DOI: https://doi.org/10.4038/jdrra.v2i1.27

The Journal of Desk Research Review and Analysis, Vol. 2, Issue 1, 2024, 58-66

IMPROVING CARDIOVASCULAR RISK PREDICTION OF SRI LANKANS USING ARTIFICIAL INTELLIGENCE

C Mettananda[1], MB Solangaarachchige[2], PS Haddela[3], AS Dassanayake[4], A Kasturiratne[5], AR Wickramasinghe[6], N Kato and HJ de Silva[7]

Abstract

There are no CV risk prediction models derived from Sri Lankan cohorts. Therefore, the World Health Organization(WHO) risk charts developed for the Southeast Asia Region are being used to risk stratify Sri Lankans. However, Sri Lankans are quite different to some Southeast Asian countries and may not agree with Sri Lankans. Therefore, we aimed to develop a CV risk prediction model specific to a cohort of Sri Lankans. Using supervised machine learning of 10-year follow-up data of a randomly selected, population-based cohort of Sri Lankans, we developed a model to predict the 10-year risk of developing a cardiovascular event. We compared predictions of the new model at baseline in 2007 with the observed events in 2017 following a 10-year follow-up using receiver operating characteristic curves(ROC) to find the predictive performance. We compared the predictions of the new model and the currently used WHO risk charts. We selected 2596 Sri Lankans between 40 and 65 years old with no history of previous CV diseases (CVD) at recruitment and who had completed 10-year follow-ups. There were 179 hard CVDs recorded over the ten years. CVD included all cardiovascular deaths confirmed or presumed cases as mentioned in death certificates, non-fatal strokes, and physician-diagnosed non-fatal acute coronary syndromes, including elective percutaneous coronary interventions and coronary artery bypass grafts done on patients with symptomatic unstable angina. Any cardiac presentation except those mentioned here was excluded. Of 179 events, the ML-based model predicted 124; only 33 were predicted by the new model, while only 33 were predicted by 2019 WHO risk charts. The new ML-based model had 0.93 accuracy with an AUC-ROC of 0.74 ± 0.06. Machine learning of individual data of a Sri Lankan cohort improved CV risk prediction of Sri Lankans than using risk charts developed for an epidemiological region using a modelling approach.

Keywords: Cardiovascular Risk, Prediction, Machine Learning, Artificial Intelligence, Sri Lanka

[1] Professor, Department of Pharmacology, University of Kelaniya, Sri Lanka.
Email: chamila@kln.ac.lk https://orcid.org/0000-0002-3328-1553 [1]
[2]Technical Officer (ICT), Faculty of Medicine, University of Kelaniya, Sri Lanka.
Email: maheeka@kln.ac.lk https://orcid.org/0009-0005-9166-4869
[3]Senior Lecturer, Faculty of Computing, SLIIT, Malabe, Sri Lanka,.
Email: prasanna.s@sliit.lk https://orcid.org/0000-0002-6969-8772
[4]Professor, Faculty of Medicine, University of Kelaniya, Sri Lanka.
Email: anuradhadassa@kln.ac.lk https://orcid.org/0000-0001-5567-3542
[5]Cadre Chair and Senior Professor, Faculty of Medicine, University of Kelaniya, Sri Lanka.
Email: arwicks@kln.ac.lk
[6]National Centre for Global Health and Medicine, Toyama, Shinjuku-ku,Tokyo, Japan
Email: nokato@hosp.ncgm.go.jp
[7]Emeritus Professor of Medicine, Faculty of Medicine, University of Kelaniya, Sri Lanka.
Email: hjanakadesilva@gmail.com

Received date: 25.07.2024
Accepted date: 22.10.2024

Introduction

Noncommunicable diseases (NCDs) account for more than 80% of deaths in Sri Lanka (Mendis, 2018). Cardiovascular diseases (CVD) are the leading cause of death globally and in Sri Lanka (Medical et al., 2019). Primary prevention of CVD depends on the aggressive management of CV risk factors. Risk factor control implemented through a high-risk strategy, where individuals are treated according to their total CV risk rather than on individual risk factor measure, is more cost-effective than a population-based strategy, especially for low-middle-income countries (Babu et al., 2012). Calculation of total CV risk requires reliable risk prediction models to identify individuals at high CV risk without failing.

