Available Online at http://www.bjbmr.org

BRITISH JOURNAL OF BIO-MEDICAL RESEARCH

Cross Ref DOI: https://doi.org/10.24942/bjbmr.2022.962 **Volume 06, Issue 02, March – April 2022**

ISSN:2456-9836
IF: 5.719

Research Article

Artificial Intelligence (AI) models predicting cardiac risk. Are the developed models optimal in the accuracy of clinical prediction, population- specific and robust? Necessity for specific risk variables and new hybrid models.

Dr. Vidya Sagar[1]*, Mr. Gokul Rajan[2], Mr. Sriram Vaidyanathan[3], Mr. Deepak Batra[4], Mr. Rajaram Naganur[5], Dr Anwesh Reddy[6]

[1]Head Tech, Aegis Lifesciences Pvt. Ltd, Ahmedabad, India
[2]COO. Digital Initiative, Hinduja Leyland Finance, Chennai, India
[3]Project Manager, Ideas2IT Technologies, India
[4]Global Head - Finance/CFO, Evalue serve, India
[5]General Manager, Samsung Research Institute, Bangalore, India
[6]Senior Data Scientist, Great Learning, Hyderabad, India

ARTICLE INFO

Article History:
Received on 04th March 2022
Peer Reviewed on 18th Mar 2022
Revised on 14th April 2022
Published on 28th April 2022

Keywords:
AI, Models, Cardiac, Risk, Prediction, Variables

ABSTRACT

Artificial Intelligence (AI) algorithms have changed the landscape of Cardio Vascular Diseases (CVD) risk assessment and demonstrated a better performance mainly due to their ability to handle input nonlinear variations. Most commonly used algorithms in CVD risk predications were classification and regression tress (CART).

Though most of the developed models have shown good accuracy, but have not considered risks factors related to specific population, which play an integral role in predicting the risk of CVDs. This include gender specific clinical risk factors (hormonal changes, bone density etc.), metrological, chronological data, exposure to environmental pollutants, race, genotype, hereditary, dietary intake, physical inactivity, psychological stress, cardiac markers, post covid infection status etc. Secondly the existing models have not included the weighing and grading of the risks factors and Prediction, as all factors wont contribute equally to the Cardiac Risk. Importantly predictive models can be readily used within the populations in which they were developed but practically they often give a less satisfactory performance, when applied to another population because of the Inter genetic variations especially linked to CVDs.

Hence there is necessity to develop upgraded AI models or Hybrid models (Logistic regression with decision tree, NN etc.). Inclusion of more descriptive and apt risk factors or variables, specific to a subset of population, Race, Genotype is quite essential. Secondly allotting weighing for Risk factors and grading for Risk Prediction in the models, will provide accurate cardiac risk prediction compared to other approaches. So presently the solution should be more Data centric with equal importance to AI Models.

Corresponding Author: *Dr Vidya Sagar, Head – Tech, Medical Devices, Aegis Lifesciences Pvt. Ltd, Ahmedabad, Gujarat, India – 382213*

INTRODUCTION

CARDIOVASCULAR DISEASES - WORLD WIDE DATA AND ANALYSIS

Heart disease, alternatively known as cardiovascular disease (CVDs), encases various conditions that impact the heart and is the primary basis of death worldwide over the span of the past few decades. It associates many risk factors in heart disease and a need of the time to get accurate, reliable, and sensible approaches to make an early diagnosis to achieve prompt management of the disease. Data mining is a commonly used technique for processing enormous data in the healthcare domain.

CVDs, despite the significant advances in the diagnosis and treatments, still represents the leading cause of morbidity and mortality worldwide. In order to improve and optimize CVD outcomes, artificial intelligence techniques have the potential to radically change the way we practice cardiology, especially in imaging, offering us novel tools to interpret data and make clinical decisions. AI techniques such as machine learning and deep learning can also improve medical knowledge due to the increase of the volume and complexity of the data, unlocking clinically relevant information. Likewise, the use of emerging communication and information technologies is becoming pivotal to create a pervasive healthcare service through which elderly and chronic disease patients can receive medical care at their home, reducing hospitalizations and improving quality of life. CVDs such as ischaemic heart disease and cerebrovascular such as stroke account for 17.7 million deaths and are the leading cause in accordance with the World Health Organization. [1]

