BMC Medical Informatics and Decision Making

Evidential MACE prediction of acute coronary syndrome using electronic health records

Danqing Hu[1], Wei Dong[2], Xudong Lu[1], Huilong Duan[1], Kunlun He[2,3*] and Zhengxing Huang[1*]

Abstract

Background: Major adverse cardiac event (MACE) prediction plays a key role in providing efficient and effective treatment strategies for patients with acute coronary syndrome (ACS) during their hospitalizations. Existing prediction models have limitations to cope with imprecise and ambiguous clinical information such that clinicians cannot reach to reliable MACE prediction results for individuals.

Methods: To remedy it, this study proposes a hybrid method using Rough Set Theory (RST) and Dempster-Shafer Theory (DST) of evidence. In details, four state-of-the-art models, including one traditional ACS risk scoring model, i.e., GRACE, and three machine learning based models, i.e., Support Vector Machine, L_1-Logistic Regression, and Classification and Regression Tree, are employed to generate initial MACE prediction results, and then RST is applied to determine the weights of the four single models. After that, the acquired prediction results are assumed as basic beliefs for the problem propositions and in this way, an evidential prediction result is generated based on DST in an integrative manner.

Results: Having applied the proposed method on a clinical dataset consisting of 2930 ACS patient samples, our model achieves 0.715 AUC value with competitive standard deviation, which is the best prediction results comparing with the four single base models and two baseline ensemble models.

Conclusions: Facing with the limitations in traditional ACS risk scoring models, machine learning models and the uncertainties of EHR data, we present an ensemble approach via RST and DST to alleviate this problem. The experimental results reveal that our proposed method achieves better performance for the problem of MACE prediction when compared with the single models.

Background

Acute coronary syndrome (ACS) refers to a group of conditions due to decreased blood flow in the coronary arteries such that part of the heart muscle is unable to function properly or dies [1, 2]. Major adverse cardiac events (MACE) indicates the composite of a variety of adverse events related to the cardiovascular system [3, 4], which may lead severe or fatal outcome for ACS patients. MACE prediction, as a crucial and widely explored topic, plays a pivotal role in the optimal management for ACS patients at their early stage of hospitalization, e.g., clinical decision making of care and treatment, drug development and cost estimation [4, 5].

Over the past decades, a mountain of studies has been proposed to facilitate risk assessment [1, 4]. Many traditional ACS risk score tools, e.g., TIMI [5], PURSUIT [6] and GRACE [7], have been widely used in real clinical circumstances and shown good discriminatory accuracy in predicting MACE for ACS patients [8, 9]. However, these traditional models have several inherent limitations [10]. In particular, these models developed using data from clinical trials and registries may be not representative of a general department patient population because there are strict inclusion and exclusion criteria of the cohort [1]. In addition, to obtain a simple

* Correspondence: HEKL301@aliyun.com; zhengxinghuang@zju.edu.cn
[2]Department of Cardiology, Chinese PLA General Hospital, Fuxing Road, Beijing, China
[1]College of Biomedical Engineering and Instrument Science, Zhejiang University, Key Lab for Biomedical Engineering of Ministry of Education, Zheda Road, Hangzhou, China
Full list of author information is available at the end of the article

Hu et al. BMC Medical Informatics and Decision Making 2019, **19**(Suppl 2):61

Page 10 of 197

and easy-use tool, traditional risk scoring models are established on a small set of handy-picked risk factors based on the significant univariate relationship to the end point by univariate logistic regression, which may cause deterioration of predicting performance [4, 10, 11]. Moreover, it is hard to enroll new and more discriminatory risk factors into those traditional models, which limits their extension ability [1].

Recently, with the rapid growth of electronic health records (EHRs) data, a multitude risk prediction models utilizing the potential of EHRs have become available and achieved significant improvements in this field [4, 10–13]. Most of these models are built based on machine learning and data mining techniques. Although valuable, there are still some deficiencies to apply them on mining EHRs, particularly due to the vagueness, impreciseness and uncertain clinical information contained in EHR data. Specifically, most of these models assume that MACEs have been correctly annotated in the EHR dataset and the focus is on the learning capabilities of the MACE prediction scheme. However, unambiguous MACE annotations may be difficult and imprecise due to the lack of information required for specifying certain MACE labels to patient individuals.

Both the traditional risk scoring models and machine learning based models provide us with diverse perspectives on the problem of MACE prediction [4], so that each of them results in complementary information and could be fused to produce an integrative and reliable result. By utilizing a proper strategy for the construction of an ensemble network, it can be successfully applied to MACE prediction problem with imprecise and uncertain information. Dempster-Shafer Theory [14, 15] (DST) of evidence is a general framework for reasoning with uncertainty by combining multiple evidences together to obtain a more reliable result, which has been widely employed in sensor fusion [16], financial distress detection [17], medical diagnosis [18] and etc. To this end, we propose a hybrid method using Rough Set Theory [19] (RST) and Dempster-Shafer Theory of evidence for MACE prediction. The proposed approach integrates four state-of-the-art models, including one traditional ACS risk scoring model, i.e., GRACE, and three machine learning based models, i.e., Support Vector Machine [20] (SVM), L_1-Logistic Regression [21] (L_1-LR), and Classification and Regression Tree [22] (CART), to generate comprehensive and reliable MACE prediction results. In particular, RST is applied to determine the weights of the four single models, and then the prediction results generated by these single models are assumed as basic beliefs for the problem propositions and in this way, an ensemble MACE prediction result is generated by combine each single model's evidence such that the overall prediction performance can be enhanced.

