

Classification of Exploring the Mortality Rates of Diseases by Using Machine Learning Techniques



Imdad Ali Shah, Dr. Riaz Ahmed Shaikh, Dr. Samina Rajper

Department of Computer Science Shah Abdul University Khairpur. Sindh, Pakistan

imdad.shah@salu.edu.pk, riaz.shaikh@salu.edu.pk, Samina.rajper@salu.edu.pk

ABSTRACT

Early years, the analysis of healthcare data has become one of the most committed research fields. The several types of data in healthcare which are open several doors for the new researchers. In this study, we have used the electronic medical record database (EMRD) and compared the mortality rates of NICU Department of KMC, Civil Hospital, District Khairpur. We have identified the incident based on the diagnosis of the last year. There are several types of diseases such as Neonatal Sepsis, Pre-mature, Birth Asphyxia, Meconium Aspiration Syndrome, Neonatal Jaundice, Pneumonia, and Low birth weight, Congenital Heart Disease (HD), Surgical Problems and Respiratory Distress Syndrome (RDS). We conducted sensitivity analyses of EMRD among patients who were admitted in NICU ward, the record has been analyzed using ML techniques and applied K-NN algorithm for classification and data analysis of disease. Our data analysis result will help the Government and the private sector to give considerable focus on giving medical facilities.

Key words: Patient Data, Electric Medical Record, Machine Learning, Diagnosis and Patient's record

1. INTRODUCTION

Health-care (HC) is an expression that refers to the structure of improving or maintaining health facilities to meet people's medical needs. In HC, health records are maintained and restored by the patients, vendors, Physicians, and health companies. Several types of diseases are handled by health care including diabetes, strokes, cancer, and so on using ML. The deadliest diseases in Neonatal Sepsis, Pre-mature, Birth Asphyxia, Meconium Aspiration Syndrome, Neonatal Jaundice, Pneumonia, Low birth weight, congenital heart disease (HD), Surgical Problem and Respiratory Distress Syndrome (RDS). The controlling, measuring diagnosis

and treatment of human, it needs to analyze and predict. The exact and reliable data are necessary for analysis for treatment proposes [1,2]. SL contains a training model with a labeled record and applying the modeled that have trained for predictions, builds in light of the current record. It contains the distribution of data this involves the distribution of the data into two sets, including a training to set and a testing set. As the first model is trained to set and then the performance is tested on the testing set. The model's performance might be evaluated using performance metrics [3,4,5]. RL works by emerging a system that improves its performance by taking feedback from the situation and taking possible steps to improve them. It is a process of learning from the environment by interacting with it without human help. This is a repetitive process. Figure-1 shows the relationship between the data and output [6],[7]. The medical and genomic records have been applied through the ML algorithms to analyze those records. The aim of this study, the electronic medical record database is increasing day to day and ML is using for accuracy. In light of the accuracy of the record, future strategies are made for giving health facilities to the needy patient. In the field of healthcare and mortality rate, huge information regarding the patient's health and is impossible for humans to process. Therefore, ML gives a technique to identify patterns from large-scale data and uses algorithms to predict patients' future outcomes. ML in healthcare helps users understand the potential of existing programs and identifies treatments that provide patients with the best outcomes according to their condition. [11] A well-defined multilevel perception model was praised for predicting the two-year survival of non-small cell lung cancer patients. [12] The researcher focused on a new improved classification approach for survival prediction Hepatocellular (HCC). Picture-1 shows the classification of Diagnosis and picture-II shows the classification of cases and mortality rate using ML. Figure-1 show the name of diseases, shows tool Development and Design is given using Figure- 3 and 4.

