PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Machine learning for prediction of sudden cardiac death in heart failure patients with low left ventricular ejection fraction: study protocol for a retro-prospective multicenter registry in China
AUTHORS	Meng, Fanqi; Zhang, Zhihua; Hou, Xiaofeng; Qian, Zhiyong; Wang, Yao; Chen, Yanhong; Wang, Yilian; Zhou, Ye; Chen, Zhen; Zhang, Xiwen; Yang, Jing; Zhang, Jinlong; Guo, Jianghong; Li, Kebei; Chen, Lu; Zhuang, Ruijuan; Jiang, Hai; Zhou, Weihua; Tang, Shaowen; Wei, Yongyue; Zou, Jiangang

VERSION 1 - REVIEW

REVIEWER	Avinainder Singh Brigham and Women's Hospital Boston, MA USA
REVIEW RETURNED	12-Jun-2018

GENERAL COMMENTS	Meng and colleagues are aiming to evaluate the role of a machine learning model in prediction of SCD among patients with reduced EF among a large cohort in China. I have the following issues: Major comments: 1.The manuscript text lacks proper grammar, and sentence structure may benefit from additional English language editing (especially abstract) 2. Please explain the following discrepancies between the inclusion criteria reported in the trial registration http://www.chictr.org.cn/showprojen.aspx?proj=18229 vs. as reported in the manuscript. a. Inclusion EF <50% vs. EF<=35% b. Study design – case series vs. “retro-prospective” c. The proposed sample size in the trial protocol was 3000 patients, whereas the authors are planning to enroll only 2000 patients prospectively for this study? 3.Implantation of ICD is an exclusion criteria, can the authors comment on the selection bias, as they will be excluding the

	highest risk patients, who have received appropriate guideline directed therapy for their condition? 4.How will the authors account for physician preferences / variability / thresholds for ICD implantation for prevention of SCD? 5.Please specify which machine learning modelling approaches you will be using – e.g random forest etc. 6.Please add a paragraph on dissemination / scaling of a potential model for wider use 7.Page 17 – “Except for SCD”, are the authors not going to develop a model for SCD, the primary outcome of their study? 8. Consider collecting information on socioeconomic and educational status at baseline, as well as compliance with medications during follow-up Minor comments: Line 42-abstract, should be Cox proportional hazards instead of “COX”; similarly for instances throughout the paper. Page 5, Line 12-intro, remove obvious Page 5, Line 20-intro, SCD is responsible for over 50% of deaths, can the authors cite other data to support this statement? Page 6, Please move tables to end of the manuscript Page 8, Move list of hospitals to Supplement Page 9, Create table or figure for inclusion / exclusion criteria Page 17 – Please remove the sentence – “this project has great promise” - conjecture

REVIEWER	Dimitrios I. Fotiadis University of Ioannina Greece
REVIEW RETURNED	07-Aug-2018

GENERAL COMMENTS	The authors describe a study protocol for a retro-prospective multi-center trial for the prediction of sudden cardiac death in heart failure patients. The following minor and major issues should be addressed by the authors: 1. The authors should use "," at the end of the phrase "...(HF) patients..." in line 11 page 2 2. The meaning of the phrase "As compared to ...will be evaluated" is not clear. Probably the authors mean that the ML model will be evaluated and compared with traditional multivariable COX regression model derived from the same database. 3. The statement that "All results....at relevant conferences" is not necessary. 4. The meaning of the phrase "Sudden cardiac death...." of all HF deaths" is not clear.

	5. The meaning of the phrase "New strategies of identifying ...is urgently needed" is not clear. 6. I suggest the authors to Include the appropriate references to the Table 1. References can be places at the end of the authors name. 7. A recent review of the literature should be performed in order more recent studies to be included (studies reported after 2013). 8. I suggest authors to avoid using big sentences because it is difficult for the reader to understand them. For example see the sentence starting at line 18 and ending at line 27. 9. I suggest the authors to remove from the aims of the study the evaluation of machine learning models. Once a model is proposed it is self-evident that it should be evaluated and validated. 10. I suggest to the authors, taking into account the content of subsections, to change or remove the header "Statistical analysis". 11. The header "Data classification and pre-processing" does not correspond to the content of the paragraph that follows. I suggest the authors to use a header the phrase "Pre-processing of data" and provide more information how this will be performed. 12. The authors mention that information gain ranking method will be followed in order feature selection to be achieved. My first recommendation is to add an appropriate reference, my second recommendation is to explain why they select the specific method and my third recommendation is to test other feature ranking, as well as wrapper approach for feature selection. 13. The authors should make more clear what will be the response feature of the classifiers e.g. SCD yes vs. SCD no or SCD high risk vs. SCD low risk. 14. The authors should explain why they select the specific classifiers. For example ensemble classifiers like Random Forests may also be tested. 15. The authors mention that the algorithms will be evaluated based on the accuracy and interpretability. According to my understanding the output of a decision tree can be interpreted. However, the interpretability of the output of an SVM or a neural network is not a straightforward process. 16. Please add an appropriate reference for the COX proportional-hazards regression process.

