Journal of Medical Systems (2019) 43: 253
https://doi.org/10.1007/s10916-019-1359-5

PATIENT FACING SYSTEMS

A Machine Learning-Based Approach for the Prediction of Acute Coronary Syndrome Requiring Revascularization

Yung-Kyun Noh[1] · Ji Young Park[2] · Byoung Geol Choi[3] · Kee-Eung Kim[4] · Seung-Woon Rha[3]

Received: 12 December 2017 / Accepted: 30 May 2019 / Published online: 28 June 2019

Abstract

The aim of this study is to predict acute coronary syndrome (ACS) requiring revascularization in those patients presenting early-stage angina-like symptom using machine learning algorithms. We obtained data from 2344 ACS patients, who required revascularization and from 3538 non-ACS patients. We analyzed 20 features that are relevant to ACS using standard algorithms, support vector machines and linear discriminant analysis. Based on feature pattern and filter characteristics, we analyzed and extracted a strong prediction function out of the 20 selected features. The obtained prediction functions are relevant showing the area under curve of 0.860 for the prediction of ACS that requiring revascularization. Some features are missing in many data though they are considered to be very informative; it turned out that omitting those features from the input and using more data without those features for training improves the prediction accuracy. Additionally, from the investigation using the receiver operating characteristic curves, a reliable prediction of 2.60% of non-ACS patients could be made with a specificity of 1.0. For those 2.60% non-ACS patients, we can consider the recommendation of medical treatment without risking misdiagnosis of the patients requiring revascularization. We investigated prediction algorithm to select ACS patients requiring revascularization and non-ACS patients presenting angina-like symptoms at an early stage. In the future, a large cohort study is necessary to increase the prediction accuracy and confirm the possibility of safely discriminating the non-ACS patients from the ACS patients with confidence.

Keywords Acute coronary syndrome · Machine learning · Diagnosis

Introduction

Acute coronary syndrome (ACS) is a disease in which thrombosis occurs due to rupture or erosion of a vulnerable atheroma, resulting in a sudden closure of the lumen of the coronary artery resulting in myocardial ischemia or necrosis. Therefore, rapid revascularization of obstructive coronary arteries within the recommended time is very important to improve prognosis in patients with ACS. However, in reality, revascularization of coronary arteries is delayed due to various reasons such as diagnostic dilemma, which is an important reported risk factor for the poor prognosis of ACS patients [1–3].

Yung-Kyun Noh and Ji Young Park contributed equally to this work.

This article is part of the Topical Collection on *Patient Facing Systems*

✉ Seung-Woon Rha
swrha617@yahoo.co.kr

Yung-Kyun Noh
nohyung@hanyang.ac.kr

Ji Young Park
cisamoe74@eulji.ac.kr

Byoung Geol Choi
trv940@korea.ac.kr

Kee-Eung Kim
kekim@cs.kaist.ac.kr

1 Department of Computer Science, Hanyang University, 222 Wangsimni-ro, Seongdong-gu, Seoul 04763, South Korea

2 Division of Cardiology, Nowon Eulji Hospital, Eulji University, 68 Hangeulbiseok-ro, Nowon-gu, Seoul 01830, South Korea

3 Cardiovascular Center, Korea University Guro Hospital, 148 Gurodong-ro, Guro-gu, Seoul 08308, South Korea

4 School of Computing, Korea Advanced Institute of Science and Technology, 291 Daehak-ro, Yuseong-gu, Daejeon 34141 South Korea

In this study, we use a predictive model using a machine learning algorithm to predict whether those with acute chest pain visiting an outpatient clinic or emergency room are the patients with ACS requiring immediate reperfusion of their coronary arteries.

The difference of this study from previous studies are as follows. Most of the previous studies predicted the clinical outcomes of patients after revascularization of coronary arteries, and most of them used information such as myocardial enzyme levels and electrocardiogram (ECG) at the time of admission [4]. In this study however, using a machine learning diagnostic technique, we can predict whether patients have ACS requiring immediate revascularization based only on the patients' basic minimal information such as blood pressure, heart rate, past medical history, and laboratory findings in the absence of chest pain.

Methods

Study population

We analyze 9539 patients who underwent CAG for chest pain at Korea University Guro Hospital from January 2004 to May 2014. ACS is defined as coronary artery disease (CAD) requiring revascularization for acute myocardial infarction (MI) and unstable angina, and non-ACS is defined as CAD treatable by medication alone with no evidence of significant stenotic lesion. Exclusion criteria are defined as advanced heart failure (New York Heart Association functional class III or IV) or stage 4 and 5 chronic kidney disease (e-GFR < 30 ml / min / 1.73m^2) or stable angina. A total of 5882 patients are finally enrolled, among them, 2344 patients are considered to be in the ACS requiring revascularization group and 3538 patients are considered as the non-ACS group (Fig. 1).