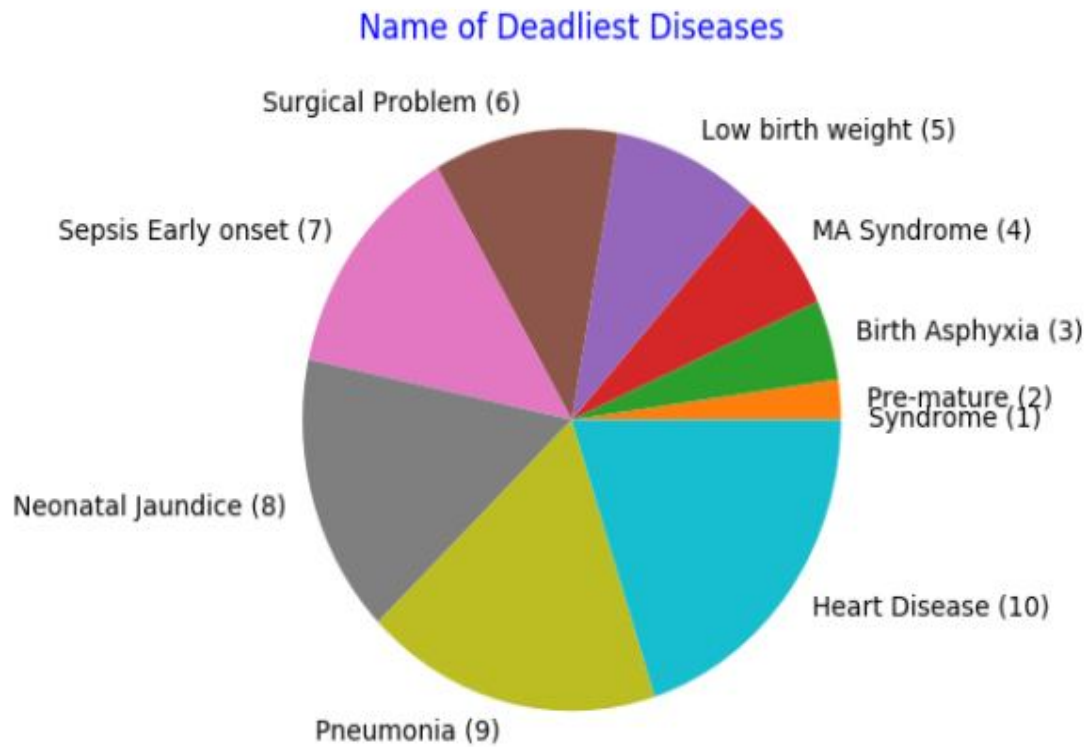


Figure-1 show the name diseases

1. Literature Review

| Year | Papers | Prose | Cones |
|------|--------|---|--|
| 2018 | [13] | Has worked for knowing the exact death cause on deathliest diseases | Suggested to continue work for knowing the Mortality rate of deathliest diseases |
| 2018 | [14] | Worked on the fibrillation and chronic kidney disease was related increasing rate of haemorrhage and ischaemic stroke on mortality cause. | Suggested to need peruse our explore findings for providing exact clinical controlling |
| 2018 | [15] | Has worked on chronic diseases and mortality rates | Has explored the limitation that our study covers from 2012 to 2014 and need further study in this regard |
| 2018 | [16] | Has worked on the diseases Colorectal cancer and analysis Cancer Registry data of the management Center from 2008 to 2011for analysis risk causes | Suggested to urgent work on this area due to limitation period |
| 2019 | [17] | Proposed to analysis of climate-mortality in parts of high malaria burden (Kenya) | Suggested to need more work in this research area. |
| 2019 | [18] | Proposed to studies on the traditional Chinese medicine and investigate renal disease & know the mortality rates users TCM treatment with DN | Further explore the TCM modalities and medications are needed still |
| 2020 | [19] | Has analysis data from the death certificate “alcohol-related mortality rate” | Explore that alcohol-related mortality increase and need for improving surveillance of alcohol-involved mortality |
| 2020 | [20] | Proposed to analyzed the mortality rate affected by the Covid-19 pandemic | Proposed the 27 regions of the countries and four regions were not express in the analyzed of the Covid-19 pandemic mortality rate |

2. DATA AND METHOD

In the prospect of the data, we have studied many associated research articles and book chapters of mortality rates in various diseases and data analysis using ML Techniques. The data has been taken from the NICU department KMC, Civil Hospital, Khairpur. Additionally, the electronic medical record database of this hospital has

been collected and used last year. Using ML techniques and applied K-NN algorithm for classification and analysis of disease. Fig-2 show the formula, Fig-3 show the flowchart of the K-NN classifier Procedure, Figure-4 show data and ML tool, Fig-5 shows tool development tool