REVIEWER	Maryam Panahiazar The University of California San Francisco, USA I have been working in this field for years.
REVIEW RETURNED	25-Aug-2018

GENERAL COMMENTS	Dear Authors, Your study protocol sounds very interesting and appropriate for the research plan and clinical trials. Your study protocol is very organized and well written and considers so many details which are necessary for such an important study. But I do suggest you for a major basic revision before I can give more detail suggestions/comments. Everything in this protocol is backed to at least 2 years ago. It seems you wanted to submit this study 2 years ago. The statistics, the facts/hypothesis, state of the art machines learning methods are back to 2 years and even more. They are more studies have been published recently using ML for outcome prediction, etc. I would like you to consider them in such

	an important protocol. Even the plan for data collections is back to 2016. I suggest to update this protocol as a present situation and state of the art machine learning methods and other studies which are happened in the last 2-3 years and I will be more than happy to review this protocol again after your revision.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Avinainder Singh
Institution and Country: Brigham and Women's Hospital, Boston, MA, USA
Major comments:

1.The manuscript text lacks proper grammar, and sentence structure may benefit from additional English language editing (especially abstract)

Response: We have revised the whole manuscript with the help of professional agency.

2. Please explain the following discrepancies between the inclusion criteria reported in the trial registration http://www.chictr.org.cn/showprojen.aspx?proj=18229 vs. as reported in the manuscript.
a. Inclusion EF <50% vs. EF<=35%
b. Study design – case series vs. “retro-prospective”
c. The proposed sample size in the trial protocol was 3000 patients, whereas the authors are planning to enroll only 2000 patients prospectively for this study?

Response:

a. In the initial stages of study design, we intended to predict the SCD in heart failure patients with LVEF<50%, namely heart failure with reduced ejection fraction (HFrEF) . We registered the study with this version protocol in the Chinese Clinical Trial Registry. However, the inclusion criteria in this manuscript was changed to EF≤35%, according to the suggestion of experts during the project application process. The reasons are as follows. Firstly, the SCD risk in patients with EF≤ 35% is higher than that with EF>35%. So the higher event rate will be observed. In the limited study timeframe, the study efficiency will be improved. Secondly, according to the present ICD indication, EF≤35%, a large number of HF patients are the candidates for ICD, however, the use of ICD in China is much lower than would be expected. It is very important to stratify the risk of SCD in patients with HF who meet the current indications for ICD, which should improve the cost-effectiveness of this treatment.

b. In order to improve the study efficiency, in the new version of protocol, according to the suggestion of experts, about 500 cases (2016-2017) will be collected retrospectively. Because the retrospective cases will also be prospectively followed up, the bias will be reduced. The 500 retrospective cases and the first 1000 prospective cases(2018-2019) will be used to develop the prediction models. And the next 1000 prospective cases will be used for model validation.

For your convenience, the related statements are pasted here:

The cases from January 2016 to December 2017 in the First Affiliated Hospital of Nanjing Medical University and Xiamen Cardiovascular Hospital Xiamen University will be collected retrospectively and followed-up prospectively. About 500 retrospective cases meet the inclusion criteria according to preliminary estimation. The prospective recruitment has started in the above 14 hospitals since January 2018. The retrospective cases and the first 1000 prospective cases will be used to develop the prediction models. And the next 1000 prospective cases will be used for model validation. The flow diagram of the progress is illustrated in Figure 1. (page 7, line 5-12)

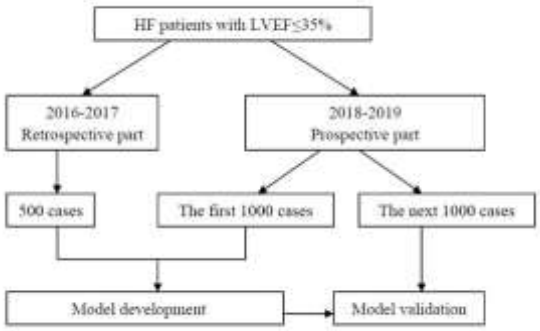

Figure 1 Flow diagram of progress

c. Due to the higher event rate in EF≤ 35%, sample size required can be reduced. So, the proposed sample size was adjusted from 3000 to 2500, including 500 retrospective cases and 2000 prospective cases.