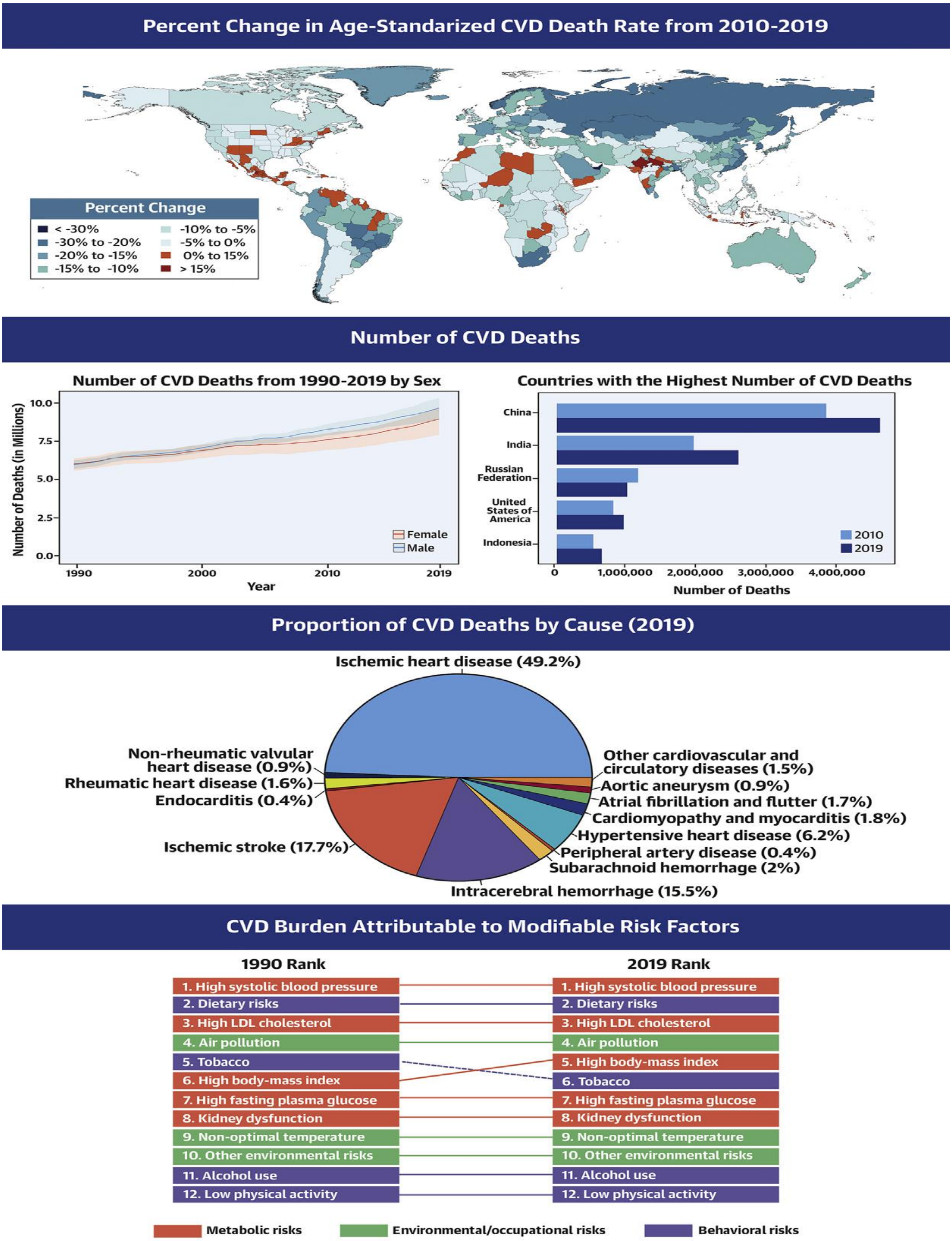
CVDs are common, have poor survival, and are increasing worldwide (Figure 1). Prevalent cases of total CVD nearly doubled from 271 million (95% UI: 257 to 285 million) in 1990 to 523 million (95% UI: 497 to 550 million) in 2019, and the number of CVD deaths steadily increased from 12.1 million (95% UI: 11.4 to 12.6 million) in 1990, reaching 18.6 million (95% UI: 17.1 to 19.7 million) in 2019 (Figure 1A). The global trends for DALYs and YLLs also increased significantly, and YLDs doubled from 17.7 million (95% UI: 12.9 to 22.5 million) to 34.4 million (95% UI: 24.9 to 43.6 million) over that period. [2-4]

At the country level, age-standardized mortality rates for total CVD were highest in Uzbekistan, Solomon Islands, and Tajikistan and were lowest in France, Peru, and Japan, where rates were 6-fold lower in 2019. From 1990 to 2019, large declines in the age-standardized rates of death, DALYs, and YLLs, together with small gradual reductions in age standardized rates for prevalent cases and YLDs, suggest that population growth and aging are major drivers of the increase in total CVD. In 2019, total CVD DALYs were higher in men than women before age 80 to 84 years. After this age, the pattern reverses. The sex differences in DALYs is most striking between ages 30 and 60 years (men greater) and age >80 years (women greater). [5-7]

The excess CVD deaths in women beginning at ages 80 to 84 years should focus attention to cause-specific mortality at older ages and have implications for secondary prevention strategies. Among women, the age-standardized rates for DALYs were highest in Central Asia, Oceania, North Africa and the Middle East, and Eastern Europe; and lowest in High-Income Asia Pacific, Australasia, and Western Europe. Among men, age-standardized rates for DALYs were highest in Central Asia, Eastern Europe, and Oceania; and lowest in High-Income Asia Pacific, Australasia, Western Europe, and Andean Latin America. At the country level, the highest age-standardized rates were estimated for many of the islands of Oceania, Uzbekistan, and Afghanistan, while the lowest rates for DALYs were seen in Japan, France, and Israel. These regional and national differences in total CVD burden and mortality reflect differences in prevalence of CVD risk factors as well as access to health care [8] Differences in access to effective primary and secondary prevention strategies may also play a role in differences in total CVD burden, especially in low- and middle-income countries (LMICs). [9]