However, no specific risk prediction models are derived from Sri Lankan cohorts or explicitly developed for Sri Lankans. Therefore, Sri Lankan physicians use risk prediction models derived from White Caucasians like Framingham general cardiovascular risk score (FRS), Atherosclerotic cardiovascular disease risk calculators (ASCVD), SCORE charts, etc. or the models developed for the whole of South East Asia Region (SEAR), like the World Health Organization (WHO) risk charts for risk prediction of Sri Lankans. These risk prediction models may not predict Sri Lankans' risk accurately as CV risk levels change with genetics, diet, lifestyle, socio-economic status, etc. of different nations(Mettananda et al., 2021; Ranawaka et al., 2016; Volkman et al., 2018).

In addressing the issue of some countries not having specific risk stratification models, the WHO collaboratively developed a set of CV risk perdition charts in 2007 for 14 epidemiological sub-regions using large cohort data instead of individual follow-up data and recalibrating them to estimate the 10-year risk of fatal or non-fatal major cardiovascular events for those sub-regions (WHO, 2007). WHO recently revised the risk charts in 2019 using individual-participant data from 85 prospective cohort studies and recalibrating the charts using age-specific and gender-specific incidence rates to suit 21 global epidemiological sub-regions (WHO, 2019). Therefore, WHO risk charts are the best available to date for risk stratification of Sri Lankans (Volgman et al., 2018). WHO/International Society of Hypertension (ISH) risk charts developed in 2007 have been validated among Sri Lankans, and it showed that it is predictive among 81% of the population but is less predictive in high-risk individuals and women (Thulani et al., 2021).
However, in the WHO epidemiological categorisation, Sri Lanka is categorised under South East Asia with Indonesia, Cambodia, Lao, Maldives, Myanmar, Malaysia, Philippines, Thailand, Timor-Leste, Viet Nam, Mauritius and Seychelles (WHO, 2019). The socio-economic and cultural backgrounds of these countries are different from those of Sri Lanka and, therefore, may not predict the CV risk of Sri Lankans perfectly. Moreover, these risk charts were developed using cohort data with a modelling approach, but individual patient follow-up data may be more informative and predictive.

Therefore, we aimed to develop a CV risk prediction model specific to a cohort of Sri Lankans using individual follow-up data of a randomly selected community cohort of Sri Lankans using artificial intelligence.

Methods

We used deidentified data from the Ragama Health Study (RHS), a prospective, community-based, follow-up study initiated in 2007 by the Ragama Medical Officer of Health (MOH) to obtain information on the prevalence of NCDs and risk factors that may be modified to formulate effective and efficient NCD intervention strategies, the protocol of which is previously published in the literature(Dassanayake et al.,

2009). Using stratified random sampling, RHS selected a random sample of Sri Lankans aged 35-65. Of this cohort, we selected participants who were 40 years or older, as most of the CVD risk calculators are developed for risk stratification of above 40-year-olds at the baseline in 2007 and completed a 10-year follow-up by July 2017 to develop a new risk prediction model. The initial database consisted of 770 variables with data on demographics, past medical history, family history, social history, anthropometry, blood pressure, laboratory investigations including fasting blood glucose, lipid profile, haemoglobin A_{1c} level, ECG and an ultrasound scan of the abdomen. Patients were followed up in 2010, 2014 and 2017 by the RHS and data on all hard CVDs, namely non-fatalnon-fatal strokes, non-fatalnon-fatal acute coronary syndromes, including elective coronary stenting and coronary artery bypass graft (CABG), fatal cardiovascular events developed during the 10-year follow-up period (documented in medical records or death certificates) had been collected and was available on the database. Following data wrangling and cleaning, 75 variables with less than 50% missing values and with a possibility to affect cardiovascular risk were selected for the model development. We imputed missing data using the mean/median appropriately, and the final database consisted of 75 variables that could have an impact on CV health.