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Fig-2 Show the formula of K-NN algorithm

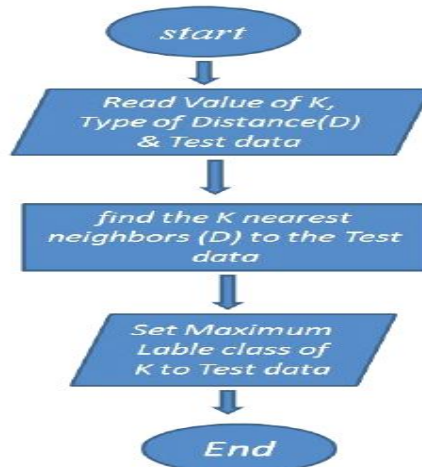


Fig-3 Show the flowchart of K -NN classifier procedure

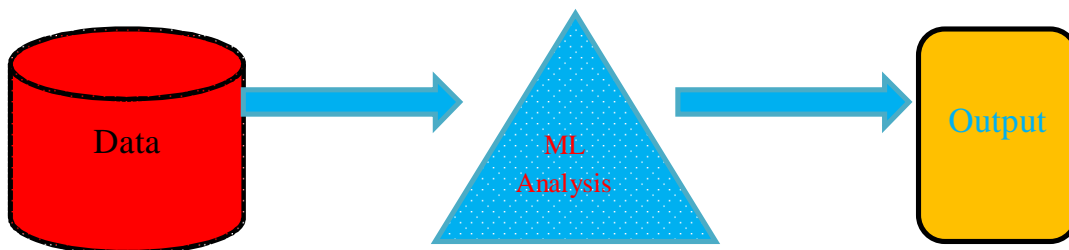


Figure-4 show the connectivity of data and ML tool

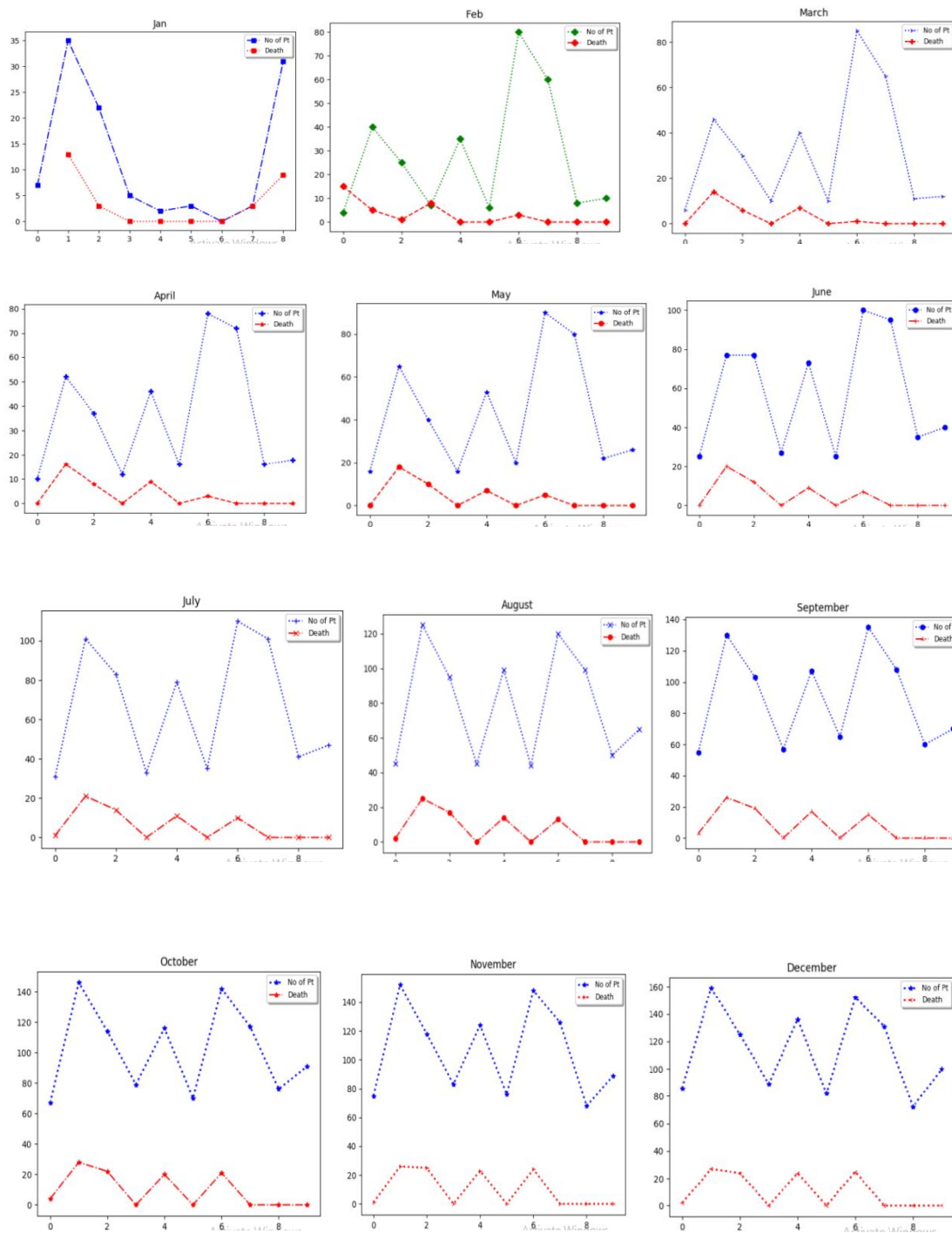


Figure-4 shows the monthly classification of the patient year 2020

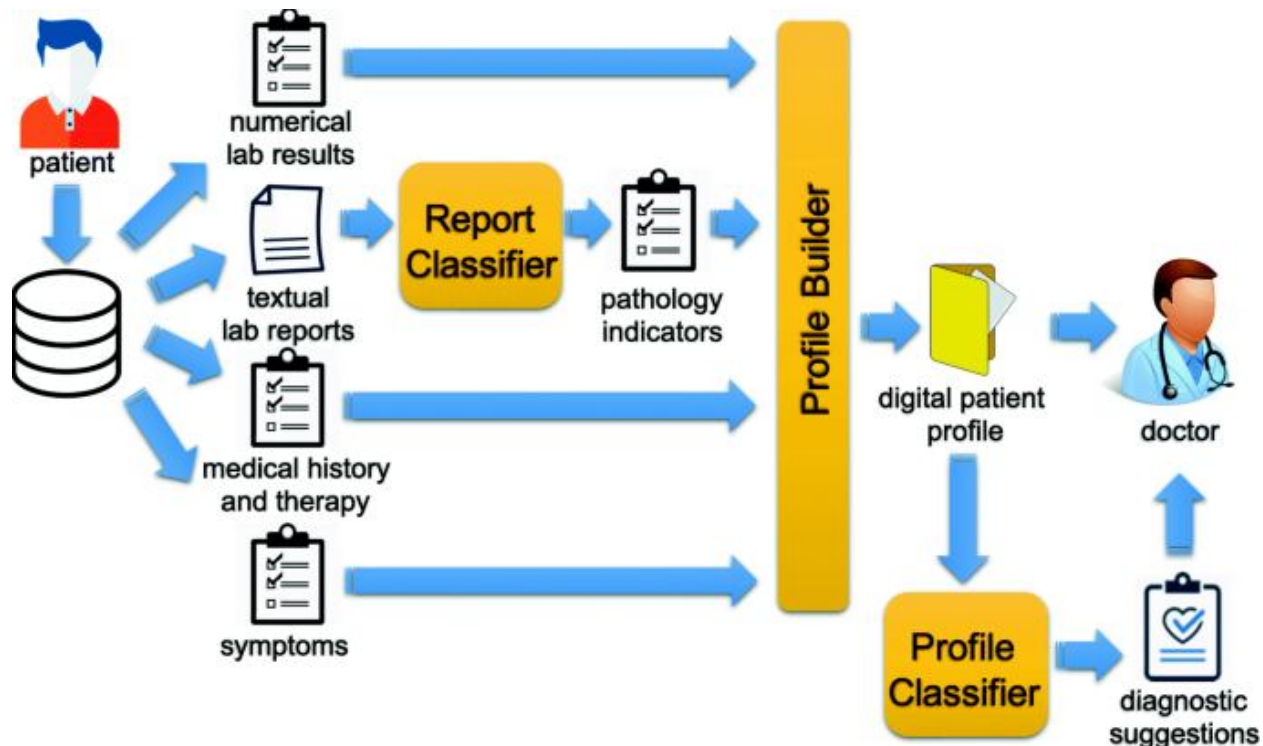


Figure-5 Shows the connectivity of tool Development and Design

3. CONCLUSION AND RESULTS

In healthcare and mortality rate, different types of data are present, they can analyze a variety of data several Machine Learning algorithms. There are several diseases such as Neonatal Sepsis, Pre-mature, Birth Asphyxia, Meconium Aspiration Syndrome, Neonatal Jaundice, Pneumonia, and Low birth weight, Congenital Heart Disease (HD), Surgical Problems and Respiratory Distress Syndrome (RDS). This study has

identified the incident based on the year of diagnosis of the last year. Additionally, this study has compared the mortality rates of disease Department of NICU KMC, Civil Hospital, Khairpur for potential confounders between the cohorts. We conducted sensitivity data analyses among patients, using Machine Learning (ML) and applied K-NN algorithm for classification and data analysis of disease Figure-5 show the result.