The standard algorithms, SVMs and LDA, are used for prediction, and the classification performance of ACS requiring revascularization is reported along with analysis of the selected features. The study protocol was approved by the institutional review board at Korea University Guro Hospital.

Experiments with data

In this study, we select features for developing a predictive model of ACS according to the globally accepted guidelines and prior literature. In the guidelines of the American Society of Cardiology and the European Society of Cardiology, age and gender; the patient's heart rate and blood pressure (BP), clinical history of cardiovascular disease, and past medical history were used for the clinical decision making of ACS [5, 6]. And these features were used in tools of early risk stratification of ACS, such as TIMI (Thrombolysis In Myocardial Infarction), GRACE (Global Registry of Acute Coronary Events), EMMACE (Evaluation of Methods of Management of Acute Coronary Syndrome) and GUSTO (Global Utilization of Streptokinase and Tissue Plasminogen Activator for Occluded Coronary Arteries) [7]. These features are listed in Table 1: two epidemiological data (gender, age), three clinical data at admission to the emergency department or outpatient clinic (systolic BP, diastolic BP, heart rate), eleven past medical history features (history of coronary artery disease, myocardial infarction, coronary artery

Fig. 1 Flow chart of study population

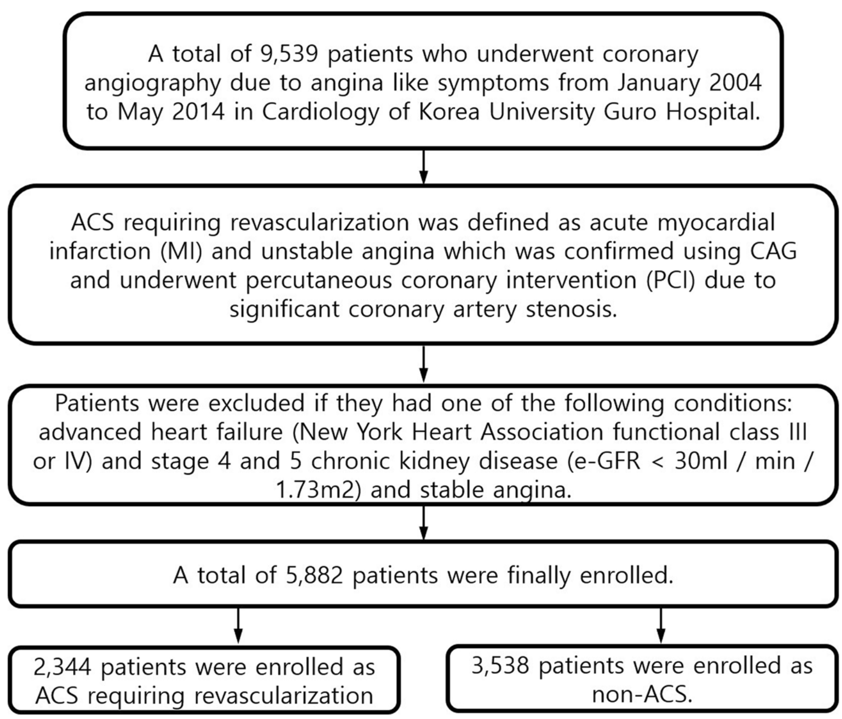

Table 1 Features for analysis

		Features
Epidemiological data	1	Gender
	2	Age
Clinical data at emergency department or outpatient clinic	3	Systolic BP
	4	Diastolic BP
	5	HR
Past medical history before presenting chest pain	6	CAD
	7	MI
	8	CABG
	9	PCI
	10	Hypertension
	11	DM
	12	Hyperlipidemia
	13	CVA
	14	PCI
	15	History of Smoking
	16	Current smoking
Laboratory data before presenting chest pain	17	TC
	18	LDL- cholesterol
	19	HDL-cholesterol
	20	TG

BP, Blood Pressure; *CABG*, Coronary Artery Bypass Graft; *CAD*, Coronary Artery Disease; *CVA*, Cerebrovascular Accidents; *DM*, Diabetes Mellitus; *HDL*, High Density Lipoprotein; *HR*, Heart Rate; *LDL*, Low Density Lipoprotein; *MI*, Myocardial Infarct; *PCI*, Percutaneous Coronary Intervention; *TC*, Total Cholesterol; *TG*, Triglyceride

bypass graft, percutaneous coronary intervention, hypertension, diabetes mellitus, dyslipidemia, cerebrovascular disease, history of smoking, and current smoking, and four laboratory data before presenting chest pain (total cholesterol, low density lipoprotein cholesterol, high density lipoprotein cholesterol, triglyceride). The classification is performed using two different feature sets: The first set includes all 20 features consisting of epidemiological data, clinical data, medical history, and laboratory data, and the second set consists of 16 elements without four of the laboratory data features. Because laboratory data have many missing elements, the number of patients with ACS requiring revascularization and the non-ACS patients are 884 and 2738, respectively for the first set without any missing elements, while the number of patients in each group to 2311 and 3527, respectively for the second set.

