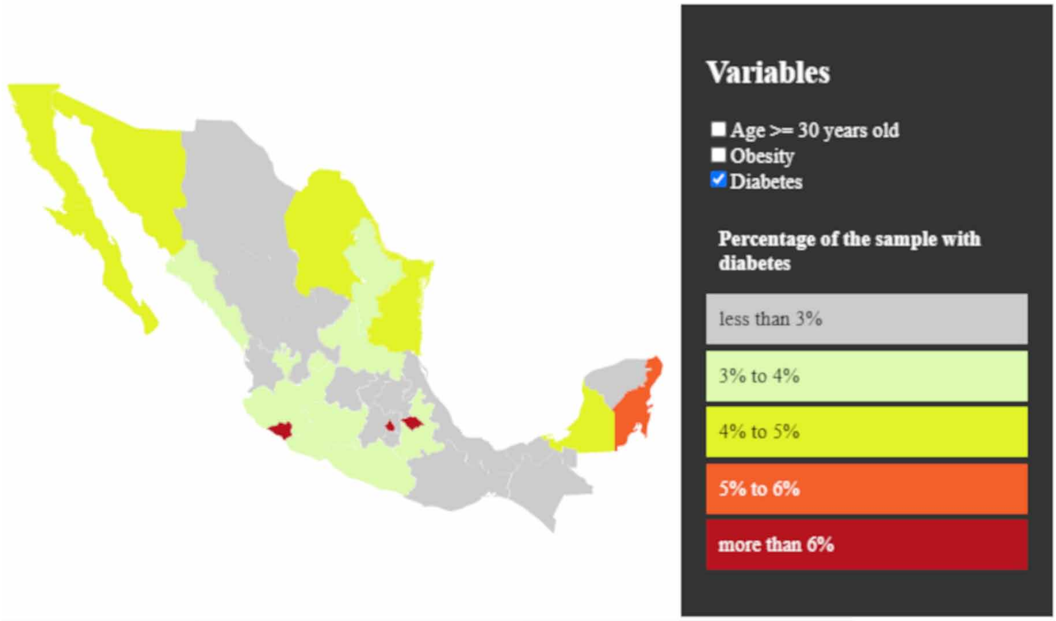


Figure 6. Percentage of the sample with diabetes



REFERENCES

AHA. (2004). *Statistical fact sheet: International cardiovascular disease statistics*. American Heart Association. Retrieved from http://www.sld.cu/galerias/pdf/servicios/hta/international_cardiovascular _disease_statistics.pdf

Al-Saffar, H. B., Nassir, H., Mitchell, A., & Philipp, S. (2015). Microalbuminuria in non-diabetic patients with unstable angina/non ST-segment elevation myocardial infarction. *BMC Research Notes*, *8*(1), 1–7. doi:10.1186/s13104-015-1347-x PMID:26362770

AlJaroudi, W., Mansour, M. J., Chedid, M., Hamoui, O., Asmar, J., Mansour, L., & Chammas, E. (2019). Incremental value of stress echocardiography and computed tomography coronary calcium scoring for the diagnosis of coronary artery disease. *The International Journal of Cardiovascular Imaging*, *35*(6), 1133–1139. doi:10.1007/s10554-019-01577-x PMID:30874979

Alshraideh, H., Otoom, M., Al-Araida, A., Bawaneh, H., & Bravo, J. A web based cardiovascular disease detection system. In R. Hervás, S. Lee, C. Nugent, & J. Bravo (Eds.), *Ubiquitous Computing and Ambient Intelligence: Personalisation and User Adapted Services (LNCS)* (Vol. 8867, pp. 243–250). Springer. doi:10.1007/978-3-319-13102-3_40

Anstiss, T., & Passmore, J. (2020). Health and wellbeing coaching. In W. Leal Philo, T. Wall, A. M. Azul, L. Brandli, & P. Gokcin Ozuyar (Eds.), *Good Health and Well-Being* (pp. 266–277). Springer. doi:10.1007/978-3-319-95681-7_100

Arnett, D. K., Blumenthal, R. S., Albert, M. A., Buroker, A. B., Goldberger, Z. D., Hahn, E. J., & Ziaeian, B. et al. (2019). *ACC/AHA guideline on the primary prevention of cardiovascular disease*. American College of Cardiology/American Heart Association.