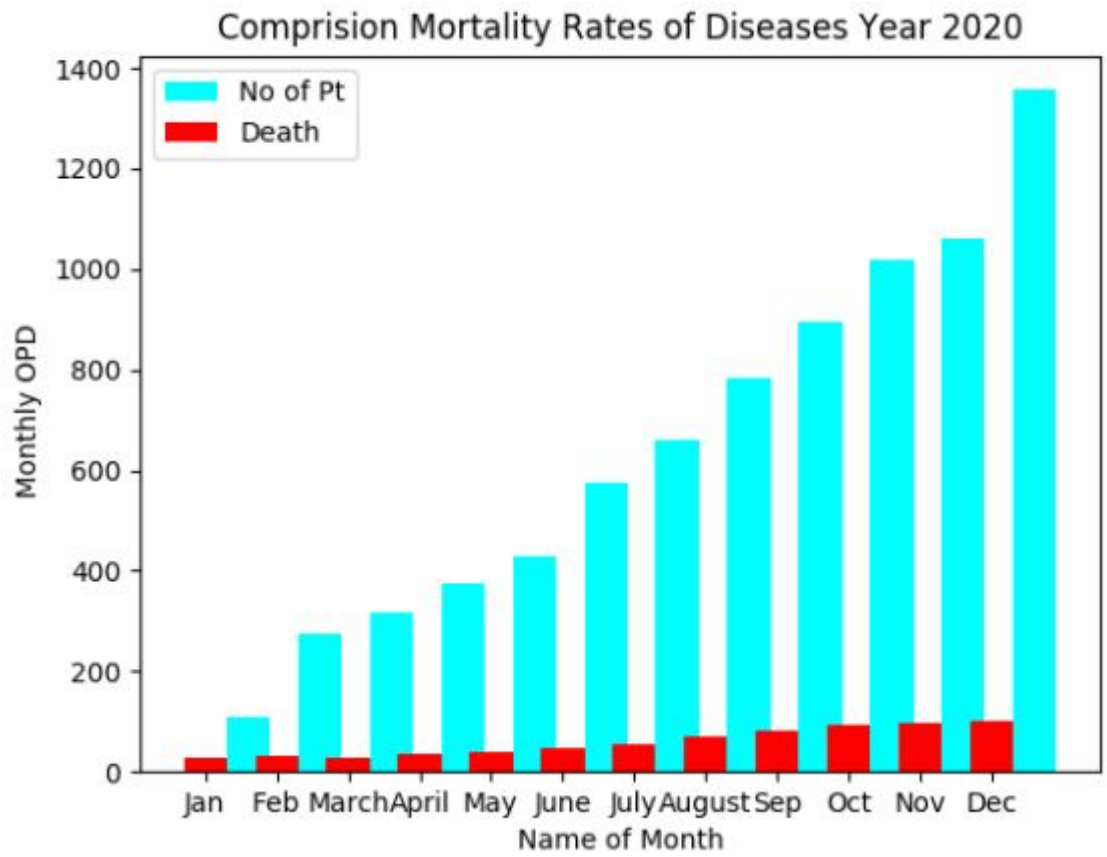


Figure-5 shows the results

REFERENCES

1. Dhillon, A., & Singh, A. (2019). Machine learning in healthcare data analysis: a survey. *Journal of Biology and Today's World*, 8(6), 1-10. <https://www.iomcworld.org/articles/machine-learning-in-healthcare-data-analysis-a-survey-44184.html>
2. Bisaso, K. R., Anguzu, G. T., Karungi, S. A., Kiragga, A., & Castelnovo, B. (2017). A survey of machine learning applications in HIV clinical research and care. *Computers in biology and medicine*, 91, 366-371. <https://www.sciencedirect.com/science/article/abs/pii/S001048251730361X>
3. Linthicum, K. P., Schafer, K. M., & Ribeiro, J. D. (2019). Machine learning in suicide science: Applications and ethics. *Behavioral sciences & the law*, 37(3), 214-222. <https://onlinelibrary.wiley.com/doi/full/10.1002/bsl.2392>
4. Tahmassebi, A., Wengert, G. J., Helbich, T. H., Bago-Horvath, Z., Alaei, S., Bartsch, R., ... & Morris, E. A. (2019). Impact of machine learning with multiparametric magnetic resonance imaging of the breast for early prediction of response to neoadjuvant chemotherapy and survival outcomes in breast cancer patients. *Investigative radiology*, 54(2), 110. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6310100/>
5. Ohkuma, T., Zoungas, S., Jun, M., Lisheng, L., Mancina, G., Marre, M., ... & ADVANCE Collaborative Group. (2020). Intensive glucose-lowering and the risk of vascular events and premature death in patients with decreased kidney function: The ADVANCE trial. *Diabetes, Obesity and Metabolism*, 22(3), 452-457. <https://dom-pubs.onlinelibrary.wiley.com/doi/full/10.1111/dom.13878>
6. Patel, N. M., Michelini, V. V., Snell, J. M., Balu, S., Hoyle, A. P., Parker, J. S., ... & Xu, J. (2018). Enhancing next-generation sequencing-guided cancer care through cognitive computing. *The oncologist*, 23(2), 179. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5813753/>
7. Gross, J., Baumgartl, H., & Buettner, R. (2020). A Novel Machine Learning Approach for High-Performance Diagnosis of Premature Internet Addiction Using the Unfolded EEG Spectra. *AMCIS'20 Proc.* <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0214365&LinkSource=PasleApp>