3.Implantation of ICD is an exclusion criteria, can the authors comment on the selection bias, as they will be excluding the highest risk patients, who have received appropriate guideline directed therapy for their condition?

Response: The inclusion criteria (EF≤ 35%) of this study is the same as ICD implantation indication. The SCD risk of enrolled population will not be different from that of excluded ICD population. And the rate of implantation of ICD is really low in China. Based on current ICD indications, It has a total number of 1903 (or 3235 if CRT-D included) in 2013, with only 45% for primary prevention (55% for secondary prevention), which will make the selection bias minimized.

Reference:

[13] Shen L, Jhund P S, Petrie M C, et al. Declining Risk of Sudden Death in Heart Failure[J]. N Engl J Med, 2017,377(1):41-51.

4.How will the authors account for physician preferences / variability / thresholds for ICD implantation for prevention of SCD?

Response: In China, physicians will recommend ICD therapy for patients who meet the current indications. However, due to the expensive device and the high cost during follow-up, only a very limited number of patients accept it, especially for primary prevention. There will be no conflicts between this study and the actual medical activity.

5.Please specify which machine learning modelling approaches you will be using – e.g random forest etc.

Response: Two machine learning techniques will be used:

1) Variable selection, such as information gain and random forest.

2) Machine learning derived prediction models, including decision trees, logistic regression, support vector machine, random forest, and artificial neural network.

For your convenience, the related statements are pasted here(page 12-13, Machine learning section):

Variable selection is the process of selecting a subset of relevant variables for use in model construction, which can substantially reduce the abundant information and decrease the number of variables that are input to the prediction model. In this study, the technique named as “information gain ranking” will be used to select appropriate variables. Information gain represents the effectiveness of a variable based on entropy, which characterizes the unpredictability of a system. The information gain of a variable is evaluated as the entropy difference of the system when including and excluding this variable. Then the variables whose information gain scores are less than a threshold are considered to be insignificant and will be excluded in the prediction.

Prediction models for SCD in HF patients will be developed by the following classification algorithms respectively: decision trees, logistic regression, support vector machine, random forest, and artificial neural network. [29] The performance and general error estimation of these ML models will be assessed by 10-fold cross-validation. The dataset will be randomly divided into 10 equal folds. 9 folds will be used as training set with the remaining one fold as validation set. The validation results from 10 repeats will be combined to provide a measure of the overall performance. The prediction models derived from the four classification algorithms above will be evaluated based on the accuracy and interpretability. Finally, clinical experts and computer specialists will discuss and choose the best model to predict the prognosis of SCD in HF patients and then perform the further validation with the prospective dataset.

6.Please add a paragraph on dissemination / scaling of a potential model for wider use

Response: “Dissemination” was put together with “Ethics” in penultimate section of this manuscript. The last sentence of this paragraph is: All results of this study will be published in international peer-reviewed journals and presented at relevant conferences.(page 14, Ethic and dissemination section)

7.Page 17 – “Except for SCD”, are the authors not going to develop a model for SCD, the primary outcome of their study?

Response: We are sorry for the mistake and have revised as follows: A broad range of outcomes, including SCD, all-cause death, lethal arrhythmia, sudden cardiac arrest, cardiopulmonary resuscitation and rehospitalization due to HF, will be evaluated in this study, and the corresponding prediction models will be developed. (page 15, Discussion section, line 14-16)

8.Consider collecting information on socioeconomic and educational status at baseline, as well as compliance with medications during follow-up

Response: Thanks for your suggestion. We have added the information collection on socioeconomic and educational status at baseline, and the compliance with medication during regular follow-up. For your convenience, the related statements are pasted here: Socioeconomic and educational status: marital status, educational status, monthly income, sources of medical expenses, medical insurance. (page11, line 1-2). Table 2 has been revised accordingly. (page 16, table 2) The compliance with medications will be evaluated. (page 11, line 6-7)

Minor comments:

Line 42-abstract, should be Cox proportional hazards instead of “COX”; similarly for instances throughout the paper.