We comparatively evaluate the performance of the proposed model on a clinical dataset consisting of 2930 ACS patients and collected from the cardiology department of Chinese PLA General Hospital. The experimental results demonstrate that, in terms of reducing uncertainty caused human subjective cognition on patient data recording and annotation, our proposed method performs better than traditional single models.

Preliminaries
Rough set theory
Rough set theory was first proposed by Pawlak [19], which is widely used to deal with problem containing uncertainty. In RST, an information system is defined as a pair $\mathbb{I} = (U, A \cup R)$, where $U = \{u_1, u_2, \ldots, u_t\}$ is a non-empty set of finite objects, $A = \{a_1, a_2, \ldots, a_n\}$ is a non-empty set of finite attributes, $R = \{r_1, r_2, \ldots, r_m\}$ is a nonempty set of finite results. With each subset P A, there is an indiscernibility relation (also called equivalence relation) defined as $IND(P) = \{(x, y) \in U^2 | \forall a_i \in P, a_i(x) = a_i(y)\}$. The set of objects U can be partitioned based on the relation $IND(P)$, which is denoted by $U/IND(P)$, where an element from $U/IND(P)$ is called an equivalence class. According to equation above, the indiscernibility relation of A, R, and $A - \{a_j\}$, are defined as $IND(A) = \{(x, y) \in U^2 | \forall a_i \in A, a_i(x) = a_i(y)\}$, $IND(R) = \{(x, y) \in U^2 | \forall r_i \in R, r_i(x) = r_i(y)\}$, and $IND(A - \{a_j\}) = \{(x, y) \in U^2 | \forall a_i \in A, a_i \neq a_j, a_i(x) = a_i(y)\}$, $j = 1, 2, \ldots, m$. Depending on the theory of entropy, the dependence of R to A can be defined as:

$$D(IND(R)/IND(A)) = -\sum_{[x] \in U/IND(R)} p[x] \sum_{[y] \in U/IND(A)} p([y]/[x]) \ln(p([y]/[x]))$$

(1)

where $p[x] = \frac{card[x]}{card[U]}$, $p([y]/[x]) = \frac{card([y] \cap [x])}{card[x]}$. The significance of attribute a_j can be defined as:

$$\omega(a_j, A, R) = |D(IND(R)/IND(A - \{a_j\})) - D(IND(R)/IND(A))|, j = 1, 2, \ldots, m.$$

(2)

Finally, the weight of attribute a_j is defined as follows:

$$w(a_j) = \frac{\omega(a_j, A, R)}{\sum_{j=1}^{m} \omega(a_j, A, R)}$$

(3)

Dempster-Shafer theory
Let Θ be the frame of discernment, which represents all possible mutually exclusive states of a system. The power set 2^Θ is the set of all subset of Θ, including

Hu *et al. BMC Medical Informatics and Decision Making* 2019, **19**(Suppl 2):61

Page 11 of 197

the empty set $\varnothing$, which represents propositions related to actual state of the system. The basic probability assignment (BPA) is defined as $m : 2^{\Theta} \rightarrow [0, 1]$, where m satisfies: $m(\varnothing) = 0$, $\sum_{A \subseteq X} m(A) = 1$ and $m(A)$ is called BPA of proposition A. If $m(A) > 0$, the subset A is called focal element. The belief function of proposition A denoted as $Bel(A)$ is defined as $Bel(A) = \sum_{B \subseteq A} m(B), \forall A \subseteq \Theta$. The plausibility function of proposition A denoted as $Pl(A)$ is defined as $Pl(A) = 1 - Bel(\overline{A}) = \sum_{B \cap A \neq \varnothing} m(B), \forall A \subseteq \Theta$. The belief function and plausibility function represent the minimal and maximal support of A based on the BPA, respectively.

When the system has more than one basic probability assignment functions, Dempster's combinational rule can combine them together. Let m_1 and m_2 be the two different BPA functions, and the evidences are A_1, A_2, ... , A_m with respect to m_1 and B_1, B_2, ... , B_n with respect to m_2, if $\sum_{A_i \cap B_j = \varnothing} m_1(A_i) m_2(B_j) < 1$, we have:

$$m_{1,2}(C) = m_1 \oplus m_2(A) = \begin{cases} \dfrac{1}{1-K} \sum_{A_i \cap B_j = C} m_1(A_i) m_2(B_j), \forall C \subseteq \Theta, C \neq \varnothing \\ 0, C = \varnothing \end{cases}$$

(4)

where $K = \sum_{A_i \cap B_j = \varnothing} m_1(A_i) m_2(B_j)$, which indicates the conflict between the evidences, called conflict probability. And the coefficient $\frac{1}{1-K}$ is a normalization factor.