Statistical method

The classification is performed using SVMs and LDA [1], which are the standard classification methods in machine learning; the SVM classifier uses the boundary with the largest margin to both competing classes, and the solution is obtained using convex optimization; kernel methods are powerful methods used with the SVMs to make a nonlinear boundary [8]. Data are divided into training and testing sets for five-fold cross validation, and the classification is performed using data without any missing features. After removing the data containing missing features, smaller sets (884 for ACS and 2728 for non-ACS) are used with all 20 features, and larger sets (2311 for ACS and 3527 for non-ACS) are used with those excluding the four laboratory data features because many elements are missing in laboratory data features.

For treating missing data, imputation is performed using the information in the data, and the result is compared with those obtained from the training with fewer but unmodified data with no missing features. We use the two most well-known imputation methods: nearest neighbor imputation and the mean imputation [9, 10]. Nearest neighbor imputation uses the element of the nearest datum among the identically-labeled data in terms of the Euclidean distance using the remaining non-missing elements. Mean imputation uses the mean of the element in all the identically-labeled data. Nearest neighbor imputation uses the local information near the point, while the mean imputation uses the global information. Both imputations are relevant when the missing element has either a strong (nearest neighbor) or weak (mean) correlation with the non-missing elements.

Finally, we investigate the learned parameters of LDA to show the covariance structure of the class-conditional data how co-occurrence patterns of the features can be used for prediction in the selected 20 features. LDA finds the linear boundary considering the mean and the covariance configurations for all classes. For LDA with two classes, the separation is maximized with a generative model where each class is a Gaussian having identical covariance structures. The orthogonal direction to the boundary is considered as the one-dimensional subspace with discriminative information from LDA, and the classification is performed after the projection onto this one-dimensional subspace, which can then be investigated. For comparison, the projected data on this subspace are also shown onto the direction of the mean difference. The feature weights in LDA are shown because they represent the relative contribution of the features to the discrimination. The individual content of the discriminative information can be obtained from the mean difference of the normalized data with unit variance for each feature. The difference between the weight vector in LDA and the mean difference vector is due to the correlation between features, which can amend the mean difference vector to obtain a more robust LDA discriminating direction.

Results

Table 2 shows the five-fold cross validation results using support vector machines (SVMs) and linear discriminant analysis (LDA). Both algorithms show a similarly high prediction area under the ROC curve (AUC), and the best AUC of 0.860 was acquired using SVMs with the mean imputation. The first feature set includes the data with no missing elements for all features, and the second feature set contains features without laboratory data. Laboratory data have many missing elements, and the second set includes 2311 and 3527 data for ACS and non-ACS, respectively, compared with 884 and 2738 in the first set. Imputation further increases the number of available data, but at the same time, the imputed data inevitably distort the information in the true underlying data-generating distributions producing a bias on the predicted results. The exact effect on the bias resulting from those imputation methods should be further investigated.

Figure 2 shows the ROC curves for the experiments using SVMs and LDA for the eight experiments listed in Table 2.

Table 2 Area under receiver operating characteristic (ROC) curve for classification

	SVMs	LDA
1st feature set (with laboratory data)	0.800 ± 0.003	0.798 ± 0.004
2nd feature set (without laboratory data)	0.797 ± 0.004	0.781 ± 0.005
Nearest neighbor imputation	0.800 ± 0.005	0.784 ± 0.006
Mean imputation	0.860 ± 0.004	0.828 ± 0.004

SVMs, Support Vector Machines; *LDA*, Linear Discriminant Analysis

The SVM results with mean imputation show the best outcome in terms of AUC, but more actionable information comes from the prediction with the second set (no laboratory data). According to the detailed figure shown in the bottom-left corner of the ROC curve in Fig. 2, information from many data successfully achieves high specificity with SVMs; a classifier could be made which would predict 2.60% of the non-ACS patients correctly but would successfully predict all ACS patients. By using this classifier, 2.60% of non-ACS patients would not have to undergo an unnecessary coronary artery angiography (CAG). Otherwise, considering the fatal risk of ACS, CAG would be performed automatically regardless of their diagnosis classification.