Response: It’s our mistake. We have revised throughout the manuscript.

Page 5, Line 12-intro, remove obvious
Response: We have removed “obvious” from this sentence. For your convenience, the revised sentence is pasted here: Although the survival rate after HF diagnosis has been increased due to (obvious) improvement in medical therapy, the mortality of HF remains high. (page 4, line 5-7)

Page 5, Line 20-intro, SCD is responsible for over 50% of deaths, can the authors cite other data to support this statement?

Response: We have revised this sentence and cited a new paper to support this statement. For your convenience, the revised sentence is pasted here: The two most common causes of death in patients with HF are sudden cardiac death (SCD) and progressive pump failure. SCD in HF patients is usually caused by lethal arrhythmias such as ventricular tachycardia or ventricular fibrillation, and is reported to be responsible for nearly 50% of all cardiovascular death in HF patients.[3, 4] (page 4, line 8-12)

Added reference:

[4] Solomon S D, Wang D, Finn P, et al. Effect of candesartan on cause-specific mortality in heart failure patients: the Candesartan in Heart failure Assessment of Reduction in Mortality and morbidity (CHARM) program[J]. Circulation, 2004,110(15):2180-2183.

Page 6, Please move tables to end of the manuscript

Response: We have moved tables to the end of the manuscript. (page 15-16)

Page 8, Move list of hospitals to Supplement

Response: We have moved the list of hospitals to Supplement, which has been uploaded.

Page 9, Create table or figure for inclusion / exclusion criteria
Response: Thank you for your suggestion. But I think the inclusion/exclusion criteria is very concise. If we create a table or figure for it, the content will be duplicate.

Page 17 – Please remove the sentence – “this project has great promise” - conjecture

Response: We have removed this sentence. (page 15, Discussion section, line 18) Thank you for your suggestion.

Reviewer: 2
Reviewer Name: Dimitrios I. Fotiadis
Institution and Country: University of Ioannina, Greece

The authors describe a study protocol for a retro-prospective multi-center trial for the prediction of sudden cardiac death in heart failure patients. The following minor and major issues should be addressed by the authors:

1. The authors should use "," at the end of the phrase "...(HF) patients..." in line 11 page 2

Response: We have added “,” at the end of the phrase"...(HF) patients..." according to your suggestion. For your convenience, the revised sentence is pasted here: Left ventricular ejection fraction (LVEF) ≤35%, as current significant implantable cardioverter-defibrillator (ICD) indication for primary prevention of sudden cardiac death (SCD) in heart failure (HF) patients, has been widely recognized to be inefficient. (page 2, Abstract section, line 3-6)

2. The meaning of the phrase "As compared to ...will be evaluated" is not clear. Probably the authors mean that the ML model will be evaluated and compared with.

Response: We are very sorry for our vague statement. We have revised this sentence:

Both ML and traditional multivariable COX proportional hazards regression models will be developed and compared in the prediction of SCD. Moreover, the ML model will be validated in a prospective study. (page 2, Abstract section, line 21-23)

3. The statement that "All results....at relevant conferences" is not necessary.

Response: All results of this study will be published in international peer-reviewed journals and presented at relevant conferences. This is the part of the statement of Ethics and dissemination section. (page 2) We think this statement is necessary.

4. The meaning of the phrase "Sudden cardiac death...." of all HF deaths" is not clear.

Response: We have revised this sentence and cited a new paper to support this statement. For your convenience, the revised sentence is pasted here: The two most common causes of death in patients with HF are sudden cardiac death (SCD) and progressive pump failure. SCD in HF patients is usually caused by lethal arrhythmias such as ventricular tachycardia or ventricular fibrillation, and is reported to be responsible for nearly 50% of all cardiovascular death in HF patients.[3, 4] (page 4, line 8-12)

Added reference:

[4] Solomon S D, Wang D, Finn P, et al. Effect of candesartan on cause-specific mortality in heart failure patients: the Candesartan in Heart failure Assessment of Reduction in Mortality and morbidity (CHARM) program[J]. Circulation, 2004,110(15):2180-2183.