Global patterns of total CVD have significant implications for clinical practice and public health policy development. [10] Prevalent cases of total CVD are likely to increase substantially as a result of population growth and aging, especially in Northern Africa and Western Asia, Central and Southern Asia, Latin America and the Caribbean, and Eastern and South eastern Asia, where the share of older persons is projected to double between 2019 and 2050. [11-12] Increased attention to promoting ideal cardiovascular health and healthy aging across the lifespan is necessary. [13] Equally importantly, the time has come to implement feasible and affordable strategies for the prevention and control of CVD and to monitor results. [14]

Figure 1: Central Illustration of Cardiovascular Disease Burden Across Location, Cause, and Risk Factors (Courtesy: Roth, G.A. et al. J Am College of Cardiology 2020) [10]



AI MODELS IN HEART DISEASE PREDICTION

There is an increasing interest in predicting the probability of adverse events for patients hospitalized for medical or surgical treatment. Accurately predicting the probability of adverse events allows for effective patient risk stratification, thus permitting more appropriate medical care to be delivered to patients. [14-19] Furthermore, accurately predicting the probability of an adverse event allows for risk-adjusted outcomes to be compared across providers of health care. [15]

Logistic regression is the most commonly used method for predicting the probability of an adverse outcome in the medical literature. Recently, data-driven methods, such as classification and regression trees (CART) have been used to identify subjects at increased risk of adverse outcomes or of increased risk of having specific diagnoses. [6-41] Advocates for CART have suggested that these methods allow the construction of easily interpretable decision rules that can easily be applied in clinical practice. Furthermore, CART methods are adept at identifying important interactions in the data [31, 34, 40] and in identifying clinical subgroups of subjects at very high or very low risk of adverse outcomes. [41]

Several studies have compared the performance of regression trees and logistic regression for predicting outcomes. These studies can be grouped into three broad categories. First, studies that compared the variables identified by logistic regression as significant predictors of the outcome with those variables identified by a regression tree analysis as predictors of the outcome. [6-16] Second, studies that compared the sensitivity and specificity of logistic regression with that of regression trees. [6, 12, 17-30] Third, a small number of studies that compared the predictive accuracy, as measured by the area under the receiver operating characteristic (ROC) curve, of logistic regression with that of regression trees. [13-14, 31-39, 42]

The first category of studies does not allow one to compare the predictive ability of the two different prediction methods. Rather, it compares agreement on which factors are prognostically important. Since each model uses variables in a different manner, it is possible that the methods could differ in predictive accuracy, yet agree on which factors are prognostically important. The second category of studies compares sensitivity and specificity of regression trees with that of logistic regression. However, computing sensitivity and specificity from a logistic regression model requires specifying a probability threshold, and then assuming that the response will be positive if the predicted probability exceeds this probability threshold. [43] In particular, it is highly dependent upon the probability threshold chosen for a positive prediction. Furthermore, it is an insensitive and inefficient measure of predictive accuracy. [44]

Only a small number of studies have compared the predictive ability of regression trees with that of logistic regression using the area under the ROC curve. [13-14, 31-39, 42] Among these studies, the conclusions were inconsistent. Six studies concluded that regression trees and logistic regression had comparable performance. [13, 31, 33, 36-38]; five studies concluded that logistic regression had superior performance to regression trees [14, 32, 34, 39, 42]; while one study arrived at the opposite conclusion. [35] Only one recent study, using a relatively small sample, employed repeated split sample validation to examine the robustness of the findings to the particular splitting of the sample in derivation and validation samples. The authors of this study suggested that similar methods be applied in other disciplines and other data sets to test the validity of their findings. [38]

BRIEF NOTE ON SOME IMPORTANT, LATEST AI MODELS IN HEART DISEASE PREDICTION

C1. Heart Disease Prediction using Machine Learning Techniques: Devansh Shah *etal* work presents various attributes related to heart disease, and the model on basis of supervised learning algorithms as Naïve Bayes, decision tree, K-nearest neighbour, and random forest algorithm. It uses the existing dataset from the Cleveland database of UCI repository of heart disease patients.

The dataset comprises 303 instances and 76 attributes. Of these 76 attributes, only 14 attributes are considered for testing, important to substantiate the performance of different algorithms. This research paper aims to envision the probability of developing heart disease in the patients. The results portray that the highest accuracy score is achieved with K-nearest neighbour. [45]

C2. Machine Learning Technology-Based Heart Disease Detection Models: Different machine learning technologies based on heart disease detection by Umarani Nagavelli *etal*. Firstly, Naive Bayes with a weighted approach is used for predicting heart disease. Second one, according to the features of frequency domain, time domain, and information theory, is automatic and analyse ischemic heart disease localization/detection. Two classifiers such as support vector machine (SVM) with XGBoost with the best performance are selected for the classification in this method. Third one is the heart failure automatic identification method by using an improved SVM based on the duality optimization scheme also analysed.