Figure 3a and b shows the plot of data of the first set along the directions of LDA and the mean difference. The figures show the two-dimensional subspace of the 20-dimensional data space, and the horizontal axis shows the directions of LDA (Fig. 3a) and the mean difference (Fig. 3b). Both figures show that both directions separate data with some amount of overlap. The LDA direction is obtained considering the discriminative information as well as the correlation effect between features, while the mean difference direction represents only the discriminative information of each feature individually.

Figure 3c shows the weight vector elements which make the LDA discriminant direction (yellow bar) and their absolute values (blue plot) to represent the amounts. The position of the yellow bar shows whether the element tends toward ACS (positive) or non-ACS (negative). Figure 3d shows the elements of the mean difference vector (yellow bar) and their absolute values (blue plot). The difference of Fig. 3c from 3d represents the effect of correlation when we consider the multivariate information as a whole for discrimination. According to both figures, age and systolic blood pressure (BP) are the most important features for the discrimination of ACS and non-ACS. By considering the correlation among features, the age and systolic BP contribution is emphasized even more. In Fig. 3d, low density lipoprotein (LDL)-cholesterol alone cannot be used for discriminating, but as in Fig. 3c, LDL-cholesterol strongly contributes to positively to discriminating ACS when it is considered along with other features. In Fig. 3c, the effect of diastolic BP and past medical history of coronary artery disease (CAD), myocardial infarction (MI), coronary artery bypass graft (CABG), and percutaneous coronary intervention (PCI) are significantly reduced. The set of significant features contributes to the discrimination differently from the features utilized individually; when the features are used together using the correlation information between them, the important features can be obtained for the prediction of ACS requiring revascularization even with features which are considered to be noise when analyzed individually.

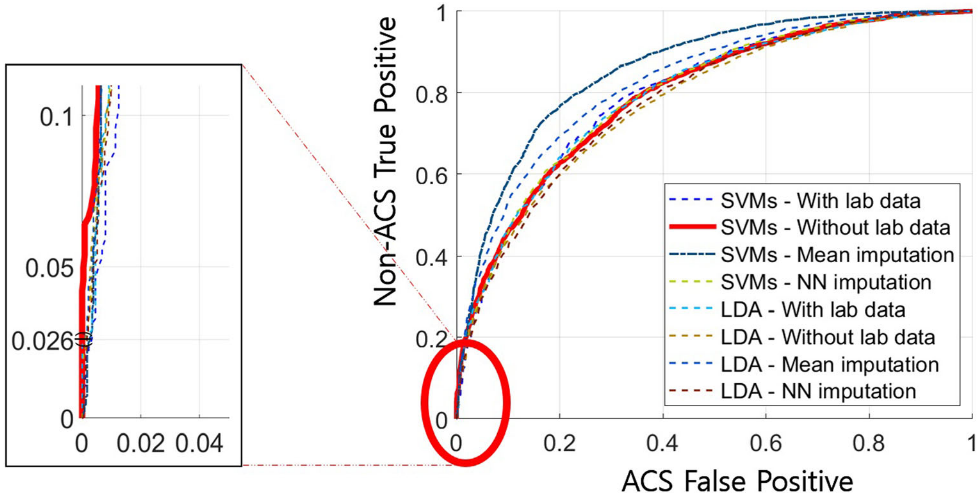

Fig. 2 Receiver operating characteristic (ROC) curves with SVM and LDA classification using each set in Table 2. Data include epidemiological data, clinical data, past medical history, and laboratory data. The first set is made using all features, but only 3618 (ACS 884 + Non-ACS 2738) number of data having all features without missing are used, and the second set is made using all features except laboratory data using 5838 (ACS 2311 + Non-ACS 3527) number of data without missing. Two sets are additionally used with imputed data in all missing features using two different imputation methods. SVMs and LDA are trained and validated using these four sets. The area under curve (AUC) for the set with mean imputation achieves the best performance of 0.860. A detailed figure of the ROC curves on the left shows that the SVM prediction without using laboratory data discriminates part of non-ACS patients (2.60%) from ACS patients without failing to capture any ACS patients

Discussion

In this study, we used machine learning algorithms to predict patients with ACS requiring immediate coronary intervention through only basic patient information before admission. The most encouraging result was that the predictive value was 86.0%, even though the myocardial enzyme and ECG at the time of admission were excluded.