8. Diller, G. P., Kempny, A., Babu-Narayan, S. V., Henrichs, M., Brida, M., Uebing, A., ... & Dimopoulos, K. (2019). Machine learning algorithms estimating prognosis and guiding therapy in adult congenital heart disease: data from a single tertiary centre including 10 019 patients. *European heart journal*, 40(13), 1069-1077. <https://discovery.ucl.ac.uk/id/eprint/10076628/>
9. Krishna, M. R., & Kumar, R. K. (2020). Diagnosis and management of critical congenital heart diseases in the newborn. *The Indian Journal of Pediatrics*, 1-7. <https://link.springer.com/article/10.1007/s12098-019-03163-4>
10. Gupta, N., Potdar, V., Praharaj, I., Giri, S., Sapkal, G., Yadav, P., ... & Munivenkatappa, A. (2020). Laboratory preparedness for SARS-CoV-2 testing in India: Harnessing a network of virus research & diagnostic laboratories. *The Indian Journal of Medical Research*, 151(2-3), 216. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7258754/>
11. Rahil Pasha, S. A., Bhat, P., Abdul Naseer, A., & Ved, N. C. Ventilator-Associated Pneumonia (VAP) with Multidrug-Resistant (MDR) Pathogens in Geriatric Patients: Risk Factors and their Antibiotic Susceptibility Pattern with Detection of MRSA, ESBLs and MBLs in Intensive Care Unit. <https://pdfs.semanticscholar.org/a048/cb475566b7c0f37d5f44c97717c7b4404025.pdf>
12. Occhi-Alexandre, I. G. P., Cruz, P. V., Bendo, C. B., Paiva, S. M., Pordeus, I. A., & Martins, C. C. (2020). Prevalence of dental caries in preschool children born preterm and/or with low birth weight: A systematic review with meta-analysis of prevalence data. *International Journal of Paediatric Dentistry*, 30(3), 265-275. <https://onlinelibrary.wiley.com/doi/full/10.1111/ipd.12610>
13. D'Amato, V., Di Lorenzo, E., & Sibillo, M. (2018). Dread disease and cause-specific mortality: Exploring new forms of insured loans. *Risks*, 6(1), 13. <https://www.mdpi.com/2227-9091/6/1/13>
14. Kumar, S., de Lusignan, S., McGovern, A., Correa, A., Hriskova, M., Gatenby, P., ... & Camm, A. J. (2018). Ischaemic stroke, haemorrhage, and mortality in older patients with chronic kidney disease newly started on anticoagulation for atrial fibrillation: a population based study from UK primary

