

Predicting The Heart Attack From Accessible Patients Medical Datasets Using Data Mining Technique

S. Ajmal Mohamed

Mr. Balamurali

Prist University, Puducherry Campus, Deemed University,
Pondichery, India

Abstract: Today medical services had come a long way to treat a patient with different diseases. Among them, one leading disease is myocardial infarction. Which cannot be seen through the naked eyes. It comes immediately when it has reached its limitations. The cost to treat this problem is very high and cannot affordable to all patients. So this system typically create a huge amount of datasets. The creation of dataset is taken as input from the patients and generate a result using kmean clustering algorithm. Clustering is the process of portioning a group of data points into a small number of clusters. A quantitative approach would be to measure certain features of the risk prediction and predict the percentage. The goal is to assign a cluster to each data points. The number of clusters should match the data. An incorrect choice of the number of cluster will invalidate the whole process. An empirical way to find the best number of clusters is to try kmean clustering with different number of clusters and measure the result. We have enhanced this system by adding the features of datasets with 12 attribute for predicting the risk (blood pressure, lack of exercise, poor diet, heredity, blood test report etc..). Finally this system will predict the heart attack with remedies for overcoming it and with suggestion from the specialist.

Index Terms: Classification, clinical decision support system, clinical risk prediction, medical screening, myocardial infarction.

I. INTRODUCTION

The best practice to avoid human mortality caused by life threatening diseases like myocardial infarction (MI) is to detect them early and prevent their onset. One approach is to devise computational methods that capitalize on clinical biomarkers to better screen the patients for their potential risk of experiencing (future) MI. Broadly, clinical screening/risk prediction tools are very important as it could potentially lead to the following benefits at the individual patient-level: for example, 1) when patients become knowledgeable of their health risk and with good physician-patient therapeutic relationship, they would be more willing to make changes to their lifestyle and adhere to treatment regimens [1], 2) allows clinicians to promptly recommend effective therapeutic or preventive measures (e.g., lifestyle changes, treatment of

subclinical manifestation, etc.) to their patients [2], and 3) if such screening tools were to be integrated into electronic health record system and executed automatically to analyze individuals health risk, the number of unscreened patients who are at risk of a disease could be reduced dramatically [3]. The key ramification of wide adoption of clinical screening tools is the possibility of significantly reducing the number of avoidable mortality. However, the development of versatile, reliable, and accurate computer aided MI screening tools which the clinicians can use in the clinics/hospitals to instantly predict patients risk remains a challenge.

The conventional approaches for assessing the risk of individuals experiencing MI include risk scoring system and survival curves [4][6]. These, however, have limitations like the inability to substantially identify minority of individuals with subsequent risk of experiencing MI [7]. Moreover, clinical

biomarkers and symptoms seldom follow a linear relationship and the expected outcome at the individual patient-level does not always abide by the rules of epidemiology [8]. As a result, conventional risk scoring systems which model relationships in a linear manner often flounder in view of these challenges [9], [10].

In recent years, there is an exponential increase in the amount of clinical and molecular data collected from routine medical examination. To overcome the challenges associated with human scale of thinking and analysis, data mining techniques which have been postulated as a central feature for future health-care system [11] became a popular method for extracting insights from this data deluge. Advantages of using data mining techniques include the capability of dealing with plethora of information, solving nontrivial problems, producing data-driven prediction models, and handling nonlinear relationships among biomarkers. Examples of data mining techniques used to estimate disease risk include work from: 1) Wiens et al. [12] who employed support vector machine (SVM) to identify patients who are at high risk of experiencing hospital acquired *Clostridium difficile* (*C. diff*); and 2) Khan et al. [13] who used artificial neural network (ANN) for discriminating small, round blue-cell tumors (SRBCTs).