We developed new risk prediction models with supervised machine learning (ML) of the follow-up cohort selected. We used the Train-Test Split method to develop the models. Google Colab ML platform with Scikit-learn library and Python programming language were used (9,10). We randomly split participants into two groups: the training and testing samples. The training sample was used to build the ML-based models, and the testing sample was used to assess the efficacy of the built algorithms. Predictions were studied against observed events using the receiver operating characteristic curve (ROC). The total CVD to No-CVDs ratio in the cohort was 7:93, and therefore, to prevent over-fitting with unbalanced data, we used 10-fold stratified cross-validation with ROC curves, and the mean area under the receiver operating characteristic curve (AUC-ROC) was taken as the AUC-ROC of the new ML-based model developed.

We trialled six standard ML classification algorithms with different modelling approaches: Decision tree, Random forest, k-nearest neighbour, 2D neural networks, AdaBoost and gradient boosting; the best-fitting model with the highest AUC-ROC was selected. Using the new ML-based model, we predicted the risk of developing a CVD over the next ten years of individual participants using baseline risk factor data in the new model. At the same time, we calculated predictions of the same using WHO 2019 risk charts. We compared the predictions of the new models and the WHO risk charts (Southeast Asia, 2019).

ERC approval for the RHS study was obtained from the ERC of the Faculty of Medicine, University of Kelaniya, Sri Lanka. Informed Witten consent was obtained from participants at the recruitment stage of the study. Data were analysed using SPSS version 28.

Results

Our study cohort consisted of 2596 participants who were followed up for ten years. There were 179 CV events recorded of the participants over the ten years. The cohort consisted of 1162 (44.8%) males, with a mean age of 53.5 (SD: 6.9) years in 2007 at baseline. Table 1 shows the baseline data of the cohort.

Table 1 Baseline Characteristics of the Cohort

	Male n = 1162	**Female n = 1434**	**Total n = 2596**
Ethnicity n (%)			
Sinhalese	1118 (96.2)	1375 (95.9)	2493 (96.0)
Tamil	15 (1.3)	27 (1.9)	42 (1.6)
Muslim	2 (0.2)	2 (0.1)	4 (0.2)
Burgher	15 (1.3)	19 (1.3)	34 (1.3)
Other	12 (1.0)	11 (0.8)	23 (0.9)
Age groups (years), n (%)			
40-49.9	360 (30.9)	456 (31.8)	816 (31.4)
50-59.9	526 (45.3)	669(46.7)	1195 (46.0)
≥ 60.0	276 (23.8)	309(21.5)	585 (22.6)
Smoking, n (%)	416 (35.8)	0 (0.0)	416 (16.0)
Smoking pack-years, median (IQR)	4.5 (2.0-9.0)	0	4.5 (2.0-9.0)
Hypertension, n (%)	185 (15.9)	336 (23.4)	521 (20.1)
On anti-hypertensive medications, n (%)	142 (12.2)	276 (19.2)	418 (16.1)
Diabetes, n (%)	165 (14.2)	249 (17.4)	414 (15.9)
On anti-diabetic medications, n (%)	129 (11.1)	193 (13.5)	322 (12.4)
Hyperlipidaemia, n (%)	98 (8.4)	209 (14.6)	307 (11.8)
On lipid-lowering medications, n (%)	49 (4.2)	141 (9.8)	190 (7.3)
Family history of premature CVD, n (%)	331 (28.5)	463 (32.3)	794 (30.6)
Engagement in recommended physical activities per week*, n (%)	255 (21.9)	335 (23.4)	590 (22.7)
Systolic blood pressure (mmHg), n (%)			
<139.9	766 (65.9)	869 (60.6)	1635 (62.9)
140-159.9	260 (22.4)	365 (25.5)	625 (24.1)
160-179.9	88 (7.6)	132 (9.2)	220 (8.5)
≥180.0	48 (4.1)	68 (4.7)	116 (4.5)
Total cholesterol (mmol/L), n (%)			