In our research, we predicted ACS using machine learning with a focus on the tradeoffs between the size and the quality of data. If we compare the first and the second datasets in Table 2, the first set includes all features which are considered to be relevant, but the number of data is small when used without any missing elements. On the other hand, the second dataset excludes the laboratory data features which are the cholesterol and triglyceride (TG) information because they have many missing elements. As a result, the number of data for the second set is 3622 (884 for ACS and 2738 for non-ACS) compared with the size of the first set, which are 5838 (2311 for ACS and 3527 for non-ACS). The prediction performance of the first set is better in terms of the AUC, but the actionable information comes from the SVM prediction with the second set. About 2.6% of the non-ACS patients were correctly classified (true negative) and there was no misclassification of ACS patients (no false negative). Based on the prediction results, physicians can advise non-ACS patients not to undergo an unnecessary coronary angiography. Imputation for treating missing values was beneficial in improving the overall prediction performance. Mean imputation achieved an AUC of 0.860 with SVMs which is a significant improvement over the results without imputation. However, in the experiments with imputation, the aforementioned actionable information disappeared.

In patients with ACS, rapid revascularization of obstructive coronary arteries is very important to improve prognosis [11, 12]. However, acute chest pain is associated with a variety of cardiac or non-cardiac disease, including aortic dissection, pericarditis, musculoskeletal pain, and reflux esophagitis, and as well as ACS. Therefore, even well-trained physicians cannot easily distinguish ACS from other non-ACS diseases [13, 14]. In an emergency room or outpatient clinic, most physicians distinguish the cause of acute chest pain through clinical symptoms, ECG, and myocardial markers and then decide whether to perform immediate revascularization [15]. However, in an outpatient clinic or outside the hospital, when patients present acute chest pain, but ECG and myocardial marker testing cannot be performed, physicians have to base their diagnoses only upon clinical symptoms. In such ambiguous situations, the patient may experience a delay in being transfered to the hospital or in treatment of reperfusion, which is one of the main causes of poor prognosis [15]. Due to such risks from uncertainty and the clinical significance of CAD [16], doctors tend to perform CAG for any patient who claims acute chest pain. However, only 30% of the patients with acute chest pain who underwent CAG had actually been diagnosed with CAD, that the chest pain was indeed due to CAD, and the remaining 70% had non-cardiac chest pain [17, 18].