One important component of risk prediction tools is to provide clinicians with the flexible to customize (e.g., change the range and how far into the future the prediction would be) and use a risk prediction model that they deemed most beneficial for their patients. To this end, we explore the possibility of customizing MI risk prediction models to better meet the patients needs and clinicians expectation. Particularly, the effect of sample age and prediction resolution two aspects that are not commonly examined in the literature on the performance of MI risk prediction models constructed using SVM [14][16] and evolutionary data-conscious artificial immune recognition system (EDC-AIRS) [17] algorithms were investigated. Here, sample age refers to the average age of individuals found in the baseline (i.e., input) dataset used to construct the clinical risk prediction model while prediction resolution refers to the prediction scale (i.e., number of years into the future where prediction of MI occurrence begins) and interval (i.e., time duration, in years, that marks the start and end of MI outcomes to be considered) employed by the clinical risk prediction model. In the view of the rapid aging population worldwide and the relatively high prevalence of MI among the elderly, participants amassed from the Cardiovascular Health Study (CHS) [18] consisting of subjects aged 65 and above were analyzed. Further, with the wide range of clinical measurements and risk factors accrued during the CHS observational study, it makes the CHS dataset a valuable source of information for this paper. The rest of the paper is organized as follows. Section II provides details of CHS dataset, and delineates the methodology involved in developing the predictive models. Section III provides the experimental results achieved by the risk prediction models developed using different combinations of sample age and prediction resolution. Key results are discussed in Section IV and conclusions are drawn in Section V.

II. MATERIALS AND METHODS

In Section II-A, details of CHS dataset are provided. This dataset, however, consists of a significant percentage of missing data and a highly skewed data distribution (commonly known as the class imbalanced data problem). Hence, for effective analysis, data imputation and class data balancing are performed and described in Sections II-B and II-C, respectively. Section II-D explains how the various MI risk prediction models based on different combinations of baseline data and prediction resolution were developed and validated.

A. CHS DATASET

The CHS dataset, as described in [18], is an epidemiology study of risk factors for cardiovascular diseases in elderly aged 65 and above. It contains two cohorts recruited at different phases. The first cohort consists of 5201 subjects from four U.S. communities, namely Forsyth County, North Carolina; Sacramento County, California; Washington County, Maryland; and Pittsburgh, Pennsylvania. An additional 687 African Americans were subsequently recruited forming the second cohort. Eligible individuals were sampled from Medicare eligibility lists in each area. Eligible participants include all individuals sampled from the Health Care Financing Administration (HCFA) sampling frame they were 65 years or older at the time of examination, non-institutionalized, expected to remain in the area for the next 3 years, and able to give informed consent and did not require a proxy respondent at baseline. Individuals who were wheelchair bound at home at baseline or receiving hospice treatment, radiation therapy or chemotherapy for cancer were excluded. Eligible individuals were examined yearly from

B. DATA IMPUTATION

Data imputation is the process of substituting missing entries in a dataset with plausible values and aims to improve the quality of the data. It was performed using weighted K-nearest neighbor (KNN) because of its excellent performance in estimating missing values [22], [23]. Moreover, it has the capability to estimate both qualitative and quantitative attributes. Hence, it is highly suitable for interpolating the missing values in the CHS dataset. Individuals with unknown MI status and clinical features that were uninformative (i.e., features with consistent value through-out) were first removed from the analysis. Individuals and clinical features with high percentage of missing entries were also removed. This is to ensure that there is an adequate supply of complete entries for weighted KNN to reference when estimating the missing values, which in turn promotes a more accurate data imputation process [22][24]. The resulting dataset was next normalized to unit variance to ensure that the attributes with large scale do not dominate the (Euclidean) distance measure [25]. Subsequently, the optimal value of K for each clinical feature was determined by 10-fold cross validation and used for the data imputation process. The type of replacement method used by weighted KNN depends on the data type. For instance, if categorical (continuous) data were encountered, the weighted-mode (weighted-mean) of the KNN was used to assign the value

for the missing entries. The use of weighted KNN estimation has been demonstrated in [22], [26] to be robust and accurate.