Methods

In this study, we propose an ensemble approach to integrate traditional risk scoring models and advanced machine learning based models together to alleviate the limitations we mentioned above. Figure 1 shows the outline of our proposed method. As depicted in Fig. 1, we firstly calculated the weights for the four single models, i.e., GRACE, SVM, CART, and L_1-LR, based on RST. After that, we employed the DST to integrate the weighted outputs of each model together as our ensemble MACE prediction result.

To give a more understandable explanation for our proposed method, we employed a subset of our real world dataset to show how we implemented our method step by step. Table 1 shows 10 patient samples from the collected dataset with their corresponding outputs from models trained in our previous work.

Weights calculation using rough set theory

Before calculating the weight of each single prediction model, we need to transform the models' outputs into dichotomous variables, such that we can apply RST to calculate the dependence of each model to the final prediction results. We choose the output that is closest to the top-left point in the area under the curve (AUC) figure as our threshold to transform the model's outputs. Experimentally on all patient samples we have, the thresholds are 0.2348, 0.22689, 0.2584 and 106.5 for SVM, L_1-LR, CART and GRACE, respectively. We tend to use the data obtained from our work to give a more practical description in this and following sections. According to the dichotomized outputs, we can calculate the weight for each single model based on Eq. (1–3). The weights are 0.5363, 0.1765, 0.1177 and 0.1696 for SVM, L_1-LR, CART and GRACE. Table 2 shows the dichotomized outputs, optimal thresholds and weights of the 4 single models.

Model fusion using Dempster-Shafer evidence theory

Before using the Dempster-Shafer Theory to combine the four models' outputs together, we need to transform the models' outputs into basic probability assignments (BPA). However, in our study, we notice that the range of GRACE's outputs is from 2 to 258, which cannot be
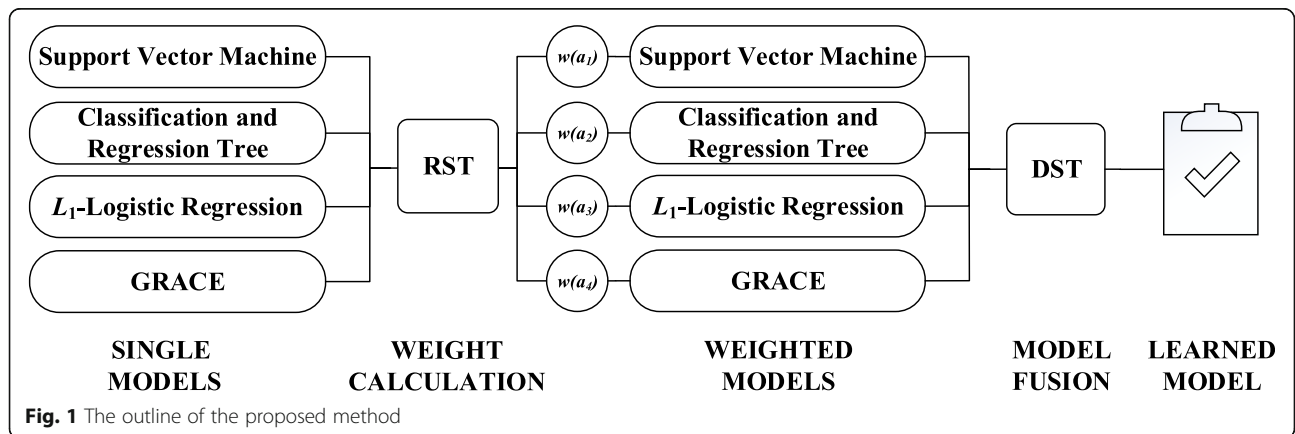

Fig. 1 The outline of the proposed method

Hu et al. BMC Medical Informatics and Decision Making 2019, **19**(Suppl 2):61

Page 12 of 197

Table 1 The original outputs of single models for 10 patient samples

Instances	Single Models				Actual MACE results
	SVM	L_1-LR	CART	GRACE	
1	0.7434	0.6032	0.6716	201	Y
2	0.1250	0.1884	0.1890	85	N
3	0.2651	0.1798	0.1890	56	N
4	0.1735	0.3272	0.3277	92	N
5	0.1608	0.3347	0.3277	119	N
6	0.7260	0.6531	0.6716	132	N
7	0.1601	0.3104	0.3346	133	Y
8	0.1171	0.1927	0.1890	137	N
9	0.4829	0.3041	0.3346	92	Y
10	0.1743	0.2050	0.3277	97	N

directly used as the BPA, and moreover, the four single models we employed have different optimal thresholds which may influence the combination results. To alleviate these problems, we first normalize the GRACE's outputs to between 0 and 1 by Eq. (5), and then apply Eq. (6) to adjust the threshold of each single model to the same value, i.e. 0.5, to eliminate the influence caused by different optimal thresholds.

$$A_{GRACE,j} = \frac{O_{GRACE,j} - min_{GRACE}}{max_{GRACE} - min_{GRACE}} ; j = 1, 2, 3, ..., n \quad (5)$$

where n is the number of patients, $O_{GRACE, j}$ and $A_{GRACE, j}$ indicate the original and normalized output of the GRACE model for the jth patient, respectively.