Bendermacher, B. L. W., Teijink, J. A. W., Willigendael, E. M., Bartelink, M. L., Peters, R. J. G., Langenberg, M., Büller, H. R., & Prins, M. H. (2012). Applicability of the ankle-brachial-index measurement as screening device for high cardiovascular risk: An observational study. *BMC Cardiovascular Disorders*, *12*(59), 1–7. doi:10.1186/1471-2261-12-59 PMID:22846150

Bikakis, N., Papastefanatos, G., & Papaemmanouil, O. (2019). Big data exploration, visualization and analytics. *Big Data Research*, *18*, 1. doi:10.1016/j.bdr.2019.100123

Cannataro, M. (2019). Big data analysis in bioinformatics. In S. Sakr & A. Y. Zomaya (Eds.), *Encyclopedia of Big Data Technologies* (pp. 161–180). Springer. doi:10.1007/978-3-319-77525-8_139

Cappiello, C., Samá, W., & Vitali, M. (2018). Quality awareness for a successful big data exploitation. In *Proceedings of the 22nd International Database Engineering & Applications Symposium (IDEAS '18)*. Villa San Giovanni, Italy: ACM doi:10.1145/3216122.3216124

CDC. (2020, Mar 27). *Centers for Disease Control and Prevention*. Retrieved from https://www.cdc.gov/

Cheng, L.-S., Tu, W.-J., Shen, Y., Zhang, L.-J., & Ji, K. (2018). Combination of high-sensitivity C-reactive protein and homocysteine predicts the post-stroke depression in patients with ischemic stroke. *Molecular Neurobiology*, *55*(4), 2952–2958. doi:10.1007/s12035-017-0549-8 PMID:28456936

Cipresso, C., Rundo, F., Conoci, S., & Parenti, R. (2019). Big data in preclinical ECG alterations research. *Biomedical Journal of Scientific & Technical Research*, *13*(2), 9895–9896. doi:10.26717/BJSTR.2019.13.002384

Deebukkhum, S., Pingmuangkaew, P., Tangvarasittichai, O., & Tangvarasittichai, S. (2012). Estimated creatinine clearance, homocysteine and high sensitivity-c-reactive protein levels determination for early prediction of nephropathy and atherosclerosis risk in type 2 diabetic patients. *Indian Journal of Clinical Biochemistry*, *27*(3), 239–245. doi:10.1007/s12291-012-0192-0 PMID:26405381

Dinter, F., Burdukiewicz, M., Schierack, P., Lehmann, W., Nestler, J., Dame, G., & Rödiger, S. (2019). Simultaneous detection and quantification of DNA and protein biomarkers in spectrum of cardiovascular diseases in a microfluidic microbead chip. *Analytical and Bioanalytical Chemistry*, *411*(29), 7725–7735. doi:10.1007/s00216-019-02199-x PMID:31760445

Edward, S. G., & Sabharwal, N. (2015). *Practical MongoDB; Architecting, Developing, and Administering MongoDB*. Apress.

Emmanuel, I., & Stanier, C. (2016). Defining big data. In *Proceedings of the International Conference on Big Data and Advanced Wireless Technologies (BDAW '16)*. Blagoevgrad, Bulgaria: ACM. doi:10.1145/3010089.3010090

ENSANUT. (2020, Mar 27). *National Survey for Health and Nutrition*. Retrieved from https://ensanut.insp.mx/

Hernesniemi, J. A., Mahdiani, S., Lyytikäinen, L. P., Lehtimäki, T., Eskola, M., Nikus, K., . . . Oksala, N. (2018). Cohort description for MADDEC: Mass data in detection and prevention of serious adverse events in cardiovascular disease. In *Proceedings of the International Federation for Medical and Biological Engineering: EMBEC & NBC 2017* (Vol. 65, pp. 1113-1116). Singapore: Springer.

Hoffman, S. (2016). *Electronic Health Records and Medical Big Data*. Cambridge University Press. doi:10.1017/9781316711149

Hong, L., Luo, M., Wang, R., Lu, P., Lu, W., & Lu, L. Big data in healthcare: applications and challenges. *Data and Information Management, 2*(3), 1-23.