C. CLASS IMBALANCE DATA PROBLEM

In order to create an unbiased dataset for SVM and EDC-AIRS algorithms to learn from, under-sampling of the majority class is necessary. The Kennard Stone (KS) algorithm [27] was employed to perform this task because of its excellent performance as demonstrated in a comparative study [28]. This algorithm sequentially selects representative data that are uniformly scattered across the data domain space. This is carried out by first selecting the data object that is closest to the mean of the dataset and is included as the first data candidate. Subsequently, the data object that is most distant from the first one (based on Euclidean distance) is included as the second data candidate. The next data object is chosen by identifying the one farthest away from the previously selected data candidates. This process repeats until the desired number of candidates has been identified [28], [29].

In this study, the KS algorithm was used to under-sample the majority class found in the imputed CHS dataset. The number of candidates to select is equivalent to the number of samples in the minority class. In other words, after this process, the number of controls and cases would be identical.

D. MI RISK PREDICTION MODELS

Two algorithms (SVM and EDC-AIRS) were employed to develop MI risk prediction models. SVM algorithm is a robust supervised learning algorithm that is capable of yielding excellent generalization performance on an extensive area of problems [30][32]. It is derived from statistical learning theory and is capable of solving linearly and nonlinearly separable problems. Fundamentally, SVM performs classification through the construction of an N-dimensional hyperplane that optimally separates the data into two or more categories whereby the margin of separation between the different categories is maximized. EDC-AIRS algorithm [17] is a supervised classification algorithm inspired by the principles and processes associated with the human immune system. It performs classification by first constructing a pool of memory cells (i.e., candidate solutions in the form of data vectors) that are representative of the training data through repetitive optimization of the (values of the) memory cells. Optimization was carried out by robustly adapting the memory cells to the different density, distribution, and characteristics exhibited by each data class in the training data. Finally, with the utilization of the generated memory cells pool, KNN is used to classify unseen data observations. This algorithm, when tested on several widely benchmarked datasets, has demonstrated highly competitive classification performance [17]. To adopt a ceteris paribus experimental design, the parameters for both algorithms were first tuned using genetic algorithm (GA) and subsequently, feature selection was conducted (using GA) to identify predictive biomarkers. The GA parameters were determined experimentally to work well with this clinical prediction problem and kept constant for all experiments. The setup details of GA are as follows: population size: 100; maximum

generation: 100; natural selection: stochastic universal sampling; crossover type: discrete recombination; crossover probability: 0.8; mutation rate: $1/P$, where P is the number of parameters/features. The parameter details for SVM are kernel function: radial basis function (RBF); cost: $[2^{-5}, 2^{13}]$; gamma:

$[2^{-15}, 2^3]$; and for EDC-AIRS are seed: 1; clonal rate: 10; hypermutation rate: 2; stimulation threshold: 0.9; initial memory pool size: [0, 200]; KNN value: [1, 15]; affinity threshold scalar: [0, 1]; total resource: [150, 300]; Radiusdensity = [0, 3]; Radiusmax = [0, 3].

Clinical data recorded during the 5th to 11th year in which the CHS clinical study was undertaken were utilized. The reason for using clinical data recorded from year 5 onward was because clinical examinations taken by the two different cohorts recruited at different phases synchronized from that year onward. The reason for ending the prediction at year 11 is because from year 12 onward, participants were only monitored annually via phone calls and no clinical examinations were conducted.

To test the hypothesis, prediction models using different baseline datasets (with different sample age) capable of predicting the risk of experiencing MI at various prediction scales and intervals were developed. As illustrated in Fig. 1, eight different prediction models were designed to investigate how time factor in relation to the onset of MI would affect the performance of the prediction model. Three different baseline datasets were used. These datasets contain clinical examination results recorded in year 5, year 7, and year 9 of the CHS study. Each of these datasets was used to predict future. Three different prediction scales (1, 3, and 5 years) and three different prediction intervals (2, 4, and 6 years) were investigated. Specifically, healthy individuals present in year 5 of the CHS dataset were used as the baseline data to predict whether one would experience MI from year 6 to 11 (prediction scale: 1 year; prediction interval: 6 years), year 6 to 7 (prediction scale: 1 year; prediction interval: 2 years), year 8 to 9 (prediction scale: 3 years; prediction interval: 2 years) and year 10 to 11 (prediction scale:

5 years; prediction interval: 2 years). Similarly, clinical examination results of healthy participants in year 7 was initialized as the baseline data, where prediction of whether one would suffer from MI whether an individual would experience MI in the near from year 8 to 11, year 8 to 9, and year 10 to 11 were conducted. Likewise, clinical data recorded in year 9 were utilized to perform prediction of MI occurrence from year 10 to 11.