Finally, for a clinical decision support system (CDSS), an effective heart disease prediction model (HDPM) is used, which includes density-based spatial clustering of applications with noise (DBSCAN) for outlier detection and elimination, a hybrid synthetic minority over-sampling technique-edited nearest neighbour (SMOTE-ENN) for balancing the training data distribution, and XGBoost for heart disease prediction. [46]

C3: Using machine learning to improve survival prediction after heart transplantation: This particular study investigates the use of modern machine learning (ML) techniques to improve prediction of survival after orthotopic heart transplantation (OHT). Retrospective study of adult patients undergoing primary, isolated OHT between 2000 and 2019 as identified in the United Network for Organ Sharing (UNOS) registry was performed. The primary outcome was 1-year post-transplant survival. Patients were randomly divided into training (80%) and validation (20%) sets. Dimensionality reduction and data re-sampling were employed during training. Multiple machine learning algorithms were combined into a final ensemble ML model. The discriminatory capability was assessed using the area under receiver-operating characteristic curve (AUROC), net reclassification index (NRI), and decision curve analysis (DCA). Results indicate that a total of 33,657 OHT patients were evaluated. One-year mortality was11% (n = 3738). In the validation cohort, the AUROC of singular logistic regression was 0.649 (95% CI, 0.628–0.670) compared to 0.691 (95% CI, 0.671–0.711) with random forest, 0.691 (95% CI, 0.671–0.712) with deep neural network, and 0.653 (95% CI, 0.632–0.674) with Ada boost. A final ensemble ML model was created that demonstrated the greatest improvement in AUROC: 0.764 (95% CI, 0.745–0.782) ($p < .001$). The ensemble ML model improved predictive performance by 72.9% ±3.8% ($p < .001$) as assessed by NRI compared to logistic regression. DCA showed the final ensemble method improved risk prediction across the entire spectrum of predicted risk as compared to all other models ($p < .001$). Modern ML techniques can improve risk prediction in OHT compared to traditional approaches. This may have important implications in patient selection, programmatic evaluation, allocation policy, and patient counselling and prognostication. [47]

C4: Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants was performed. Data-driven techniques based on machine learning (ML) might improve the performance of risk predictions by agnostically discovering novel risk predictors and learning the complex interactions between them. The Team tested (1) whether ML techniques based on a state-of-the-art automated ML framework (AutoPrognosis) could improve CVD risk prediction compared to traditional approaches, and (2) whether considering non-traditional variables could increase the accuracy of CVD risk predictions. Using data on 423,604 participants without CVD at baseline in UK Biobank, we developed a ML-based model for predicting CVD risk based on 473 available variables. Our ML-based model was derived using AutoPrognosis, an algorithmic tool that automatically selects and tunes ensembles of ML modelling pipelines (comprising data imputation, feature processing, classification and calibration algorithms). The group compared model with a well-established risk prediction algorithm based on conventional CVD risk factors (Framingham score), a Cox proportional hazards (PH) model based on familiar risk factors (i.e., age, gender, smoking status, systolic blood pressure, history of diabetes, reception of treatments for hypertension and body mass index), and a Cox PH model based on all of the 473 available variables.

Predictive performances were assessed using area under the receiver operating characteristic curve (AUC-ROC). Overall, our AutoPrognosis model improved risk prediction (AUCROC: 0.774, 95% CI: 0.768-0.780) compared to Framingham score (AUC-ROC: 0.724, 95% CI: 0.720-0.728, $p < 0.001$), Cox PH model with conventional risk factors (AUC-ROC: 0.734, 95% CI: 0.729-0.739, $p < 0.001$), and Cox PH model with all UK Biobank variables (AUC-ROC: 0.758, 95% CI: 0.753-0.763, $p < 0.001$). Out of 4,801 CVD cases recorded within 5 years of baseline, AutoPrognosis was able to correctly predict 368 more cases compared to the Framingham score. AutoPrognosis model included predictors that are not usually considered in existing risk prediction models, such as the individuals' usual walking pace and their self-reported overall health rating. Furthermore, our model improved risk prediction in potentially relevant sub-populations, such as in individuals with history of diabetes.