Hooker, S. A. (2013). Cardiovascular disease prevention. In M. D. Gellman & J. R. Turner (Eds.), *Encyclopedia of Behavioral Medicine* (pp. 337–340). Springer. doi:10.1007/978-1-4419-1005-9_1611

Kaur, S., Singh, P., Indu, V., & Singla, G. (2012). Fibrinogen, Lp(a), microalbuminuria and left ventricular mass index: Cardiovascular disease risk factors in diabetes. *Indian Journal of Clinical Biochemistry*, *27*(1), 94–96. doi:10.1007/s12291-011-0184-5 PMID:23277719

Leone, P., Cicco, S., Prete, M., Solimando, A. G., Susca, N., Crudele, L., Buonavoglia, A., Colonna, P., Dammacco, F., Vacca, A., & Racanelli, V. (2020). Early echocardiographic detection of left ventricular diastolic dysfunction in patients with systemic lupus erythematosus asymptomatic for cardiovascular disease. *Clinical and Experimental Medicine*, *20*(1), 11–19. doi:10.1007/s10238-019-00600-8 PMID:31848778

Manonmani, M., & Balakrishnan, S. (2020). A review of semantic annotation models for analysis of healthcare data based on data mining techniques. In P. Venkata Krishna & M. Obaidat (Eds.), *Emerging Research in Data Engineering Systems and Computer Communications: Advances in Intelligent Systems and Computing* (Vol. 1054, pp. 231–238). Springer. doi:10.1007/978-981-15-0135-7_22

Mayo, C. S., Matuszak, M. M., Schipper, M. J., Jolly, S., Hayman, J. A., & Ten Haken, R. K. (2017). Big data in designing clinical trials: Opportunities and challenges. *Frontiers in Oncology*, *7*(187), 1–7. doi:10.3389/fonc.2017.00187 PMID:28913177

Mogha, G., Ahlawat, K., & Singh, A. M. (2018). Performance analysis of machine learning techniques on big data using apache spark. Data Science and Analytics. In B. Panda, S. Sharma, & N. Roy (Eds.), *Data Science and Analytics: Communications in Computer and Information Science* (Vol. 799, pp. 17–26). Springer. doi:10.1007/978-981-10-8527-7_2

Mohammad Forkan, A. R., Khalil, I., & Atiquzzaman, M. (2017). Visibid: A learning model for early discovery and real-time prediction of severe clinical events using vital signs as big data. *Computer Networks*, *113*, 244–257. doi:10.1016/j.comnet.2016.12.019

NCVD. (2020, Mar 27). *National Cardiovascular Disease Database*. Retrieved from http://www.acrm.org.my/ncvd/

Nichols, S. J. (2018). Inflammatory markers and novel risk factors. In P. Toth & C. Cannon (Eds.), *Comprehensive Cardiovascular Medicine in the Primary Care Setting* (pp. 87–98). Humana Press.

Nishimura, H., Miura, T., Minamisawa, M., Ueki, Y., Abe, N., Hashizume, N., Mochidome, T., Harada, M., Shimizu, K., Shoin, W., Yoshie, K., Oguchi, Y., Ebisawa, S., Motoki, H., Izawa, A., Koyama, J., & Ikeda, U. (2017). Prognostic value of ankle brachial index for future incident heart failure in patients without previos heart failure: Data from the impressive predictive value of ankle brachial index for clinical long term outcome in patients with cardiovascular disease examined by ABI study. *Heart and Vessels*, *32*(3), 295–302. doi:10.1007/s00380-016-0873-3 PMID:27412398

Okore, S. K., Bakyarani, E. S., & Phil, M. (2015). Big data personal health records management and analysis on cloud using no sql-mongo db. *International Journal of Computer Trends and Technology*, *20*(2), 68–74. doi:10.14445/22312803/IJCTT-V28P113

Pfohl, S., Marafino, B., Coulet, A., Rodriguez, F., Palaniappan, L., & Shah, N. H. (2019). Creating fair models of atherosclerotic cardiovascular disease. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (AIES '19)*. Honolulu, HI: ACM. doi:10.1145/3306618.3314278

Piepoli, M. F., Hoes, A. W., Agewall, S., Albus, C., Brotons, C., & Binno, S. et al. (2016). 2016 European Guidelines on cardiovascular disease prevention in clinical practice. *European Heart Journal*, *37*(29), 2315–2381. doi:10.1093/eurheartj/ehw106 PMID:27222591