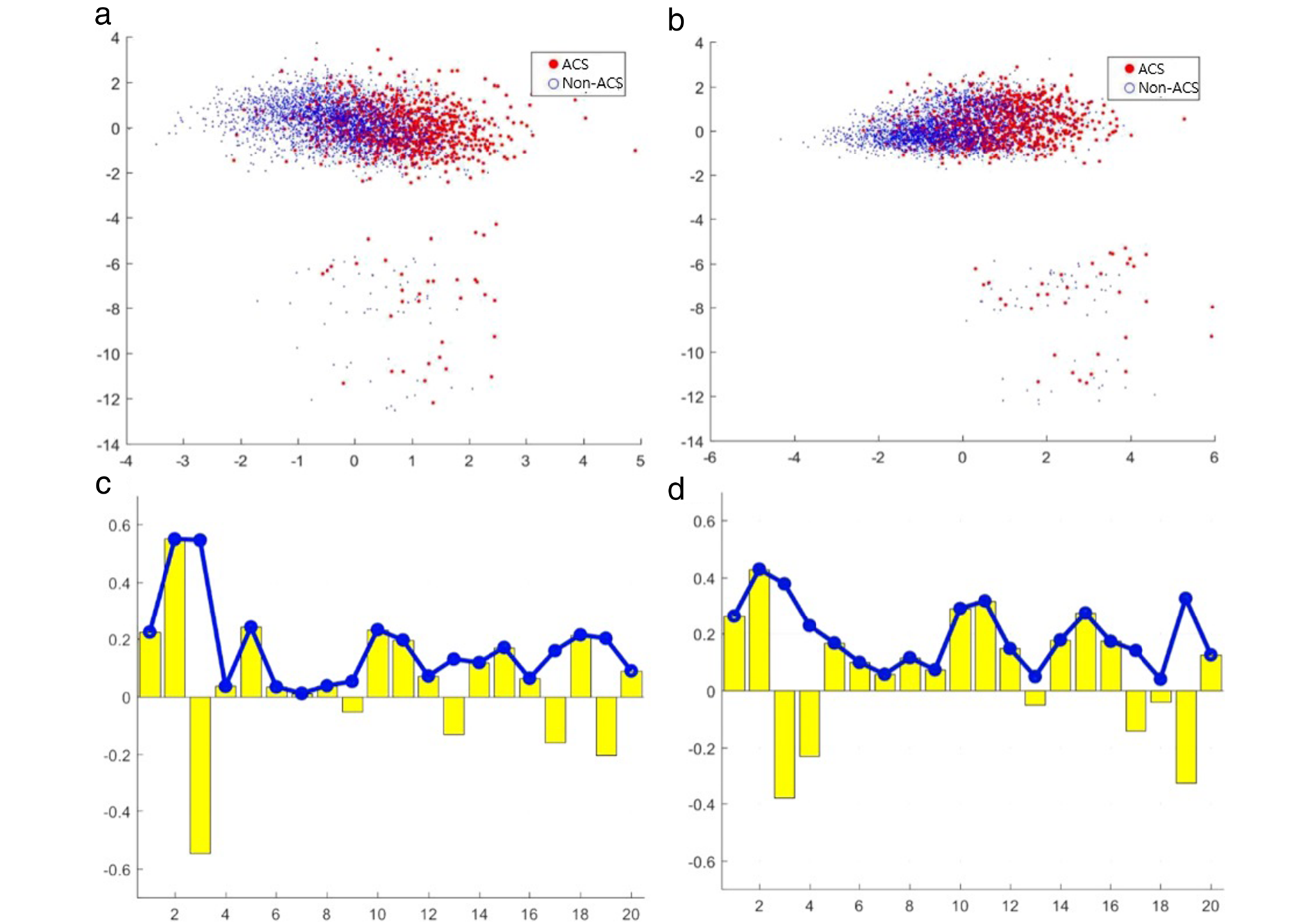

Fig. 3 Projection of data along the extracted features. Horizontal directions in (**a**) and (**b**) denote the discriminant directions obtained from LDA and the direction of mean difference between the two classes. Vertical directions in (**a**) and (**b**) are the directions of maximum variance within the orthogonal space to the horizontal directions. Figures (**c**) and (**d**) show the weight parameters of the horizontal directions of (**a**) and (**b**), respectively. The features have been indexed using those in Table 1: 1. Gender, 2. Age, 3. Systolic BP, 4. Diastolic BP, 5. HR, 6. CAD, 7. MI, 8. CABG, 9. PCI, 10. Hypertension, 11. DM, 12. Hyperlipidemia, 13. CVA, 14. PCI, 15. History of Smoking, 16. Current smoking, 17. TC, 18. LDL- cholesterol, 19. HDL-cholesterol, and 20. TG. The yellow bars are the weight parameters for the extracted discriminant features, and the blue plot displays the absolute values of the weights

In this study, we developed an ACS prediction model using data from the electronic medical record (EMR). The reason is as follows. In clinical trials of previous studies, early risk stratification of ACS was developed using data from clinical trials or registries. However in clinical trials, elderly patients over 70 or multi-morbid patients were excluded [19]. In registry studies, it is difficult to obtain sufficient negative data because the inclusion criteria depend on the final diagnosis and this dependency leads to problems in the development of an ACS prediction model [20]. Therefore, this study used EMR data that can reflect the reality and accuracy of real world situations for diagnosis of disease in real time. In this study, we select algorithms, which can perform with good generalization by resisting overfitting in building an ACS prediction model. The results show a 0.860 AUC in the prediction of ACS and non-ACS, which indicates that the selected features apparently have enough information for discriminating between the two different clinical sets of patients. However, more actionable information comes from training with many data, which can be used by eliminating the features with many missing data even though they are considered to be important. Using experiments with SVMs, 2.60% of the non-ACS patients were correctly classified while all ACS patients were classified correctly. For non-ACS patients, physicians can recommend against performing an unnecessary coronary angiography.

Finally, the learning with LDA using the first feature set in Table 2 provides the following information: according to Fig. 3d, LDL is not an important factor by itself. Traditionally, in order to prevent cardiovascular disease, it has been a priority to reduce the level of LDL first. However, residual cardiovascular risk still exists even after optimal reduction of LDL using statins. The results in Fig. 3c show that the decrease of LDL alone is an unimportant factor for predicting the risk of ACS. When other features are considered together with LDL including high density lipoprotein (HDL) and triglyceride (TG) which are known to be associated with residual cardiovascular risk, then LDL becomes a significant factor for the prediction of ACS [21].