The working group highlighted the relative benefits accrued from including more information into a predictive model (information gain) as compared to the benefits of using more complex models (modelling gain). Auto Prognosis model improves the accuracy of CVD risk prediction in the UK Biobank population. This approach performs well in traditionally poorly served patient subgroups. Additionally, Auto Prognosis uncovered novel predictors for CVD disease that may now be tested in prospective studies. We found that the "information gain" achieved by considering more risk factors in the predictive model was significantly higher than the "modelling gain" achieved by adopting complex predictive models. [48]

C5: Detection of Cardiovascular Disease using Machine Learning Classification Models: The project intends to automatically detect cardiovascular disease using two datasets through a deep learning network and a variety of machine learning classification models. The performance evaluated based on the accuracy, precision, recall, and f-score for each of the models. Random Forest model achieved the highest performance at 94% accuracy in the heart diseases dataset, while Gradient Boosting model achieved the highest performance at 73% accuracy, 73% Recall, 73% F1-score, and 74% Precision in Cardiovascular Disease Dataset. [49]

C6: Machine learning model for predicting out-of-hospital cardiac arrests using meteorological and chronological data: The study evaluates a predictive model for robust estimation of daily out-of- Hospital cardiac arrest (OHCA) incidence using a suite of machine learning (ML) approaches and high-resolution meteorological and chronological data. Methods in this population-based study, we combined an OHCA nationwide registry and high-resolution meteorological and chronological datasets from Japan. We developed a model to predict daily OHCA incidence with a training dataset for 2005–2013 using the extreme Gradient Boosting algorithm. A dataset for 2014–2015 was used to test the predictive model.

The main outcome was the accuracy of the predictive model for the number of daily OHCA events, based on mean absolute error (MAE) and mean absolute percentage error (MAPE). In general, a model with MAPE less than10% is considered highly accurate. Results Among the 1 299 784 OHCA cases, 661 052OHCA cases of cardiac origin (525 374 cases in the training dataset on which fourfold cross-validation was performed and 135 678 cases in the testing dataset) were included in the analysis.

Compared with the ML models using meteorological or chronological variables alone, the ML model with combined meteorological and chronological variables had the highest predictive accuracy in the training (MAE 1.314 and MAPE 7.007%) and testing datasets (MAE 1.547 and MAPE 7.788%). Sunday, Monday, holiday, winter, low ambient temperature and large interday or intraday temperature difference were more strongly associated with OHCA incidence than other the meteorological and chronological variables. So a ML predictive model using comprehensive daily meteorological and chronological data allows for highly precise estimates of OHCA incidence. [50]

Table 1: Comparison of some important AI Models in Heart Disease Prediction

S. No	Title of the Study	Note on the AI model	Publication details
1	Heart Disease Prediction using Artificial Intelligence	K Neighbours, Support Vector, Decision Tree, Random Forest algorithms.	Zaibunnisa L. H. Malik, International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181, Published by, www.ijert.org, NREST – 2021 [51]
2	Machine Learning Outperforms	ML Risk Calculator based on Support Vector Machines	Ioannis A. Kakadiaris, PhD; Michalis Vrigkas, PhD; Albert A. Yen, MD; Tatiana Kuznetsova, MD; Matthew Budoff, MD; Morteza Naghavi,

S. No	Title of the Study	Note on the AI model	Publication details
	ACC/AHA CVD Risk Calculator in MESA		MD (J Am Heart Assoc. 2018;7:e009476. DOI: 10.1161/JAHA.118.009476.) [52]
3	Association of Fine Particulate Matter Exposure with Bystander-Witnessed Out-of-Hospital Cardiac Arrest	Logistic Regression	Sunao Kojima, MD, PhD; Takehiro Michikawa, MD, JAMA Network Open. 2020;3(4):e203043. doi:10.1001/jamanetworkopen.2020.3043 [53]
4	Machine learning prediction in cardiovascular diseases: a meta-analysis	SVM and boosting algorithms	Chayakrit Krittanawong, Scientific Reports (2020)10:16057, Nature Research https://doi.org/10.1038/s41598-020-72685-1 [54]
5	Deep-learning-based risk stratification for mortality of patients with acute myocardial infarction	Deep-learning-based risk stratification	Kwon J-m et al. (2019) Deep-learning-based risk stratification for mortality of patients with acute myocardial infarction. PLoS ONE 14(10): e0224502. https://doi.org/10.1371/journal.pone.0224502 [55]
6	An Algorithm Based on Deep Learning for Predicting In-Hospital Cardiac Arrest	Recurrent neural network AUROC), AUPRC) and the net reclassification index.	Joon-myoung Kwon, MD; Youngnam Lee, MSDOI: 10.1161/JAHA.118.008678, Journal of the American Heart Association [56]
7	A comparison of regression trees, logistic regression, generalized additive models, and multivariate adaptive regression splines for predicting AMI mortality	Classification and regression trees (CART), data-driven models: generalized additive models (GAMs) and multivariate adaptive regression splines (MARS)	Peter C. Austin STATISTICS IN MEDICINE, Statist. Med. 2007; 26:2937–2957 Published online 21 December 2006 in Wiley Inter Science (www.interscience.wiley.com) DOI: 10.1002/sim.2770 [57]

CONCLUSION

Identifying people at risk of cardiovascular diseases (CVD) is a cornerstone of preventative cardiology. Different approaches include Risk prediction models, currently recommended by clinical guidelines, typically based on a limited number of predictors with sub-optimal performance across all patient groups. Other Approaches in AI models can be used but are more generalized to all populations with inclusion of traditional risk factors or markers. In Indian context, aggressive screening tests should begin at an early age and will be beneficial for early detection and treatment to reduce the mortality.