<4.0	211 (18.2)	207 (14.4)	418 (16.1)
4-4.9	297 (25.6)	269 (18.8)	566 (21.8)
5-5.9	391 (33.6)	476 (33.2)	867 (33.4)
6-6.9	192 (16.5)	322 (22.5)	514 (19.8)
7-7.9	66 (5.7)	123 (8.6)	189 (7.3)
≥8.0	5 (0.4)	37 (2.5)	42 (1.6)
BMI ≥ 23Kg/m^2, n (%)	590 (50.8)	945(65.9)	1535 (59.1)
BMI ≥ 30Kg/m^2, n (%)	47 (4.0)	166 (11.6)	213 (8.2)

CVD – cardiovascular diseases
BMI- body mass index
IQR – Interquartile range
*Strenuous physical activity ≥ 75min/week or moderate activity ≥ 150 min/week

Data from Randomly selected 2336 participants were used to create an ML-based model, and the data of the remaining 260 participants were used to test the model. Baseline data of participants who developed CVDs and did not develop CVDs were compared, and six ML-based models were developed. The Random Forest models showed the highest accuracy (0.9311), and AUC-ROC (0.74 ± 0.06) was selected as the final model. The feature importance of the variables is given in Table 2.

Table 2 Feature importance of variables studied

Variable	Importance
Age	0.08666
Current smoking	0.06200
Height	0.05601
Systolic blood pressure (mean)	0.05274
Duration of smoking	0.05246
Sex	0.05149
Sugar control for three months	0.03583
Hip circumference	0.03004
Diastolic blood pressure (mean)	0.02795
Serum triglyceride level	0.02524
Number of packs smoked a day	0.02387
History of hypertension	0.02246
Baseline insulin level	0.02220
Low-density lipoprotein cholesterol	0.02200
Fasting blood sugar	0.02166
Total cholesterol level	0.01910
Weight	0.01904
Alcohol use once a week or more	0.01901
Waist circumference	0.01798

Body mass index	0.01788
Aspartate amino transferase level	0.01781
HDL/LDL ratio	0.01648

HDL high-density lipoprotein level, LDL low-density lipoprotein level

Using the new model and the WHO risk charts, we predicted the expected CVD over the next ten years for the baseline cohort. Using a confusion matrix, we also compared the predictions and observed events between the two models. The new ML-based model predicted 124/179 cases while only 10/179 cases were predicted by WHO risk charts (Fig. 1). The ML-based model developed using data from a Sri Lankan cohort predicted 91 more cases than the latest WHO risk charts. However, the non-case prediction was equally good with the new ML-based model and the WHO risk charts, predicting 2293/2417 and 2372/2417 non-cases by the new ML-based model and WHO risk charts, respectively.

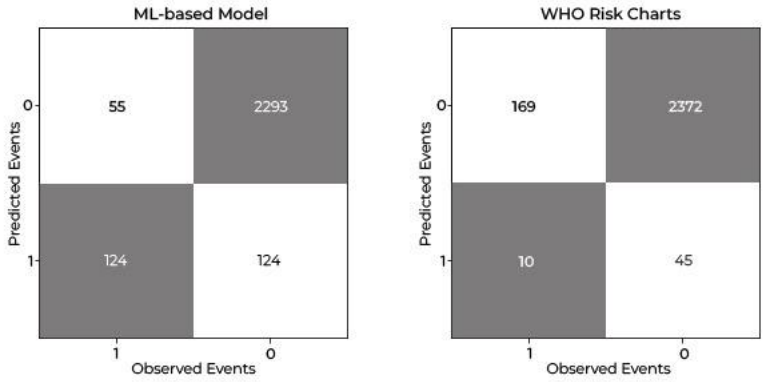

Figure 1 Comparison of the Predictions of the Machine-learning Model and the World Health Organization 2019 Cardiovascular Risk Charts

ML – machine learning, WHO – World Health Organization, CV - cardiovascular

Conclusion

The results show that the new ML-based risk prediction model derived using data from a cohort of Sri Lankans improved the prediction of high-risk individuals compared to the WHO risk charts. However, the new model could not predict 30.7% of events, indicating the need for using bigger databases to train the model to improve its accuracy further.