Table 2 The dichotomized outputs, optimal thresholds and weights of single models for 10 patient samples

Instances	Single Models				Actual MACE results
	SVM	L_1-LR	CART	GRACE	
1	1	1	1	1	1
2	0	0	0	0	0
3	1	0	0	0	0
4	0	1	1	0	0
5	0	1	1	1	0
6	1	1	1	1	0
7	0	1	1	1	1
8	0	0	0	1	0
9	1	1	1	0	1
10	0	0	1	0	0
Threshold	0.2348	0.2689	0.2584	106.5	NA
Weight	0.5363	0.1765	0.1177	0.1696	NA

max_{GRACE} and min_{GRACE}, the maximum value and minimum value of the original output of GRACE, are 37 and 201 in our study, respectively.

$$A^*_{i,j} = \begin{cases} 0.5 \times \frac{A_{i,j}}{Threshold_i}, A_{i,j} < Threshold_i \\ 0.5, A_{i,j} = Threshold_i \\ 0.5 \times \frac{A_{i,j} - Threshold_i}{1 - Threshold_i} + 0.5, A_{i,j} > Threshold_i \end{cases} ; j = 1, 2, ..., n \quad (6)$$

where $A^*_{i,j}$ is the adjusted output of ith model for the jth patient with i∈{SVM, L_1-LR, CART, GRACE}, $Threshold_i$ is the ith model's optimal threshold utilized in the dichotomization procedure for weights calculation using RST. Table 3 shows the adjusted outputs of each single model based on Eqs. (5, 6).

Based on the adjusted outputs, we can obtain the BPA for each patient. In our method, we combined the weights calculated by RST into the BPA using the following functions:

$$m_{A^*_{i,j}}(\varnothing) = 0 \quad (7)$$

$$m_{A^*_{i,j}}(1) = \frac{w_i \times A^*_{i,j}}{w_i \times A^*_{i,j} + w_i \times (1 - A^*_{i,j}) + 1} \quad (8)$$

$$m_{A^*_{i,j}}(0) = \frac{w_i \times (1 - A^*_{i,j})}{w_i \times F_{i,j} + w_i \times (1 - A^*_{i,j}) + 1} \quad (9)$$

$$m_{A^*_{i,j}}(\Theta) = \frac{1}{w_i \times A^*_{i,j} + w_i \times (1 - A^*_{i,j}) + 1} \quad (10)$$

where w_i is the weight of the ith model with i∈{SVM, L_1-LR, CART, GRACE}.

According to the weighted BPA obtained by Eqs. (7–10), we can employ the Dempster's combinational rule to combine the four models' BPA functions together. Based on Eq. (4), we have:

Table 3 The adjusted outputs of single models for 10 patient samples

Instances	Models				Actual MACE results
	SVM	L_1-LR	CART	GRACE	
1	0.8323	0.7286	0.7786	1.0000	1
2	0.2663	0.3503	0.3658	0.3453	0
3	0.5198	0.3344	0.3658	0.1367	0
4	0.3695	0.5399	0.5468	0.3957	0
5	0.3424	0.5450	0.5468	0.5661	0
6	0.8210	0.7628	0.7786	0.6349	0
7	0.3409	0.5284	0.5514	0.6402	1
8	0.2494	0.3583	0.3658	0.6614	0
9	0.6621	0.5241	0.5514	0.3957	1
10	0.3712	0.3812	0.5468	0.4317	0

Hu et al. BMC Medical Informatics and Decision Making 2019, **19**(Suppl 2):61

Page 13 of 197

$$m_{A^*_{all,j}}(1) = m_{A^*_{SVM,j}} \oplus m_{A^*_{L_1-LR,j}} \oplus m_{A^*_{CART,j}} \oplus m_{A^*_{GRACE,j}}(1) \tag{11}$$

$$m_{A^*_{all,j}}(0) = m_{A^*_{SVM,j}} \oplus m_{A^*_{L_1-LR,j}} \oplus m_{A^*_{CART,j}} \oplus m_{A^*_{GRACE,j}}(0) \tag{12}$$

Thus, the final decision value for the jth patient, i.e., $R_{all,\,j}$, can be simply represented as:

$$R_{all,j} = \frac{m_{A^*_{all,j}}(1)}{m_{A^*_{all,j}}(0) + m_{A^*_{all,j}}(1)} \tag{13}$$

Table 4 shows the patient sample's BPA, the combined BPA and the final decision value. Note that the

prediction results are determined by the optimal threshold of decision value, i.e., 0.4759, determined based on the same criteria as the dichotomization procedure. After all the procedures above, we can obtain the ensemble prediction model, which can consider the weight of each single model calculated by RST when combining the BPA by DST.