Each baseline dataset was randomly split into two subsets having balanced class distribution. The first subset contains 70% of the initial data. Using this subset, the prediction model was trained and tuned based on 10-fold cross validation. The second subset, which contains the remaining 30% of the data, was used to validate the developed model. This splitting process was repeated three times and independently used to develop and test the respective prediction model. It is highly encouraged to do so to avoid the developed model from capturing not only the true associations, but, also, idiosyncratic features of the training data, which often produces an overly optimistic model [33]. Three commonly

used performance measurements were employed to evaluate the prediction models developed namely sensitivity, specificity, and balanced accuracy (i.e., average between sensitivity and specificity).

Finally, to determine whether the prediction models developed using SVM and EDC-AIRS algorithms are statistically different from each other, McNemar's test was conducted. This statistical test was chosen as it has been demonstrated to have low type-1 error [34]. For each prediction model, this test was carried out by first recording the prediction outcomes obtained (by each algorithm) when tested using each validation dataset. The results obtained from each algorithm were then used to algorithms.

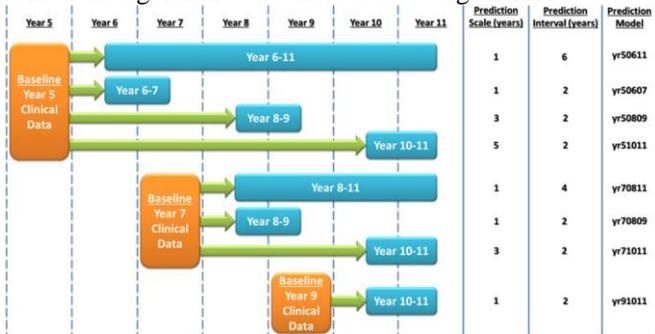


Figure 1: MI risk prediction models of various prediction scales and intervals. MI risk prediction at various time scales and intervals using the CHS dataset was performed. Prediction scale refers to the number of years into the future where prediction of MI occurrence begins while prediction interval refers to the time duration (in years) that marks the start and end of MI outcomes to be considered

		EDC-AIRS	
		Misclassification	Correct Classification
SVM	Misclassification	a	b
	Correct Classification	c	d

Figure 2: Contingency table for McNemar's test. "a" indicates the number of data items misclassified by both SVM and EDC-AIRS algorithms; "b" represents the number of data items misclassified by SVM algorithm but correctly classified by EDC-AIRS algorithm; "c" denotes the number of data items misclassified by EDC-AIRS algorithm but correctly classified by SVM algorithm; "d" dictates the number of data items correctly classified by both SVM and EDC-AIRS

construct the contingency table shown in Fig. 2. Referring to the figure, if the sum of b and "c" is greater than 25, chi-square test with 1 degree of freedom is used for performing McNemar's test. Otherwise, to provide a better estimation of small sample (i.e., $b + c \geq 25$), binomial distribution is used for (exact) McNemar's test. The prediction model is considered to be statistically different from the ground truth if the p-value computed using McNemar's test is smaller than 0.05.

III. EXPERIMENTAL RESULTS

A. DATA PREPROCESSING

Table I provides the details of the resulting CHS datasets

after the removal of records and clinical features with significant missing entries.

Table II offers the details of the datasets used to develop and test the MI prediction models after data imputation and class data balancing were performed.