Discussion

This is the first CV risk prediction model developed using individual follow-up data of Sri Lankans and the only specific risk prediction model for Sri Lankans. We showed that the new model developed by machine learning of individual data of Sri Lankans is more predictive for Sri Lankans than the WHO model developed using estimates for different countries included in the Southeast Asia epidemiological region. The new model was more predictive in high-risk individuals. In contrast, the WHO risk score was all right in predicting low-risk individuals but ineffective in predicting high-risk individuals who should be the priority for CVD prevention interventions.

A model developed using individual follow-up data, and of individuals of Sri Lanka only being more predictive in Sri Lankans is no surprise. The fact that the WHO risk chats were meant for the whole of Southeast Asia and that those were developed by estimating from population-based data of Sri Lankans and not using individual patient data explains why the new model is better in predicting CVD risk in Sri Lankans than the WHO risk charts. There is evidence coming up showing the benefits of CV risk predictions developed by machine learning of all available data (Alaa et al., 2019; Chen et al., 2023; Liu et al., 2023; Mohd Faizal et al., 2021) than using traditional approaches considering few data and our findings also support the same. Machine learning models have been shown to achieve high accuracy in predicting cardiovascular diseases, outperforming traditional risk assessment methods (Alaa et al., 2019; Dalal et al., 2023; Peng et al., 2023; Weng et al., 2017; You et al., 2023). ML allows the models to appreciate subtle, complex interactions between variables in predicting outcomes rather than using conventional logistic regression. This identifies the real clinical associations rather than concentrating only on well-known risk factors. The study by Alaa et al. using the UK biobank data showed that the predictive capacity of the ML model when using all available 476 variables was better than that when using only the traditional variables (Alaa et al., 2019).

The potential of artificial intelligence in cardiovascular medicine is vast. However, rigorous validation and optimisation are necessary before clinical practice application. A recent meta-analysis of ML algorithms utilised for CVD prediction has highlighted the importance of using the optimal algorithm for the datasets being used due to the heterogeneity among ML algorithms (Krittanawong et al., 2020). A recent review on artificial intelligence (AI) and CV risk prediction has shown that AI-based predictive models may overcome some of the limitations of classic regression models. However, the successful application of AI requires knowledge of the potential pitfalls in AI techniques to guarantee their safe and effective use in daily clinical practice (Chiarito et al., 2022).

There are several strengths in our study. We studied data from a genuinely community-based random sample of Sri Lankans at the right age to develop CVDs. Participants were prospectively followed up with for ten years, and individual outcome data was studied. The dropout rate was very low, and only the data of participants who completed 10-year follow-ups were used to develop the ML models. The outcome endpoints were objective, and therefore, bias was minimal.

However, our study has some limitations. Our study cohort was drawn from a semi-urban area without representation from the estate sector. According to the 2012 census, Sri Lanka comprises urban-18.2%, rural-77.4% and estate sectors-4.4%(Central-bank-of-Sri-Lanka, 2022; Medical et al., 2019). However, as the estate sector is a very small percentage, our findings could be generalisable to Sri Lanka.

Conclusion

We showed that an ML-based risk prediction model developed using individual follow-up data from a Sri Lankan cohort improved the CV risk prediction of Sri Lankans more than the WHO risk charts developed for the Southeast Asia region. However, these models need to be validated and refined before using in clinical practice. We plan to validate the new ML-based model we developed in an external cohort, fine-tune it and develop a mobile app for easy use in future.