This study has several limitations. First, there are a lot of missing laboratory data for the patients not yet admitted to an ED or outpatient clinic. In the machine learning algorithms for prediction, there is no standard method of treating missing data. We apply two widely-used imputation methods, but in the future, we can try diverse methods of treating data with missing elements including intensive investigation of the effect of imputation. In this work, the imputation did not preserve the actionable information that can be used for selecting a portion of non-ACS patients with high-confidence. In addition, a large cohort study should be followed to increase the performance and to confirm the effect of actionable information. The second limitation concerns uncertain labeling of ACS requiring revascularization; in this study, there was difficulty in checking the vulnerable plaque burden of all coronary arteries. In order to improve the accuracy of diagnosis of ACS, imaging tools, such as intravascular ultrasound or optical coherence tomography, should be considered. Third, we only use the pre-admission data from before admission to an ED or outpatient clinic. However, post-admission data, such as Electrocardiography (EKG) and cardiac enzymes, are very important features for diagnosis of ACS. In the future, we are planning to analyze the prediction of ACS using EKG and cardiac enzymes as well to improve the accuracy of the diagnosis for ACS. Fourth, there is an important concern about the population selected. In this study, we selected the patients who had undergone coronary angiography due to acute chest pain. As a result, these patients already had a high suspicion of having ACS, so it creates selection bias. In the future, patients with atypical or other non-cardiac chest pain should be analyzed together to overcome selection bias and the representativeness problem of ACS.

We investigated one prediction algorithm to select the high risk group of patients with ACS requiring revascularization and another to select the low risk group of non-ACS patients presenting angina-like symptoms at an early stage. Machine learning is an important technique for the prediction of these events, and our study of predicting ACS requiring revascularization using patient data before observed chest pain can be used to reduce the diagnostic dilemma. The main obstacles to applying the machine learning techniques are the missing data and uncertain labeling of ACS causing the small number of available data and the significant overlap between different classes, respectively. Further research would be required to overcome these limitations and to solve the delayed diagnosis problem.

Acknowledgments This research was supported by EMBRI Grants 2017SN0002 from the Eulji University and partly supported by the National Research Foundation of Korea Grant NRF/MSIT-2017R1E1A1A03070945 funded by the Korea government.

Compliance with ethical standards

Conflict of interests The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Chang, C. C., and Lin, C. J., LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST).* 2:27, 2011.
2. Noh, Y. K., Hamm, J., Park, F. C., Zhang, B. T., and Lee, D. D., Fluid dynamic models for Bhattacharyya-based discriminant analysis. *IEEE Trans Pattern Anal Mach Intell.* 40:92–105, 2018.
3. Iosifidis, A., Tefas, A., and Pitas, I., On the optimal class representation in linear discriminant analysis. *IEEE Trans Neural Netw Learn Syst.* 24:1491–1497, 2013.
4. Antman, E. M., Cohen, M., Bernink, P. J., McCabe, C. H., Horacek, T., Papuchis, G., Mautner, B., Corbalan, R., Radley, D., and Braunwald, E., The TIMI risk score for unstable angina/non-ST elevation MI: A method for prognostication and therapeutic decision making. *JAMA.* 284:835–842, 2000.
5. Anderson, J. L., Adams, C. D., Antman, E. M., Bridges, C. R., Califf, R. M., Casey, Jr., D. E., Chavey, 2nd, W. E., Fesmire, F. M., Hochman, J. S., Levin, T. N., Lincoff, A. M., Peterson, E. D., Theroux, P., Wenger, N. K., Wright, R. S., Smith, Jr., S. C., Jacobs, A. K., Halperin, J. L., Hunt, S. A., Krumholz, H. M., Kushner, F. G., Lytle, B. W., Nishimura, R., Ornato, J. P., Page, R. L., Riegel, B., American College of C, American Heart Association Task Force on Practice G, American College of Emergency P, Society for Cardiovascular A, Interventions, Society of Thoracic S, American Association of C, Pulmonary R, and Society for Academic Emergency M, ACC/AHA 2007 guidelines for the management of patients with unstable angina/non ST-elevation myocardial infarction: A report of the American College of Cardiology/American Heart Association task force on practice Guidelines (writing committee to revise the 2002 Guidelines for the Management of Patients with Unstable Angina/non ST-elevation myocardial infarction): Developed in collaboration with the American College of Emergency Physicians, the Society for Cardiovascular Angiography and Interventions, and the Society of Thoracic Surgeons: Endorsed by the American Association of Cardiovascular and Pulmonary Rehabilitation and the Society for Academic Emergency Medicine. *Circulation.* 116:e148–e304, 2007.
6. Hamm, C. W., Bassand, J. P., Agewall, S., Bax, J., Boersma, E., Bueno, H., Caso, P., Dudek, D., Gielen, S., Huber, K., Ohman, M., Petrie, M. C., Sonntag, F., Uva, M. S., Storey, R. F., Wijns, W., Zahger, D., and Guidelines, E. S. C. C. f. P., ESC Guidelines for the management of acute coronary syndromes in patients presenting

without persistent ST-segment elevation: The task force for the management of acute coronary syndromes (ACS) in patients presenting without persistent ST-segment elevation of the European Society of Cardiology (ESC). *Eur Heart J.* 32:2999–3054, 2011.