Hence there is necessity to develop upgraded AI models or Hybrid models (Logistic regression with decision tree etc.). Inclusion of more descriptive and apt risk factors or variables, specific to a subset of population, Race (Caucasoid / Dravidian / Mongolian, Black, Red Indian etc..) is quite essential. Secondly allotting weighing (0 to 5) for Risk factors and grading for Risk Prediction (example 0 for mild, 1 for moderate, 2 for severe etc.) in the models, will provide accurate cardiac risk prediction compared to other approaches. So presently the solution should be more Data centric with equal importance to AI Models

Acknowledgements:

We sincerely thank Great Learning, Great Lakes, University of Texas, Austin and the team of AI / ML/ Data Science professionals, Dr. Sarkar, Dr. Kumar, Dr. Jitendra, Dr. Narayana, Dr. Anwesh Reddy and Team. Also many thanks to Ms. Sahana V our program coordinator and whole batch of PGPAIFL – Nov 2021, from Great Learning, Great Lakes.

REFERENCES

1. Vos T, Lim SS, Abbafati C. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020; 396:1204–22.
2. Mensah GA, Roth GA, Fuster V. The global burden of cardiovascular diseases and risk factors:2020 and beyond. J Am Coll Cardiol 2019;74: 2529–32.
3. Mensah GA, Wei GS, Sorlie PD, Decline in cardiovascular mortality. Circ Res 2017; 120: 366–80.
4. Murray CJL, Aravkin AY, Zheng P, et al. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020; 396:1223–49.
5. GBD Compare. Available at: https://vizhub. healthdata.org/gbd-compare/. Accessed November 11, 2020.
6. Global Health Data Exchange. Available at: http://ghdx.healthdata.org/. Accessed November 11, 2020.
7. Tolonen Hanna, Mähönen Markku, Asplund Kjell, Do trends in population levels of blood pressure and other cardiovascular risk factors explain trends in stroke event rates? Stroke 2002; 33:2367–75.
8. GBD 2017 Risk Factor Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018; 392:1923–94.
9. Husain MJ, Datta BK, Kostova D, et al. Access to cardiovascular disease and hypertension medicines in developing countries: an analysis of essential medicine lists, price, availability, and affordability. J Am Heart Assoc 2020; 9:e015302.
10. Roth GA, Forouzanfar MH, Moran AE, et al. Demographic and epidemiologic drivers of global J A C C VO L . -, NO. -, 2020 Roth et al. -, 2020 cardiovascular mortality. N Engl J Med 2015;372: 1333–41.
11. United Nations, Department of Economic and Social Affairs, Population Division. World Population Ageing 2019: Highlights. ST/ESA/SER.A/430. Available at:https://www.un.org/en/development/desa/population/publications/pdf/ageing/ WorldPopulationAgeing2019-Highlights. pdf. Accessed November 12, 2020.
12. United Nations, Department of Economic and Social Affairs, Population Division. World Population Prospects 2019: Highlights. ST/ESA/SER.A/ 423. Available at: https://population.un.org/ wpp/Publications/Files/WPP2019_Highlights.pdf . Available at: Accessed November 12, 2020.
13. Reynolds I, Page RL, Boxer RS. Cardiovascular health and healthy aging. In: Coll PP, editor. Healthy Aging: A Complete Guide to Clinical Management. Cham, Switzerland: Springer International Publishing, 2019:31–51.
14. Fuster V. Global burden of cardiovascular disease: time to implement feasible strategies and to monitor results. J Am Coll Cardiol 2014; 64:520–2.
15. Lee DS, Austin PC, Rouleau JL, Liu PP, Naimark D, Tu JV. Predicting mortality among patients hospitalized for heart failure: derivation and validation of a clinical model. Journal of the American Medical Association 2003; 290:2581–2587.
16. Tu JV, Jaglal SB, Naylor CD et al. Multicenter validation of a risk index for mortality, intensive care unit stay, and overall hospital length of stay after cardiac surgery. Circulation 1995; 91:677–684.
17. Lee KL, Woodlief LH, Topol EJ et al. Predictors of 30-day mortality in the era of reperfusion for acute myocardial infarction. Circulation 1995; 91:1659–1668.
18. Sullivan LM, Massaro JM, D'Agostino RB. Presentation of multivariate data for clinical use: the Framingham study risk score functions. Statistics in Medicine 2004; 23:1631–1660.
19. Iezzoni LI. Risk Adjustment for Measuring Healthcare Outcomes (2nd edn). Health Administration Press: Chicago, IL, 1997.
20. Geelhoed M, Boerrigter AO, Camfield P, Geerts AT, Arts W, Smith B, Camfield C. The accuracy of outcome prediction models for childhood-onset epilepsy. Epilepsia 2005; 46:1526–1532.
21. Nishida N, Tanaka M, Hayashi N, Nagata H, Takeshita T, Nakayama K, Morimoto K, Shizukuishi S. Determination of smoking and obesity as periodontitis risks using the classification and regression tree method. Journal of Periodontology 2005; 76:923–928.