5. The meaning of the phrase "New strategies of identifying ...is urgently needed" is not clear.

Response: We have revised this part. For your convenience, the revised statements are pasted here: Identifying the patients who will be most likely to benefit from primary prevention ICD, is urgently needed. Based on the latest literature, LVEF≤35% is still an independent predictor of all-cause and cardiovascular mortality in chronic systolic HF, and displays a better combination of sensitivity and specificity than 40% cut-off.[12] Finding ways to evaluate the SCD risk in patients with lower EF will be more efficient and economically significant. (page 4, Introduction section, line 21-26)

6. I suggest the authors to Include the appropriate references to the Table 1. References can be places at the end of the authors name.

Response: we added the references at the end of the author's name in Table 1. (page 15-16)

Table 1. The risk model for HF in the literature

Author	Database	Year	Variables (n)	Patients (n)	Endpoints
Agostoni[14]	MECKI	2012	6	2716	Cardiovascular death; urgent cardiac transplant
Barlera[15]	GISSI-HF	2013	14	6975	all-cause mortality
Collier[16]	EMPHASIS-HF	2013	10	2737	all-cause mortality
Komajda[17]	I-PRESERVE	2011	12	4128	all-cause mortality
Levy[18]	SHFM	2006	14	1125	Survival
O'Connor[19]	HF-ACTION	2012	4	2331	all-cause mortality
Pocock[20]	CHARM	2006	21	7599	all-cause mortality
Pocock[21]	MAGGIC	2012	13	39372	all-cause mortality
Senni[22]	CVM-HF	2006	13	292	all-cause mortality
Senni[23]	3C-HF	2013	11	2016	all-cause mortality; urgent heart transplant (1year)
Vazquez[24]	MUSIC	2009	10	992	all-cause mortality; cardiac mortality; pump failure death, sudden death
Nicole[25]	BARDICHE-index	2017	8	1811	all-cause mortality; all-cause hospitalization; CHF-related hospitalization

7. A recent review of the literature should be performed in order more recent studies to be included (studies reported after 2013).

Response: According to your suggestion, we have supplemented some new literature published in the last 2-3 years, and carefully revised the Introduction section. The supplemented literature involves the latest study on the prognosis of heart failure and the application of machine learning methods in heart failure. The papers are listed as follows. Please see the manuscript for the detailed revision. (page 4-6 section "Introduction")

The added papers are listed as follows:

[11] Aimo A, Januzzi J J, Vergaro G, et al. Left ventricular ejection fraction for risk stratification in chronic systolic heart failure[J]. Int J Cardiol, 2018.

[25] Uszko-Lencer N, Frankenstein L, Spruit M A, et al. Predicting hospitalization and mortality in patients with heart failure: The BARDICHE-index[J]. Int J Cardiol, 2017,227:901-907.

[26] Delgado V, Bucciarelli-Ducci C, Bax J J. Diagnostic and prognostic roles of echocardiography and cardiac magnetic resonance[J]. J Nucl Cardiol, 2016,23(6):1399-1410.

[27] Halliday B P, Cleland J, Goldberger J J, et al. Personalizing Risk Stratification for Sudden Death in Dilated Cardiomyopathy: The Past, Present, and Future[J]. Circulation, 2017,136(2):215-231.

[28] Kelesidis I, Travin M I. Use of cardiac radionuclide imaging to identify patients at risk for arrhythmic sudden cardiac death[J]. J Nucl Cardiol, 2012,19(1):142-152, 153-157.

[29] Martins D S M, Vidigal F M, Morao M A. Iodine-123-metaiodobenzylguanidine scintigraphy in risk stratification of sudden death in heart failure[J]. Rev Port Cardiol, 2013,32(6):509-516.

[33] Awan S E, Sohel F, Sanfilippo F M, et al. Machine learning in heart failure: ready for prime time[J]. Curr Opin Cardiol, 2018,33(2):190-195.

8. I suggest authors to avoid using big sentences because it is difficult for the reader to understand them. For example see the sentence starting at line 18 and ending at line 27.

Response: We greatly appreciate your suggestion. But we didn't find the sentence starting at line 18 and ending at line 27. After checking throughout the manuscript, we found a big sentence in page 5, and have revised it.

Original: Although currently some non-invasive factors, including mechanical dyssynchrony measured by echocardiography, myocardial fibrosis detected with cardiovascular magnetic resonance, and cardiac autonomic dysfunction assessed by 123-metaiodobenzylguanidine scintigraphy, have been evaluated to predict SCD in HF patients, [23] it is difficult to widely use them to predict SCD in large HF population.