Prediction Model	Sample Size* (Cases/Controls)	#Features	Age (Mean±SD)
yr50611	3102 (6.2%/93.8%)	237	75.7 ± 5.34
yr50607	3102 (2.4%/97.6%)	237	75.7 ± 5.34
yr50809	3034 (2.1%/97.9%)	237	75.7 ± 5.34
yr51011	2978 (2.1%/97.9%)	237	75.7 ± 5.36
yr70811	2407 (2.1%/97.9%)	233	77.2 ± 5.40
yr70809	2407 (2.1%/97.9%)	233	77.2 ± 5.40
yr71011	2362 (2.0%/98.0%)	233	77.2 ± 5.40
yr91011	1909 (1.9%/98.1%)	242	78.8 ± 5.09

This sample size refers to the number of individuals that remain in the CHS dataset after removal of records with significant missing entries. yrXYYZZ denotes that the prediction model uses clinical measurements observed in year X to make prediction of whether one would experience MI from year YY to ZZ.

Table 1: Details of The Imputed Chs Dataset

Prediction Model	#Training Instances	#Validation Instances	McNemar's Test [†]
			(p-value)
			SVM versus EDC-AIRS
yr50611	270	114	< 0.01
yr50607	104	42	< 0.01
yr50809	92	38	< 0.01
yr51011	88	36	< 0.01
yr70811	136	58	0.31
yr70809	70	30	0.04
yr71011	66	28	< 0.01
yr91011	52	20	0.07

All training and validation datasets contain equal number of cases and controls.

#The p-value of McNemar's test is presented examining whether the performance of the SVM algorithm is statistically different from EDC-AIRS algorithm.

Table 2: Details of Datasets Used To Build the Prediction Models

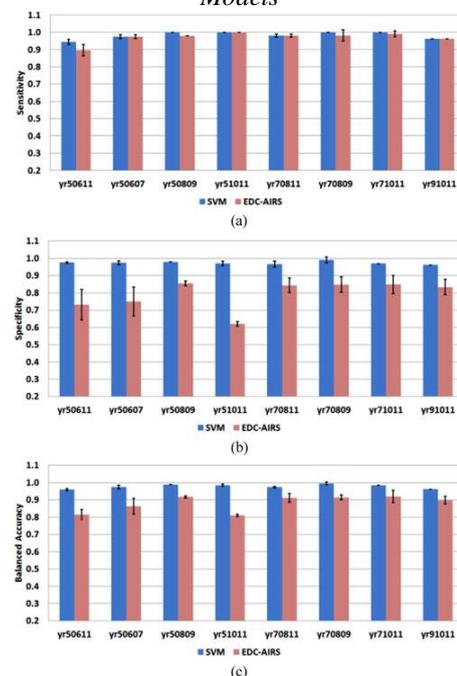


Figure 3: Classification performance of SVM and EDC-AIRS algorithms (cross validated). (a) Sensitivity performance metric. (b) Specificity performance metric. (c) Balanced accuracy performance metric. These performance measurements were obtained by performing 10-fold cross validation for each prediction model

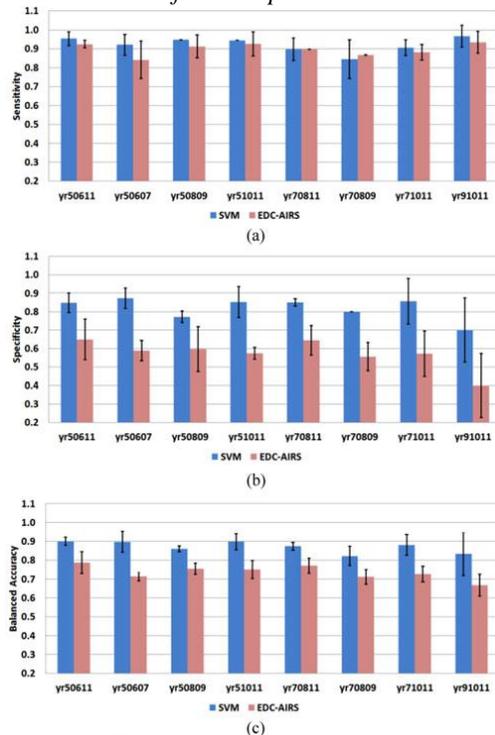


Figure 4: Classification performance of SVM and EDC-AIRS algorithms (tested with validation dataset). (a) Sensitivity performance metric. (b) Specificity performance metric. (c) Balanced accuracy performance metric. These performance measurements were obtained by evaluating each developed prediction model with their respective validation dataset.