22. Dessein PH, Joffe BI, Veller MG, Stevens BA, Tobias M, Reddi K, Stanwix AE. Traditional and non-traditional cardiovascular risk factors are associated with atherosclerosis in rheumatoid arthritis. Journal of Rheumatology 2005; 32:435–442.
23. Kuchibhatla M, Fillenbaum GG. Alternative statistical approaches to identifying dementia in a community-dwelling sample. Aging and Mental Health 2003; 7:383–389.
24. Avila PC, Segal MR, Wong HH, Boushey HA, Fahy JV. Predictors of late asthmatic response. Logistic regression and classification tree analyses. American Journal of Respiratory and Critical Care Medicine 2000; 161: 2092–2095.
25. Bhavnani SM, Drake JA, Forrest A, Deinhart JA, Jones RN, Biedenbach DJ, Ballow CH. A nationwise, multicenter, case–control study comparing risk factors, treatment, and outcome for vancomycin-resistant and -susceptible enterococcal bacteremia. Diagnostic Microbiology and Infectious Disease 2000; 36:145–158.
26. Wietlisbach V, Vader JP, Porchet F, Costanza MC, Burnand B. Statistical approaches in the development of clinical practice guidelines from expert panels: the case of laminectomy in sciatica patients. Medical Care 1999; 37:785–797.
27. Roehrborn CG, Malice M, Cook TJ, Girman CJ. Clinical predictors of spontaneous acute urinary retention in men with LUTS and clinical BPH: a comprehensive analysis of the pooled placebo groups of several large clinical trials. Urology 2001; 58:210–216.
28. Knuiman MW, Vu HT, Segal MR. An empirical comparison of multivariable methods for estimating risk of death from coronary heart disease. Journal of Cardiovascular Risk 1997; 4:127–134.
29. Litvan I, Campbell G, Mangone CA, Verny M, McKee A, Chaudhuri KR, Jellinger K, Pearce RK, D'Olhaberriague L. Which clinical features differentiate progressive supranuclear palsy (Steele–Richardson– Olszewski syndrome) from related disorders? A clinicopathological study. Brain 1997; 120:65–74.
30. Pilote L, Miller DP, Califf RM, Rao JS, Weaver WD, Topol EJ. Determinants of the use of coronary angiography and revascularization after thrombolysis for acute myocardial infarction. New England Journal of Medicine 1996; 335:1198–1205.
31. Bhattacharyya S, Siegel ER, Petersen GM, Chari ST, Suva LJ, Haun RS. Diagnosis of pancreatic cancer using serum proteomic profiling. Neoplasia 2004; 6:674–686.
32. Taylor WJ, Marchesoni A, Arreghini M, Sokoll K, Helliwell PS. A comparison of the performance characteristics of classification criteria for the diagnosis of psoriatic arthritis. Seminars in Arthritis and Rheumatism 2004; 34:575–584.
33. Seligman DA, Pullinger AG. Improved interaction models of temporomandibular joint anatomic relationships in asymptomatic subjects and patients with disc displacement with or without reduction. Journal of Orofacial Pain 2004; 18:192–202.
34. Stalans LJ, Yarnold PR, Seng M, Olson DE, Repp M. Identifying three types of violent offenders and predicting violent recidivism while on probation: a classification tree analysis. Law and Human Behavior 2004; 28:253–271.
35. Arena VC, Sussman NB, Mazumdar S, Yu S, Macina OT. The utility of structure-activity relationship (SAR) models for prediction and covariate selection in developmental toxicity: comparative analysis of logistic regression and decision three models. Sar and Qsar in Environmental Research 2004; 15:1–18.
36. Schwarzer G, Nagata T, Mattern D, Schmelzeisen R, Schumacher M. Comparison of fuzzy inference, logistic regression, and classification trees (CART). Prediction of cervical lymph node metastasis in carcinoma of the tongue. Methods of Information in Medicine 2003; 42:572–577.
37. Thwaites GE, Chau TT, Stepniewska K, Phu NH, Chuong LV, Sinh DX, White NJ, Parry CM, Farrar JJ. Diagnosis of adult tuberculous meningitis by use of clinical and laboratory features. Lancet 2002; 360:1287–1892.
38. Pullinger AG, Seligman DA, John MT, Harkins S. Multifactorial modeling of temporomandibular anatomic and orthopedic relationships in normal versus undifferentiated disk displacement joints. Journal of Prosthetic Dentistry 2002; 87:289–297.
39. Pullinger AG, Seligman DA. Multifactorial analysis of differences in temporomandibular joint hard tissue anatomic relationships between disk displacement with and without reduction in women. Journal of Prosthetic Dentistry 2001; 86:407–419.
40. Seligman DA, Pullinger AG. Analysis of occlusal variables, dental attrition, and age for distinguishing healthy controls from female

patients with intracapsular temporomandibular disorders. Journal of Prosthetic Dentistry 2000; 83:76–82.