Acknowledgement

We thank all study participants and the participants of Ragama Health Study and Research Council of the University of Kelaniya for funding this study. The study was supported by the Strengthening Research Outputs Grant of the University of Kelaniya, Sri Lanka (RC/SROG/2021/01).

References

Alaa, A. M., Bolton, T., Di Angelantonio, E., Rudd, J. H. F., & van der Schaar, M. (2019). Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants. *PLoS One*, *14*(5), e0213653. https://doi.org/10.1371/journal.pone.0213653

Babu, R. B., Alam, M., Helis, E., & Fodor, J. G. (2012). Population-based versus high-risk strategies for preventing cardiovascular diseases in low- and middle-income countries. *Indian Heart Journal*, *64*(5), 439-443. https://doi.org/10.1016/j.ihj.2012.08.001

Central-bank-of-Sri-Lanka. (2022). *SRI LANKA SOCIO ECONOMIC DATA 2022*. https://www.cbsl.gov.lk/sites/default/files/cbslweb_documents/publications/otherpub/publication_sri_lanka_socio_economic_data_folder_2022_e.pdf

Chen, S.-F., Loguercio, S., Chen, K.-Y., Lee, S. E., Park, J.-B., Liu, S., Sadaei, H. J., & Torkamani, A. (2023). Artificial Intelligence for Risk Assessment on Primary Prevention of Coronary Artery Disease. *Current Cardiovascular Risk Reports*, *17*(12), 215-231. https://doi.org/10.1007/s12170-023-00731-4

Chiarito, M., Luceri, L., Oliva, A., Stefanini, G. G., & Condorelli, G. (2022). Artificial Intelligence and Cardiovascular Risk Prediction: All That Glitters is not Gold. *European Cardiology Review 2022;17:e29.* https://doi.org/10.15420/ecr.2022.11

Dalal, S., Goel, P., Onyema, E. M., Alharbi, A., Mahmoud, A., Algarni, M. A., & Awal, H. (2023). Application of Machine Learning for Cardiovascular Disease Risk Prediction. *Computational Intelligence and Neuroscience*, *2023*(1), 9418666. https://doi.org/https://doi.org/10.1155/2023/9418666

Dassanayake, A. S., Kasturiratne, A., Rajindrajith, S., Kalubowila, U., Chakrawarthi, S., De Silva, A. P., Makaya, M., Mizoue, T., Kato, N., Wickremasinghe, A. R., & de Silva, H. J. (2009). Prevalence and risk factors for non-alcoholic fatty liver disease among adults in an urban Sri Lankan population. *J Gastroenterol Hepatol*, *24*(7), 1284-1288. https://doi.org/10.1111/j.1440-1746.2009.05831.x

Krittanawong, C., Virk, H. U. H., Bangalore, S., Wang, Z., Johnson, K. W., Pinotti, R., Zhang, H., Kaplin, S., Narasimhan, B., Kitai, T., Baber, U., Halperin, J. L., & Tang, W. H. W. (2020). Machine learning prediction in cardiovascular diseases: a meta-analysis. *Scientific Reports*, *10*(1), 16057. https://doi.org/10.1038/s41598-020-72685-1

Liu, W., Laranjo, L., Klimis, H., Chiang, J., Yue, J., Marschner, S., Quiroz, J. C., Jorm, L., & Chow, C. K. (2023). Machine-learning versus traditional approaches for atherosclerotic cardiovascular risk prognostication in primary prevention cohorts: a systematic review and meta-analysis. *European Heart Journal - Quality of Care and Clinical Outcomes*, *9*(4), 310-322. https://doi.org/10.1093/ehjqcco/qcad017

Medical Statistics Unit MoH. (2019). *Annual Health Statistics - 2019*. http://www.health.gov.lk/moh_final/english/public/elfinder/files/publications/AHB/2020/AHB%202019.pdf