B. MI RISK PREDICTION MODELS

Prediction models using baseline dataset with different sample age at various time scales and intervals were developed using the training datasets. Cross validation was carried out to evaluate the performance of each prediction model. For all prediction models developed, results (as shown in Fig. 3) indicate consistently high predictive performance was achieved by both SVM and EDC-AIRS algorithms. For example, a balanced accuracy of at least 0.95 and 0.81 was achieved by SVM and EDC-AIRS algorithms, respectively.

To assess whether the prediction models developed generalize well, validation was performed using the validation datasets. Results, as presented in Fig. 4, demonstrate that a balanced accuracy of at least 0.81 and 0.71 was achieved by SVM and EDC-AIRS algorithms, respectively.

McNemar's test was conducted to determine whether the performance of SVM and EDC-AIRS algorithms is statistically different from each other. Results (as shown in Table II) indicate that for most of the prediction models (except prediction models yr70811 and yr91011), the performance of SVM and EDC-AIRS algorithms are statistically different.

IV. DISCUSSION

MI risk prediction models developed using baseline datasets with different sample age, and based on different prediction resolution combinations were analyzed. Cross validation was utilized during the training phase as an approach to evaluate and develop potent MI risk prediction models. The resultant prediction models developed by both algorithms achieved a relatively high sensitivity, specificity, and balanced accuracy (for SVM algorithm, the respective performance achieved is at least 0.94, 0.96, and 0.95; while for EDC-AIRS algorithm, the respective performance achieved is at least 0.89, 0.62, and 0.81). An investigation of whether the prediction models developed were overtrained was conducted by validating each developed model with an unseen dataset (i.e., not used to develop the prediction model). The aim of this step was to assess the generalizability of the developed models. Results indicate that SVM algorithm (and EDC-AIRS algorithm) across all prediction models tested—achieved a sensitivity, specificity, and balanced accuracy of at least 0.84, 0.70, and 0.82 (and 0.84, 0.40, and 0.67), respectively. Furthermore, it can be observed that in general there is a drop in the validation sensitivity (SVM: 0.060 ± 0.054 ; EDC-AIRS: 0.073 ± 0.052), specificity (SVM: 0.154 ± 0.058 ; EDC-AIRS: 0.219 ± 0.124), and balanced accuracy (SVM: 0.107 ± 0.036 ; EDC-AIRS: 0.146 ± 0.070) among all the prediction models developed. It is noteworthy that the drop in performance is less severe for SVM algorithm (when compared to EDC-AIRS algorithm). This shows that SVM algorithm tends to perform better on noisy data even after data imputation was conducted. This observation is supported by the results obtained from the performance of McNemar's test. From this statistical evaluation, it was demonstrated that SVM algorithm outperforms EDC-AIRS algorithm for six out of eight prediction models tested.

Prediction models developed (with SVM algorithm) using baseline dataset from year 5 (and year 7), and tested using their respective validation datasets have shown comparable sensitivity, specificity, and balanced accuracy. Analysis of variance (ANOVA) test was conducted on the respective group of prediction models (i.e., developed using either year 5 or 7 as baseline dataset) that has a prediction interval of 2 years. Results demonstrate that they are statistically comparable with p-value of

0.47 for prediction models using baseline dataset from year 5 (and 0.25 for prediction models using baseline dataset from year 7). This signifies that prediction scale does not have a significant impact on the performance of (SVM-based) prediction models developed and tested using subjects aged 65 and above. Similar analysis was performed on prediction models developed based on different prediction interval. Results indicate that these models are statistically comparable with p-value of 0.92 and 0.88 for prediction models developed using baseline dataset from year 5 and 7, respectively. This means that prediction interval does not have a significant impact on the performance of prediction models developed using SVM algorithm.