- care. *bmj*, 360. <https://www.bmj.com/content/360/bmj.k342>
15. Raghupathi, W., & Raghupathi, V. (2018). An empirical study of chronic diseases in the United States: a visual analytics approach to public health. *International journal of environmental research and public health*, 15(3), 431. <https://www.mdpi.com/1660-4601/15/3/431>
 16. Mansori, K., Mosavi-Jarrahi, A., Motlagh, A. G., Solaymani-Dodaran, M., Salehi, M., Delavari, A., ... & Asadi-Lari, M. (2018). Exploring spatial patterns of colorectal cancer in Tehran City, Iran. *Asian Pacific journal of cancer prevention: APJCP*, 19(4), 1099. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6031791/>
 17. Manthey, J., & Rehm, J. (2019). Mortality from alcoholic cardiomyopathy: Exploring the gap between estimated and civil registry data. *Journal of clinical medicine*, 8(8), 1137. <https://www.mdpi.com/2077-0383/8/8/1137>
 18. Chen, H. Y., Pan, H. C., Chen, Y. C., Chen, Y. C., Lin, Y. H., Yang, S. H., ... & Wu, H. T. (2019). Traditional Chinese medicine use is associated with lower end-stage renal disease and mortality rates among patients with diabetic nephropathy: a population-based cohort study. *BMC complementary and alternative medicine*, 19(1), 81. [file:///C:/Users/ASIAN%20COMPUTER/Downloads/Chen2019_Article_TraditionalChineseMedicineUseI%20\(2\).pdf](file:///C:/Users/ASIAN%20COMPUTER/Downloads/Chen2019_Article_TraditionalChineseMedicineUseI%20(2).pdf)
 19. White, A. M., Castle, I. J. P., Hingson, R. W., & Powell, P. A. (2020). Using death certificates to explore changes in alcohol-related mortality in the United States, 1999 to 2017. *Alcoholism: Clinical and Experimental Research*, 44(1), 178-187. <https://onlinelibrary.wiley.com/doi/abs/10.1111/acer.14239>
 20. Lima, E., Vilela, E., Peralta, A., Rocha, M. G., Queiroz, B. L., Gonzaga, M. R., ... & Piscoya, M. (2020). Exploring excess of deaths in the context of covid pandemic in selected countries of Latin America. *Open Science Framework*. [file:///C:/Users/ASIAN%20COMPUTER/Downloads/Excess_mortality_Lima_etal%20\(1\).pdf](file:///C:/Users/ASIAN%20COMPUTER/Downloads/Excess_mortality_Lima_etal%20(1).pdf)