Experiments and results

Based on our previous work, we have obtained the original outputs of the four single models, e.g., SVM, L_1-LR, CART and GRACE, for a total of 2930 ACS patient samples collected from the Cardiology Department of the Chinese PLA General Hospital. We employed

Table 4 The BPA, combined BPA and the final decision value for 10 patient samples

Instances	BPA	Single Models				Combined BPA	Decision value	Prediction results	Actual MACE results
		SVM	L_1-LR	CART	GRACE				
1	1	0.2905	0.1093	0.0820	0.1450	0.4806	0.8631	1	1
	0	0.0585	0.0407	0.0233	0.0000	0.0762			
	Θ	0.6509	0.8500	0.8947	0.8550	0.4432			
2	1	0.0929	0.0525	0.0385	0.0501	0.1549	0.2845	0	0
	0	0.2561	0.0975	0.0668	0.0949	0.3894			
	Θ	0.6509	0.8500	0.8947	0.8550	0.4557			
3	1	0.1814	0.0502	0.0385	0.0198	0.2046	0.3785	0	0
	0	0.1676	0.0998	0.0668	0.1252	0.3360			
	Θ	0.6509	0.8500	0.8947	0.8550	0.4594			
4	1	0.1290	0.0810	0.0576	0.0574	0.2260	0.4190	0	0
	0	0.2201	0.0690	0.0477	0.0876	0.3133			
	Θ	0.6509	0.8500	0.8947	0.8550	0.4607			
5	1	0.1195	0.0817	0.0576	0.0821	0.2371	0.4404	0	0
	0	0.2296	0.0683	0.0477	0.0629	0.3012			
	Θ	0.6509	0.8500	0.8947	0.8550	0.4617			
6	1	0.2866	0.1144	0.0820	0.0921	0.4396	0.7999	1	0
	0	0.0625	0.0356	0.0233	0.0529	0.1100			
	Θ	0.6509	0.8500	0.8947	0.8550	0.4504			
7	1	0.1190	0.0793	0.0581	0.0928	0.2433	0.4523	0	1
	0	0.2301	0.0707	0.0472	0.0522	0.2947			
	Θ	0.6509	0.8500	0.8947	0.8550	0.4620			
8	1	0.0870	0.0537	0.0385	0.0959	0.1834	0.3397	0	0
	0	0.2620	0.0962	0.0668	0.0491	0.3564			
	Θ	0.6509	0.8500	0.8947	0.8550	0.4603			
9	1	0.2311	0.0786	0.0581	0.0574	0.3146	0.5839	1	1
	0	0.1180	0.0714	0.0472	0.0876	0.2242			
	Θ	0.6509	0.8500	0.8947	0.8550	0.4611			
10	1	0.1296	0.0572	0.0576	0.0626	0.2124	0.3931	0	0
	0	0.2195	0.0928	0.0477	0.0824	0.3278			
	Θ	0.6509	0.8500	0.8947	0.8550	0.4598			

Hu et al. BMC Medical Informatics and Decision Making 2019, **19**(Suppl 2):61

Page 14 of 197

Table 5 The weights of single models in each fold

Folds	Models			
	SVM	L_1-LR	CART	GRACE
1	0.5363	0.1765	0.1177	0.1696
2	0.2824	0.3410	0.2943	0.0823
3	0.5583	0.2189	0.1041	0.1187
4	0.4427	0.2433	0.2177	0.0962
5	0.3183	0.2988	0.2194	0.1634

5-fold cross validation to construct both the four single models and our proposed model. To compare with other ensemble methods, we trained the Bagging [23] and AdaBoost [24] models by 5-fold cross validation as well. The metrics of area under the curve [25] (AUC), prediction accuracy (ACC) and their corresponding standard deviations (STD) are employed to evaluate all these models. All model constructions and statistical analyses were completed by R version 3.3.1 (The R Foundation for Statistical Computing, Vienna, Austria). Table 5 illustrates four single models' weights in 5-fold cross validation. Tables 6 and 7 shows the AUC value and accuracy for all models in our study.

From Table 5, we can find that each model has different weights in each fold, which indicates that the weight calculation step in our method distinguishes the discrimination ability of each single model and affects the construction of the proposed model in each fold cross validation. As illustrated in Tables 6 and 7, we can notice that our proposed method achieves the highest AUC value comparing with the 4 single models which means it can combine the output of each single model and generate a more reliable prediction result. And also, the accuracy of our model is competitive in all models with AUC values above 0.70. Moreover, when compared with the traditional ensemble methods, i.e., Bagging and AdaBoost, our models achieve a better performance with a significant margin. Furthermore, we can notice that the proposed model is the only one whose all AUC values in 5-fold are above 0.70 with a competitive standard deviation, which indicates the outstanding stability of our method.

Figures 2 and 3 presents a more understandable comparison between our proposed model and other models.

Discussion

The problem of MACE prediction plays a vital role in the optimal treatment management for ACS patients during their hospitalizations. Facing with the limitations in traditional risk scoring models, machine learning methods and the uncertainties of EHR data, we present an ensemble approach to alleviate this problem. We firstly employed RST to determine each single MACE prediction model's weight. And then, DST was applied to combine all weighted single models as our ensemble model so as to enhance the performance of MACE prediction. Experiments have been conducted on a clinical dataset collected from the Cardiology Department of the China PLA General Hospital. The experimental results show our proposed method achieves the best prediction performance with 0.715 AUC value, which indicates our model can combine various information provided by the single models to generate more reliable and stable prediction result on the MACE prediction problem.