As for prediction models developed using EDC-AIRS algorithm, similar analysis was conducted. For prediction models developed using baseline dataset from year 5 (and year

7) that are based on 2-year prediction interval, and tested using their respective validation datasets, ANOVA test was conducted. Results indicate that the prediction models in their respective group are statistically comparable having a p-value of 0.71 (for prediction model using year 5 baseline dataset) and 0.93 (for prediction model using year 7 baseline dataset). This indicates that prediction scale does not have a significant impact on prediction models developed using EDC-AIRS algorithm as well. Likewise, prediction models developed based on different prediction interval were analyzed. Results show that these models are statistically comparable having a p-value of 0.12 and 0.14 for prediction models developed using baseline dataset from year 5 and 7, respectively. This suggests that prediction interval does not have a significant impact on the performance of prediction models developed using EDC-AIRS algorithm as well. A summary of the p-values discussed is provided in Table III.

Prediction Models Compared	ANOVA Test [#] (p-value)	
	SVM	EDC-AIRS
Prediction Scale		
yr50607; yr50809; yr51011	0.47	0.71
yr70809; yr71011	0.25	0.93
Prediction Interval		
yr50611; yr50607	0.92	0.12
yr70811; yr70809	0.88	0.14

Table 3: Statistical Evaluation of Prediction Resolution

Analysis of prediction models that aim to predict the likelihood of MI occurrence in individuals subsequent 2 years (i.e., yr50607," yr70809," and yr91011) indicate comparable performance with p-value of 0.50 and 1.00 for SVM and EDC-AIRS algorithms, respectively. Comparison of age among individuals belonging to different baseline datasets indicates that they are statistically different (p-value < 0.01). This portends that sample age does not have a significant impact on the performance of prediction models.

Among all the prediction models developed, key biomarkers identified to be statistically significant by both SVM and EDC-AIRS algorithms are related to cognitive function, physical function, depression/life events, electrocardiography, general changes to health/lifestyle, and medications. These biomarkers, in general, are also identified as clinically significant in the literature [35][38]. This suggests that statistically significant biomarkers can also be clinically significant providing a promising avenue for identifying the potential cardiovascular risk factors to be evaluated in clinical trials.

One benefit of performing risk prediction using different prediction resolution and sample age is that it allows more refined and progressive risk prediction to be conducted (without compromising accuracy). This provides the advantage of estimating the seriousness of a disease one is experiencing; enabling clinicians to offer a more personalized management and/or therapeutic strategy to the patient.

The limitation of this investigation includes the use of a single dataset to evaluate the effect of sample age and prediction resolution in relation to the performance of MI risk prediction. This limits the power to conclusively state how each factor influences the performance of the prediction model. Nevertheless, it does provide some insights on whether

sample age and prediction resolution have an impact on the performance of clinical risk prediction model. In view of the observations from this study and the importance of screening since young, we aim to investigate the effect of prediction resolution and sample age on younger subjects as part of our future work.

V. CONCLUSION

Early detection of individuals with high risk of experiencing MI is very important clinically, but has proved to be elusive. To this end, we investigated the effect of sample age and prediction resolution in relation to the development of accurate clinical risk prediction model. Our experiments indicate that both sample age and prediction resolution do not have a significant impact on prediction models developed using subjects aged 65 and above. Table performance with p-value of 0.50 and 1.00 for SVM and EDC-AIRS algorithms, respectively. Comparison of age among individuals belonging to different baseline datasets indicates that they are statistically different (p-value < 0.01). This portends that sample age does not have a significant impact on the performance of prediction models.

Among all the prediction models developed, key biomarkers identified to be statistically significant by both SVM and EDC-AIRS algorithms are related to cognitive function, physical function, depression/life events, electrocardiography, general changes to health/lifestyle, and medications. These biomarkers, in general, are also identified as clinically significant in the literature [35][38]. This suggests that statistically significant biomarkers can also be clinically significant providing a promising avenue for identifying the potential cardiovascular risk factors to be evaluated in clinical trials.

One benefit of performing risk prediction using different prediction resolution and sample age is that it allows more refined and progressive risk prediction to be conducted (without compromising accuracy). This provides the advantage of estimating the seriousness of a disease one is experiencing; enabling clinicians to offer a more personalized management and/or therapeutic strategy to the patient.