Mendis, R. S. S. (2018). *Prevention and Control of Noncommunicable Diseases Think Globally-Act Locally; Lessons from Sri Lanka*. www.health.gov.lk/moh_final/english/public/elfinder/files/Downloade/NCDbook2018.pdf

Mettananda, K. C. D., Gunasekara, N., Thampoe, R., Madurangi, S., & Pathmeswaran, A. (2021). Place of cardiovascular risk prediction models in South Asians; agreement between Framingham risk score and WHO/ISH risk charts. *Int J Clin Pract*, *75*(7), e14190. https://doi.org/10.1111/ijcp.14190

Mohd Faizal, A. S., Thevarajah, T. M., Khor, S. M., & Chang, S.-W. (2021). A review of risk prediction models in cardiovascular disease: conventional approach vs. artificial intelligent approach. *Computer Methods and Programs in Biomedicine*, *207*, 106190. https://doi.org/https://doi.org/10.1016/j.cmpb.2021.106190

Peng, M., Hou, F., Cheng, Z., Shen, T., Liu, K., Zhao, C., & Zheng, W. (2023). Prediction of cardiovascular disease risk based on major contributing features. *Scientific Reports*, *13*(1), 4778. https://doi.org/10.1038/s41598-023-31870-8

Ranawaka, U., Wijekoon, N., Pathmeswaran, P., Kasturiratne, A., Gunasekara, D., Chackrewarthy, S., Kato, N., & Wickremasinghe, R. (2016). Risk estimates of cardiovascular diseases in a Sri Lankan community. *Ceylon Med J*, *61*, 11. https://doi.org/10.4038/cmj.v61i1.8253

Thulani, U. B., Mettananda, K. C. D., Warnakulasuriya, D. T. D., Peiris, T. S. G., Kasturiratne, K., Ranawaka, U. K., Chakrewarthy, S., Dassanayake, A. S., Kurukulasooriya, S. A. F., Niriella, M. A., de Silva, S. T., Pathmeswaran, A. P., Kato, N., de Silva, H. J., & Wickremasinghe, A. R. (2021). Validation of the World Health Organization/ International Society of Hypertension (WHO/ISH) cardiovascular risk predictions in Sri Lankans based on findings from a prospective cohort study. *PLoS One*, *16*(6), e0252267. https://doi.org/10.1371/journal.pone.0252267

Volgman, A. S., Palaniappan, L. S., Aggarwal, N. T., Gupta, M., Khandelwal, A., Krishnan, A. V., Lichtman, J. H., Mehta, L. S., Patel, H. N., Shah, K. S., Shah, S. H., & Watson, K. E. (2018). Atherosclerotic Cardiovascular Disease in South Asians in the United States: Epidemiology, Risk Factors, and Treatments: A Scientific Statement From the American Heart Association. *Circulation*, *138*(1), e1-e34. https://doi.org/doi:10.1161/CIR.0000000000000580

Weng, S. F., Reps, J., Kai, J., Garibaldi, J. M., & Qureshi, N. (2017). Can machine learning improve cardiovascular risk prediction using routine clinical data? *PLoS One*, *12*(4), e0174944. https://doi.org/10.1371/journal.pone.0174944

WHO. (2007). *World Health organisation/International Society of Hypertension risk prediction charts for 14 WHO epidemiological sub-regions*. WHO. https://doi.org/https://www.who.int/ncds/management/WHO_ISH_Risk_Prediction_Charts.pdf?ua=1

WHO. (2019). World Health Organization cardiovascular disease risk charts: revised models to estimate risk in 21 global regions. *Lancet Glob Health*, *7*(10), e1332-e1345. https://doi.org/10.1016/s2214-109x(19)30318-3

You, J., Guo, Y., Kang, J.-J., Wang, H.-F., Yang, M., Feng, J.-F., Yu, J.-T., & Cheng, W. (2023). Development of machine learning-based models to predict 10-year risk of cardiovascular disease: a prospective cohort study. *Stroke and Vascular Neurology*, *8*(6), 475-485. https://doi.org/10.1136/svn-2023-002332