It should be mentioned that there exist some problems needed further exploration.

In our current work, the single models we employed are based on our previous work directly with no further selection. However, the single model's outputs will have a significant impact on the final prediction results. Thus, we need to explore which single models are the most appropriate for the proposed method to combine so as to improve the prediction performances. Furthermore, resampling, a key technique to construct more single models, is also a potential direction to build more powerful and robust ensemble prediction model based on the proposed method.

In our future research, we plan to develop and deploy a continuous MACE prediction service in practice. Note that the dynamic nature of a patient status is often essential to risk stratification and subsequent treatment interventions adopted in clinical practice. Thus, it would be valuable to provide a continuous MACE prediction

Table 6 The AUC values of all models

Folds	Single Models				Ensemble Models		Proposed
	SVM	L_1-LR	CART	GRACE	Bagging	AdaBoost	
1	0.742	0.724	0.644	0.641	0.714	0.678	0.736
2	0.696	0.715	0.664	0.629	0.688	0.701	0.713
3	0.704	0.689	0.594	0.635	0.707	0.696	0.707
4	0.682	0.702	0.604	0.640	0.706	0.686	0.700
5	0.711	0.704	0.645	0.636	0.683	0.672	0.717
Average	0.707	0.707	0.630	0.636	0.700	0.687	0.715
STD	±0.022	±0.013	±0.030	±0.005	±0.013	±0.012	±0.013

Hu *et al. BMC Medical Informatics and Decision Making* 2019, **19**(Suppl 2):61

Page 15 of 197

Table 7 The accuracy values of all models

Folds	Single Models				Ensemble Models		Proposed
	SVM	L_1-LR	CART	GRACE	Bagging	AdaBoost	
1	0.715	0.725	0.693	0.625	0.674	0.630	0.724
2	0.662	0.703	0.679	0.592	0.667	0.655	0.717
3	0.635	0.684	0.715	0.560	0.659	0.671	0.674
4	0.695	0.659	0.734	0.679	0.677	0.654	0.671
5	0.676	0.677	0.696	0.601	0.676	0.689	0.686
Average	0.676	0.690	0.703	0.611	0.671	0.660	0.694
SD	±0.031	±0.025	±0.021	±0.045	±0.008	±0.022	±0.024
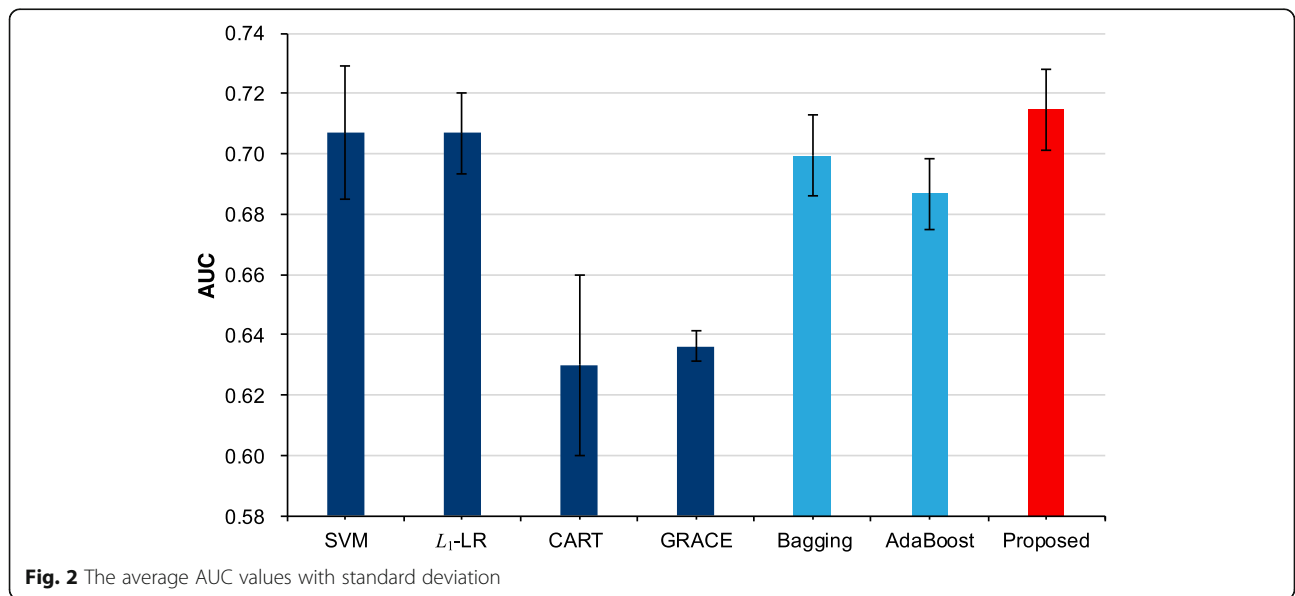
Fig. 2 The average AUC values with standard deviation
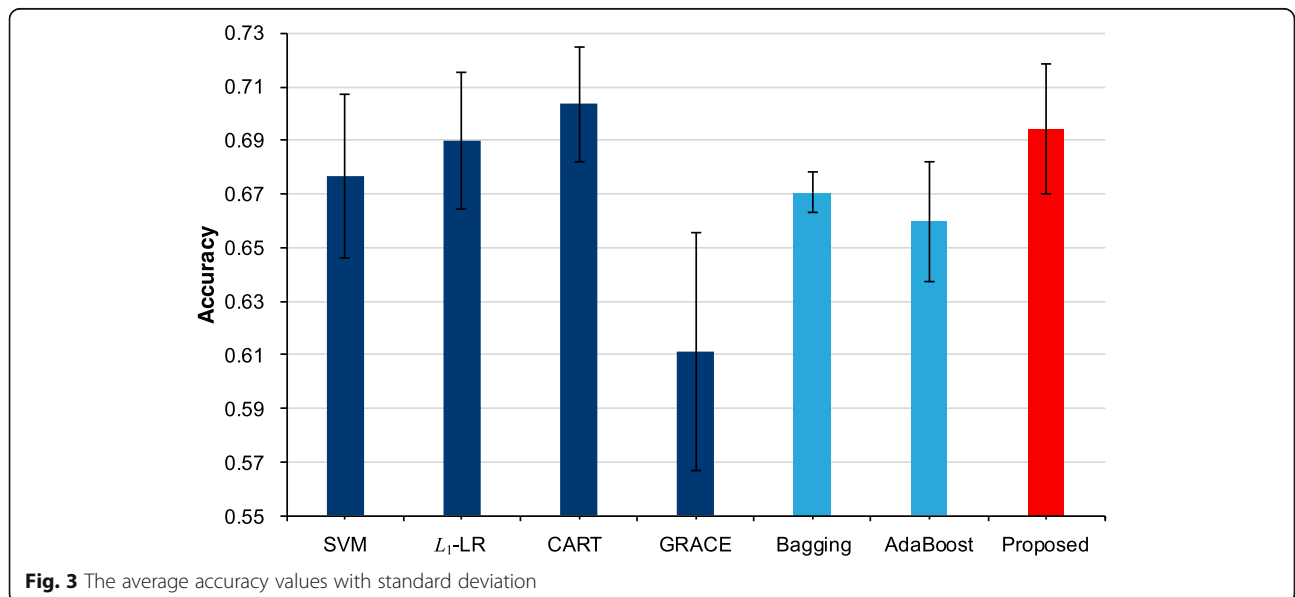
Fig. 3 The average accuracy values with standard deviation