The limitation of this investigation includes the use of a single dataset to evaluate the effect of sample age and prediction resolution in relation to the performance of MI risk prediction. This limits the power to conclusively state how each factor influences the performance of the prediction model. Nevertheless, it does provide some insights on whether sample age and prediction resolution have an impact on the performance of clinical risk prediction model. In view of the observations from this study and the importance of screening since young, we aim to investigate the effect of prediction resolution and sample age on younger subjects as part of our future work.

VI. CONCLUSION

Early detection of individuals with high risk of experiencing MI is very important clinically, but has proved to

be elusive. To this end, we investigated the effect of sample age and prediction resolution in relation to the development of accurate clinical risk prediction model. Our experiments indicate that both sample age and prediction resolution do not have a significant impact on prediction models developed using subjects aged 65 and above.

Overall, high validation sensitivity, specificity, and balanced accuracy were achieved by SVM algorithm. This opens the opportunity for constructing predictive models capable of detecting MI early, allowing clinicians to take preventative measures promptly, improving the quality of individuals' life, and reducing avoidable mortality.

In view of these results, we suggest the use of different prediction resolution to provide a more detailed health screening of elderly subjects so that more appropriate preventative measurements in relation to the individual's risk level can be taken.

REFERENCES

- [1] L. R. Martin, S. L. Williams, K. B. Haskard, and M. R. DiMatteo, "The challenge of patient adherence," *Therapeutics Clin. Risk Manage.*, vol. 1, no. 3, pp. 189–199, 2005.
- [2] M. A. Whooley, "To screen or not to screen?: Depression in patients with cardiovascular disease," *J. Amer. Coll. Cardiol.*, vol. 54, no. 10, pp. 891–893, 2009.
- [3] R. J. Hye, A. E. Smith, G. H. Wong, S. S. Vansomphone, R. D. Scott, and M. H. Kanter, "Leveraging the electronic medical record to implement an abdominal aortic aneurysm screening program," *J. Vascular Surg.*, vol. 59, no. 6, pp. 1535–1543, 2014.
- [4] T. Clayton, J. Lubsen, S. Pocock, Z. Vokac, B. Kirwan, K. Fox, and P. Poole-Wilson, "Risk score for predicting death, myocardial infarction, and stroke in patients with stable angina, based on a large randomised trial cohort of patients," *Biomed. J.*, vol. 331, pp. 869–873, 2005.
- [5] D. M. Lloyd-Jones, P. W. Wilson, M. G. Larson, A. Beiser, E. P. Leip, R. B. D'Agostino, and D. Levy, "Framingham risk score and prediction of lifetime risk for coronary heart disease," *Amer. J. Cardiol.*, vol. 94, no. 1, pp. 20–24, 2004.
- [6] W. Levy, D. Mozaffarian, D. Linker, S. Sutradhar, S. Anker, A. Cropp, I. Anand, A. Maggioni, P. Burton, M. Sullivan, B. Pitt, P. Poole-Wilson, D. Mann, and M. Packer, "The Seattle Heart Failure Model: Prediction of survival in heart failure," *Circulation*, vol. 113, pp. 1424–1433, 2006.
- [7] S. Alty, N. Angarita-Jaimes, S. Millasseau, and P. Chowienczyk, "Predicting arterial stiffness from the digital volume pulse waveform," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 12, pp. 2268–2275, Dec. 2007.
- [8] S. Chattopadhyay, "Mining the risk of heart attack: A comprehensive study," *Int. J. Biomed. Eng. Technol.*, vol. 11, no. 4, pp. 394–410, 2013.
- [9] X. Song, A. Mitnitski, J. Cox, and K. Rockwood, "Comparison of machine learning techniques with classical statistical models in predicting health outcomes," *Medinfo*, vol. 107, no. Pt 1, pp. 736–740, 2004.
- [10] J. Kim, B. Cho, S. Im, M. Jeon, I. Kim, and S. Kim, "Comparative study on artificial neural network with multiple regressions for continuous estimation of blood pressure," in *Proc. IEEE 27th Annu. Conf. Eng. Med. Biol.*, 2005, pp. 6942–6945.