Rallidis, L. S., Tellis, C. C., Lekakis, J., Rizos, I., Varounis, C., Charalampopoulos, A., Zolindaki, M., Dagres, N., Anastasiou-Nana, M., & Tselepis, A. D. (2012). Lipoprotein-associated phospholipase A_2 bound on high-density lipoprotein is associated with lower risk for cardiac death in stable coronary artery disease patients: A 3-year follow-up. *Journal of the American College of Cardiology*, *60*(20), 2053–2060. doi:10.1016/j.jacc.2012.06.057 PMID:23083783

Ravani, A., Werba, J. P., Frigerio, B., Sansaro, D., Amato, M., Tremoli, E., & Baldassarre, D. (2015). Assessment and relevance of carotid intima-media thickness in primary and secondary cardiovascular prevention. *Current Pharmaceutical Design*, *21*(9), 1164–1171. doi:10.2174/1381612820666141013121545 PMID:25312737

Romanens, M., Sudano, I., Adams, A., & Schober, E. A. (2019). Sonographic assessment of carotid atherosclerosis: Preferred risk indicator for future cardiovascular events? *Swiss Medical Weekly*, *149*, 1–14. doi:10.4414/smw.2019.20142 PMID:31846503

Rosas, E. A., Badin Castro, J., Vences Anaya, D. A., Moral Diez, J. J., Masso Bueso, J. S., Jiménez Niño, A., & Meave González, A. et al. (2017). Metabolic syndrome and diabetes mellitus in Mexico: The role of PET/CT in endothelial dysfunction and cardiovascular disease detection. *Current Cardiovascular Imaging Reports*, *10*(7), 1–8. doi:10.1007/s12410-017-9404-6

Saeed, A., Dabhadkar, K., Virani, S. S., Jones, P. H., Ballantyne, C. M., & Nambi, V. (2018). Cardiovascular disease prevention: Training opportunities, the challenges, and future directions. *Current Atherosclerosis Reports*, *20*(35), 1–7.

Setiawan, A. W., Utami, N., Mengko, T. R., & Indrayanto, A. (2014). Implementation of electronic medical record in community health center towards medical big data analytics application. In *Proceedings of the International Conference on Electrical Engineering and Computer Science (ICEECS '14)*. Kuta, Indonesia: IEEE. doi:10.1109/ICEECS.2014.7045215

Sood, E., & Gidding, S. S. (2016). Childhood psychosocial determinants of cardiovascular health. *Current Cardiovascular Risk Reports*, *10*(20), 1–8.

Tahmassebi, A., Ehtemami, A., Mohebali, B., Gandomi, A. H., Pinker, K., & Meyer-Baese, A. (2019). Big data analytics in medical imaging using deep learning. In *Proceedings of the Conference on Big Data: Learning, Analytics, and Applications*. SPIE. doi:10.1117/12.2516014

Van de Sande, D. A. J. P., Breuer, M. A. W., & Kemps, H. M. C. (2016). Utility of exercise electrocardiography in pre-participation screening in asymptomatic athletes: A systematic review. *Sports Medicine (Auckland, N.Z.)*, *46*(8), 1155–1164. doi:10.1007/s40279-016-0501-5 PMID:26893099

Wang, Y., Wang, L., Rastegar-Mojarad, M., Moon, S., Shen, F., Afzal, N., Liu, S., Zeng, Y., Mehrabi, S., Sohn, S., & Liu, H. (2018). Clinical information extraction applications: A literature review. *Journal of Biomedical Informatics*, *77*, 34–49. doi:10.1016/j.jbi.2017.11.011 PMID:29162496

WHF. (2017). Factsheet: Enfermedades cardiovasculares en México. *World Heart Federation*. Retrieved from https://www.world-heart-federation.org/wp-content/uploads/2017/05/Cardiovascular _diseases_in_Mexico__Spanish_.pdf

WHO. (2018, May 24). The top 10 causes of death. *World Health Organization.* Retrieved from https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

Yadlowsky, S., Hayward, R. A., Sussman, J. B., McClelland, R. L., Min, Y.-I., & Basu, S. (2018). Clinical implications of revised pooled cohort equations for estimating atherosclerotic cardiovascular disease risk. *Annals of Internal Medicine*, *169*(1), 20–30. doi:10.7326/M17-3011 PMID:29868850

Zaydi Anaí Acosta Chí was born in January 13 of 1984. She graduated in Nutrition and has a Master's degree in Chemical and Biochemical Sciences. she is interested in public health issues and nutritional intervention programs. She is currently working at "Universidad de la Cañada" in Oaxaca, Mexico, and is teaching subjects about community practices and public food polices.