Hu *et al. BMC Medical Informatics and Decision Making* 2019, **19**(Suppl 2):61

Page 16 of 197

service during patients' length of stay. Such a service not only anticipate MACEs at runtime, but also monitors patient treatment processes in a continuous and predictive fashion.

Conclusion

In this paper, we present an ensemble approach to alleviate the limitations in traditional ACS risk scoring models, machine learning models and the uncertainties of EHR data. We first employed RST to determine the weight for each single model. After that, DST was applied to combine the weighted outputs of single models as the final prediction results. The experimental results indicate our proposed method achieves 0.715 AUC value with a competitive standard deviation, which is a better performance for the problem of MACE prediction when compared with the single models.

Abbreviations
ACC: Accuracy; ACS: Acute Coronary Syndrome; AUC: Area Under the Curve; BPA: Basic Probability Assignment; CART: Classification and Regression Tree; DST: Dempster-Shafer Theory; EHR: Electronic Health Record; GRACE: Global Registry of Acute Coronary Events; L_1-LR: L_1-Logistic Regression; MACE: Major Adverse Cardiac Event; PURSUIT: Platelet glycoprotein IIb/IIIa in Unstable angina: Receptor Suppression Using Integrilin (eptifibatide) Therapy; RST: Rough Set Theory; STD: Standard Deviation; SVM: Support Vector Machine; TIMI: Thrombolysis in Myocardial Infarction

Availability of data and materials
The datasets generated and/or analyzed during the current study are not publicly available due to the hospital's regulations, but are available from the corresponding author on reasonable request.

About this supplement
This article has been published as part of *BMC Medical Informatics and Decision Making Volume 19 Supplement 2, 2019: Proceedings from the 4th China Health Information Processing Conference (CHIP 2018)*. The full contents of the supplement are available online at URL. https://bmcmedinformdecismak.biomedcentral.com/articles/supplements/volume-19-supplement-2.

Authors' contributions
DH, KH and ZH conceived of the proposed idea and planned the experiments. DH implemented the methods, carried out the experiments and evaluated the proposed models. DH and ZH wrote the manuscript with the comments from WD, KH, XL and HD. All authors have read and approved the final manuscript.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]College of Biomedical Engineering and Instrument Science, Zhejiang University, Key Lab for Biomedical Engineering of Ministry of Education, Zheda Road, Hangzhou, China. [2]Department of Cardiology, Chinese PLA General Hospital, Fuxing Road, Beijing, China. [3]Beijing Key Laboratory Of Chronic Heart Failure Precision Medicine, Beijing, China.