Revised: In recent years, the advances in strain echocardiography[26, 27], cardiac magnetic resonance[26, 27] and cardiac radionuclide imaging[28, 29] have provided essential insights into the mechanisms of ventricular arrhythmias, and have been recommended to predict the SCD in patients with HF. (page 5, line 13-16)

9. I suggest the authors to remove from the aims of the study the evaluation of machine learning models. Once a model is proposed it is self-evident that it should be evaluated and validated.

Response: Thank you for your advice. We have removed this section from the section AIMS. (page 6, Aim section)

10. I suggest to the authors, taking into account the content of subsections, to change or remove the header "Statistical analysis".

Response: Thanks for your suggestion. We have remove the header "Statistical analysis", and upgrade the four subheadings: "data pre-processing", "machine learning", "COX proportional hazards regression" and "model validation" to the headings. (page 12-14)

11. The header "Data classification and pre-processing" does not correspond to the content of the paragraph that follows. I suggest the authors to use a header the phrase "Pre-processing of data" and provide more information how this will be performed.

Response: Thank you for your suggestion. We have revised the header "Data classification and pre-processing" to "Data pre-processing". (page 12)

12. The authors mention that information gain ranking method will be followed in order feature selection to be achieved. My first recommendation is to add an appropriate reference, my second recommendation is to explain why they select the specific method and my third recommendation is to test other feature ranking, as well as wrapper approach for feature selection.

Response: According to your suggestion, we have added an reference as follows:

They employ computer algorithms to identify patterns in large datasets with a large number of variables, analyze rules automatically and build both linear and non-linear models in order to make data-driven predictions or decisions.[31] (page 6, line 4-7)

[31] Quinlan J R. Induction of decision trees. Machine learning 1.1[M]. 1986.

Why: To reduce a bias towards multi-valued attributes by taking the number and size of branches into account when choosing an attribute.

Currently, we intend to use information gain first, and in the subsequent experiments, we will try use random forest for feature selection to compare their performance.

13. The authors should make more clear what will be the response feature of the classifiers e.g. SCD yes vs. SCD no or SCD high risk vs. SCD low risk.

Response: The purpose of this study is to develop the model to predict SCD high risk in specific HF population (EF≤ 35%). A lot of risk factors, derived from the previous studies, have been considered valuable for the prediction of prognosis and SCD in HF patients. (line 2-14 in page 9 and figure 2) The better combination of these risk factors may will be the response feature of the high SCD risk. SCD is defined by the World Health Organization as unexpected death that occurs within 1 hour from the onset of new or worsening symptoms (witnessed arrest) or, if unwitnessed, within 24h from when the individual was last observed alive and asymptomatic[36]. The lethal arrhythmia including VT/VF, sudden cardiac arrest, cardiopulmonary resuscitation, and rehospitalization due to HF will be recorded carefully.(page 11, line 8-14)

14. The authors should explain why they select the specific classifiers. For example ensemble classifiers like Random Forests may also be tested.

Response: Several classifiers, including decision trees, logistic regression, support vector machine, random forest, artificialand artificial neural network, will be used to develop the prediction models, and will be evaluated based on the accuracy and interpretability. Finally, clinical experts and computer specialists will discuss and choose the best model to predict the prognosis and of SCD in HF patients and then perform the further validation with the prospective dataset.

For your convenience, the related statements are pasted here(page 12-13, Machine learning section):

Prediction models for SCD prediction in HF patients will be developed by the following classification algorithms respectively: decision trees, logistic regression, support vector machine, random forest, artificialand artificial neural network. [29] The performance and general error estimation of these ML models will be assessed by 10-fold cross-validation. The dataset will be randomly divided into 10 equal folds. 9 folds will be used as the training set with the remaining one fold as the validation set. The validation results from 10 repeats will be combined to provide a measure of the overall performance. The prediction models derived from the above classification algorithms above will be evaluated based on the accuracy and interpretability. Finally, clinical experts and computer specialists will discuss and choose the best model to predict the prognosis and of SCD in HF patients and then perform the further validation with the prospective dataset.

15. The authors mention that the algorithms will be evaluated based on the accuracy and interpretability. According to my understanding the output of a decision tree can be interpreted. However, the interpretability of the output of an SVM or a neural network is not a straightforward process.

Response: Yes, the interpretability of the output of an SVM and Neural network is hard. We will evaluate the performance based on the accuracy, sensitivities, specificities and the area under the Receiver operating characteristic curve.