7. Berikol, G. B., Yildiz, O., and Ozcan, I. T., Diagnosis of acute coronary syndrome with a support vector machine. *J Med Syst.* 40(84), 2016.
8. Muller, K. R., Mika, S., Ratsch, G., Tsuda, K., and Scholkopf, B., An introduction to kernel-based learning algorithms. *IEEE Trans Neural Netw.* 12:181–201, 2001.
9. Rahman, M. M., and Davis, D. N., Machine learning-based missing value imputation method for clinical datasets. *IAENG Transactions on Engineering Technologies*, Springer, pp. 245–257, 2013.
10. Bektas, J., Ibrikci, T., and Turkay Ozcan, I., The impact of imputation procedures with machine learning methods on the performance of classifiers: An application to coronary artery disease data including missing values. *Biomedical Research.* 29:2780–2785, 2018.
11. King, 3rd, S. B., Marshall, J. J., and Tummala, P. E., Revascularization for coronary artery disease: Stents versus bypass surgery. *Annual review of medicine.* 61:199–213, 2010.
12. Choi, B. G., Park, S. H., Rha, S. W., Park, J. Y., Choi, S. Y., Park, Y., Xu, S., Ngow, H. A., Ali, J., Li, H., Kim, J. B., Lee, S., Na, J. O., Choi, C. U., Lim, H. E., Kim, J. W., Kim, E. J., Park, C. G., Seo, H. S., and Oh, D. J., Five-year clinical outcomes in patients with significant coronary artery spasm: A propensity score-matched analysis. *International journal of cardiology.* 184:533–539, 2015.
13. VanHouten, J. P., Starmer, J. M., Lorenzi, N. M., Maron, D. J., and Lasko, T. A., Machine learning for risk prediction of acute coronary syndrome. *AMIA Annual Symposium proceedings. AMIA Symposium* 2014:1940–1949, 2014.
14. Goodacre, S., Cross, E., Arnold, J., Angelini, K., Capewell, S., and Nicholl, J., The health care burden of acute chest pain. *Heart.* 91: 229–230, 2005.
15. Gierl, B., Aids and tips for daily practice (I). *Quintessenz Journal.* 5:37–41, 1975.
16. Rusnak, R. A., Stair, T. O., Hansen, K., and Fastow, J. S., Litigation against the emergency physician: Common features in cases of missed myocardial infarction. *Annals of emergency medicine.* 18: 1029–1034, 1989.
17. Ekelund, U., Nilsson, H. J., Frigyesi, A., and Torffvit, O., Patients with suspected acute coronary syndrome in a university hospital emergency department: An observational study. *BMC emergency medicine.* 2(1), 2002.
18. Faybush, E. M., and Fass, R., Gastroesophageal reflux disease in noncardiac chest pain. *Gastroenterol Clin North Am.* 33:41–54, 2004.
19. Anderson, J. L., Adams, C. D., Antman, E. M., Bridges, C. R., Califf, R. M., Casey, Jr., D. E., Chavey, 2nd, W. E., Fesmire, F. M., Hochman, J. S., Levin, T. N., Lincoff, A. M., Peterson, E. D., Theroux, P., Wenger, N. K., Wright, R. S., and Smith, Jr., S. C., 2011 writing group members, and ACCF/AHA task force members, 2011 ACCF/AHA focused update incorporated into the ACC/AHA 2007 Guidelines for the Management of Patients with Unstable Angina/non-ST-elevation myocardial infarction: A report of the American College of Cardiology Foundation/American Heart Association task force on practice Guidelines. *Circulation.* 123:e426–e579, 2011.
20. Yan, A. T., Jong, P., Yan, R. T., Tan, M., Fitchett, D., Chow, C. M., Roe, M. T., Pieper, K. S., Langer, A., Goodman, S. G., and Canadian Acute Coronary Syndromes registry investigators, Clinical trial–derived risk model may not generalize to real-world patients with acute coronary syndrome. *American heart journal.* 148:1020–1027, 2004.
21. Sampson, U. K., Fazio, S., and Linton, M. F., Residual cardiovascular risk despite optimal LDL cholesterol reduction with statins: The evidence, etiology, and therapeutic challenges. *Current atherosclerosis reports.* 14:1–10, 2012.