41. Woolas RP, Conaway MR, Xu F, Jacobs IJ, Yu Y, Daly L, Davies AP, O'Briant K, Berchuck A, Soper JT et al. Combinations of multiple serum markers are superior to individual assays for discriminating malignant from benign pelvic masses. Gynecologic Oncology 1995; 59:111–1116.
42. Li D, German D, Lulla S, Thomas RG, Wilson SR. Prospective study of hospitalization for asthma. A preliminary risk factor model. American Journal of Respiratory and Critical Care Medicine 1995; 151:647–655.
43. Hasford J, Ansari H, Lehmann K. CART and logistic regression analyses of risk factors for first dose hypotension by an ACE-inhibitor. Therapie 1993; 48:479–482.
44. Stewart PW, Stamm JW. Classification tree prediction models for dental caries from clinical, microbiological, and interview data. Journal of Dental Research 1991; 70:1239–1251.
45. Devansh Shah · Samir Patel · Santosh Kumar Bharti SN. Heart Disease Prediction using Machine Learning Techniques, Computer Science (2020) 1:345, https://doi.org/10.1007/s42979-020-00365
46. Umarani Nagavelli,Debabrata Samanta , Partha Chakraborty. Machine Learning Technology-Based Heart Disease Detection Models, Hindawi, Journal of Healthcare Engineering, Volume 8, 2022, Article ID 7351061, 9 pages https://doi.org/10.1155/2022/735106147.
47. Brian Ayers, Tuomas Sandholm P,Igor Gosev, Sunil Prasad, Arman Kilic M. Using machine learning to improve survival prediction after heart transplantation, Journal of Cardiac Surgery, July 2021, Wiley, DOI: 10.1111/jocs.15917
48. Ahmed M. AlaaID, Thomas Bolton, Emanuele Di Angelantonio, James H.F. RuddI D, Mihaela van der Schaar. Cardiovascular disease risk prediction using automated machine learning, PLoS ONE, 14(5): e0213653. https://doi.org/10.1371
49. Hana H. Alalawi. Detection of Cardiovascular Disease using Machine Learning Classification Models, International Journal of Engineering Research & Technology (IJERT), http://www.ijert.org ISSN: 2278-0181, IJERTV10IS070091
50. Takahiro N, Soshiro O, Teruo Noguchi Healthcare delivery, economics and global health. https:// doi. org/ 10. 1136/ heartjnl- 2020-318726
51. Zaibunnisa L. H. Malik. Heart Disease Prediction using Artificial Intelligence, International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181, Published by, www.ijert.org, NREST – 2021
52. Ioannis A. Kakadiaris, Michalis Vrigkas, Albert A. Yen, Tatiana Kuznetsova, Matthew Budoff, Morteza Naghavi. Machine Learning Outperforms ACC/AHA CVD Risk Calculator in MESA, J Am Heart Assoc. 2018; 7:e009476. DOI: 10.1161/JAHA.118.009476.
53. Sunao Kojima, PhD; Takehiro Michikawa. Association of Fine Particulate Matter Exposure with Bystander-Witnessed Out-of-Hospital Cardiac Arrest, JAMA Network Open. 2020;3(4): e203043. doi:10.1001/jamanetworkopen.2020.3043
54. Chayakrit Krittanawong. Machine learning prediction in cardiovascular diseases: A meta-analysis Scientific Reports (2020)10:16057, Nature Research https://doi.org/10.1038/s41598-020-72685-1.
55. Kwon J-m et al. (2019) Deep-learning-based risk stratification for mortality of patients with acute myocardial infarction. PLoS ONE 14(10): e0224502. https://doi.org/10.1371/journal.pone.0224502
56. Joon-myoung Kwon, MD; Youngnam Lee, An Algorithm Based on Deep Learning for Predicting In-Hospital Cardiac Arrest, Journal of the American Heart Association, MSDOI: 10.1161/JAHA.118.008678
57. Peter C. Austin. A comparison of regression trees, logistic regression, generalizedadditive models, and multivariate adaptive regression splines for predicting AMI mortality STATISTICS IN MEDICINE, Statist. Med. 2007; 26:2937–2957 Published online 21 December 2006, Wiley Inter Science (www.interscience.wiley.com) DOI: 10.1002/sim.2770

How to cite this article:
Vidya Sagar, Gokul Rajan, Sriram Vaidyanathan, Deepak Batra, Rajaram Naganur, Anwesh Reddy ***Artificial Intelligence (AI) Models Predicting Cardiac Risk. Are The Developed Models Optimal In The Accuracy Of Clinical Prediction, Population- Specific And Robust? Necessity For Specific Risk Variables And New Hybrid Models.*** **Br J Bio Med Res, Vol.06, Issue 02, Pg.1997 - 2007, March - April 2022. ISSN:2456-9739 Cross Ref DOI : https://doi.org/10.24942/bjbmr.2022.962**

Source of Support: Great Learning,University of Texas,Austin. **Conflict of Interest:** None