Published: 9 April 2019

References
1. Amsterdam EA, Wenger NK, Brindis RG, Casey DE, Ganiats TG, Holmes DR, Jaffe AS, Jneid H, Kelly RF, Kontos MC, et al. 2014 AHA/ACC guideline for the Management of Patients with non-ST-elevation acute coronary syndromes a report of the American College of Cardiology/American Heart Association task force on practice guidelines. Circulation. 2014; 130(25):E344–426.
2. Acute Coronary Syndrome. https://en.wikipedia.org/wiki/Acute_coronary_syndrome. Accessed 12 Oct 2017.
3. Ye S. Coronary event. In: Gellman MD, Turner JR, editors. Encyclopedia of behavioral medicine. New York: Springer New York; 2013. p. 503.
4. Hu DQ, Huang ZX, Chan TM, Dong W, Lu XD, Duan HL. Utilizing Chinese admission records for MACE prediction of acute coronary syndrome. Int J Env Res Pub He. 2016;13(9):912.
5. Antman EM, Cohen M, Bernink PM, et al. The timi risk score for unstable angina/non–st elevation mi: a method for prognostication and therapeutic decision making. JAMA. 2000;284(7):835–42.
6. Boersma E, Pieper KS, Steyerberg EW, Wilcox RG, Chang WC, Lee KL, Akkerhuis KM, Harrington RA, Deckers JW, Armstrong PW, et al. Predictors of outcome in patients with acute coronary syndromes without persistent ST-segment elevation results from an international trial of 9461 patients. Circulation. 2000;101(22):2557–67.
7. Granger CB, Goldberg RJ, Dabbous O, et al. Predictors of hospital mortality in the global registry of acute coronary events. Arch Intern Med. 2003;163(19):2345–53.
8. Goncalves PDA, Ferreira J, Aguiar C, Seabra-Gomes R. TIMI, PURSUIT, and GRACE risk scores: sustained prognostic value and interaction with revascularization in NSTE-ACS. Eur Heart J. 2005;26(9):865–72.
9. D'Ascenzo F, Biondi-Zoccai G, Moretti C, Bollati M, Omede P, Sciuto F, Presutti DG, Modena MG, Gasparini M, Reed MJ, et al. TIMI, GRACE and alternative risk scores in acute coronary syndromes: a meta-analysis of 40 derivation studies on 216,552 patients and of 42 validation studies on 31,625 patients. Contemp Clin Trials. 2012;33(3):507–14.
10. Huang ZX, Dong W, Duan HL. A probabilistic topic model for clinical risk stratification from electronic health records. J Biomed Inform. 2015;58:28–36.
11. Motwani M, Dey D, Berman DS, Germano G, Achenbach S, Al-Mallah MH, Andreini D, Budoff MJ, Cademartiri F, Callister TQ, et al. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis. Eur Heart J. 2017;38(7):500–7.
12. Huang ZX, Chan TM, Dong W. MACE prediction of acute coronary syndrome via boosted resampling classification using electronic medical records. J Biomed Inform. 2017;66:161–70.
13. Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? PLoS One. 2017;12(4):e0174944.
14. Dempster AP. Upper and lower probabilities induced by a multivalued mapping. In: Yager RR, Liu L, editors. Classic works of the Dempster-Shafer theory of belief functions. Berlin, Heidelberg: Springer Berlin Heidelberg; 2008. p. 57–72.
15. Shafer G. A mathematical theory of evidence. Princeton: Princeton University Press; 1976.
16. Basir O, Yuan XH. Engine fault diagnosis based on multi-sensor information fusion using Dempster-Shafer evidence theory. Inform Fusion. 2007;8(4):379–86.

Hu *et al. BMC Medical Informatics and Decision Making* 2019, **19**(Suppl 2):61

Page 17 of 197

17. Xiao Z, Yang XL, Pang Y, Dang X. The prediction for listed companies' financial distress by using multiple prediction methods with rough set and Dempster-Shafer evidence theory. Knowl-Based Syst. 2012;26:196–206.
18. Wang JW, Hu Y, Xiao FY, Deng XY, Deng Y. A novel method to use fuzzy soft sets in decision making based on ambiguity measure and Dempster-Shafer theory of evidence: an application in medical diagnosis. Artif Intell Med. 2016;69:1–11.
19. Pawlak Z. Rough sets. Int J Comput Inform Sci. 1982;11(5):341–56.
20. James G, Witten D, Hastie T, Tibshirani R. Support vector machines. In: An introduction to statistical learning: with applications in R. New York: Springer New York; 2013. p. 337–72.
21. James G, Witten D, Hastie T, Tibshirani R. Linear model selection and regularization. In: An introduction to statistical learning: with applications in R. New York: Springer New York; 2013. p. 203–64.
22. Loh W-Y: Classification and regression trees. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 2011, 1(1):14–23.
23. Breiman L. Bagging predictors. Mach Learn. 1996;24(2):123–40.
24. Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. J Comput Syst Sci. 1997;55(1):119–39.
25. Bradley AP. The use of the area under the roc curve in the evaluation of machine learning algorithms. Pattern Recogn. 1997;30(7):1145–59.