16. Please add an appropriate reference for the COX proportional-hazards regression process.

Response: We have added the reference for the COX proportional-hazards regression process in the section “COX proportional hazards regression”. (page 13, line 20)

For your convenience, the related statements are pasted here:

Univariable COX proportional --hazards modeling will be used to identify strong independent baseline candidate predictors for the primary outcome and secondary outcomes. We will use both forward and backward stepwise procedure to derive the multivariable COX proportional --hazards model with $p<0.05$ as the inclusion criterion. Every variable in the model will be multiplied by its β-coefficient, and the products will be summed to calculate the risk score. Risk function will be used to estimate the level of risk. The calculating formula is as follows.[37]

[37] Harrell F J, Lee K L, Califf R M, et al. Regression modelling strategies for improved prognostic prediction[J]. Stat Med, 1984,3(2):143-152.

Reviewer: 3
Reviewer Name: Maryam Panahiazar
Institution and Country: The University of California San Francisco, USA
Dear Authors, Your study protocol sounds very interesting and appropriate for the research plan and clinical trials. Your study protocol is very organized and well written and considers so many details which are necessary for such an important study. But I do suggest you for a major basic revision before I can give more detail suggestions/comments. Everything in this protocol is backed to at least 2 years ago. It seems you wanted to submit this study 2 years ago. The statistics, the facts/hypothesis, state of the art machines learning methods are back to 2 years and even more. They are more studies have been published recently using ML for outcome prediction, etc. I would like you to consider them in such an important protocol. Even the plan for data collections is back to 2016. I suggest to update this protocol as a present situation and state of the art machine learning methods and other studies which are happened in the last 2-3 years and I will be more than happy to review this protocol again after your revision.

Response:

We sincerely appreciate your suggestion. We started to make the plan on this study 2 years ago. After several revisions, the study protocol was determined. According to your suggestion, we have supplemented some new literature published in the last 2-3 years. The supplemented literature involves the latest studies on the prognosis of heart failure and the application of machine learning methods in heart failure. The prediction of SCD high risk in HF patients with low LVEF (≤35%) is still a problem. Machine learning has not been reported to be applied to SCD risk prediction based on large HF population. Therefore, this study is still of great clinical significance. We carefully revised the

section “Introduction”. Please see the manuscript for the detailed revision.(page 4-6, Introduction section)

The added papers are listed as follows

[11] Aimo A, Januzzi J J, Vergaro G, et al. Left ventricular ejection fraction for risk stratification in chronic systolic heart failure[J]. Int J Cardiol, 2018.

[25] Uszko-Lencer N, Frankenstein L, Spruit M A, et al. Predicting hospitalization and mortality in patients with heart failure: The BARDICHE-index[J]. Int J Cardiol, 2017,227:901-907.

[26] Delgado V, Bucciarelli-Ducci C, Bax J J. Diagnostic and prognostic roles of echocardiography and cardiac magnetic resonance[J]. J Nucl Cardiol, 2016,23(6):1399-1410.

[27] Halliday B P, Cleland J, Goldberger J J, et al. Personalizing Risk Stratification for Sudden Death in Dilated Cardiomyopathy: The Past, Present, and Future[J]. Circulation, 2017,136(2):215-231.

[28] Kelesidis I, Travin M I. Use of cardiac radionuclide imaging to identify patients at risk for arrhythmic sudden cardiac death[J]. J Nucl Cardiol, 2012,19(1):142-152, 153-157.

[29] Martins D S M, Vidigal F M, Morao M A. Iodine-123-metaiodobenzylguanidine scintigraphy in risk stratification of sudden death in heart failure[J]. Rev Port Cardiol, 2013,32(6):509-516.

[33] Awan S E, Sohel F, Sanfilippo F M, et al. Machine learning in heart failure: ready for prime time[J]. Curr Opin Cardiol, 2018,33(2):190-195.

According to the protocol, 500 retrospective cases (2016-2017) and the first 1000 prospective cases(2018-2019) will be used to develop the prediction models. And the next 1000 prospective(2018-2019) cases will be used for model validation. The flow diagram of the progress is illustrated in Figure 1. The retrospective data collection in the two sub-centers started in March 2017, not 2 years ago, and now all these retrospective cases are being followed-up. The prospective enrollment in all 14 sub-centers has started in January 2018. The study framework and process is summarized in Figure 3.

For your convenience, the related statements are pasted here:

The cases from January 2016 to December 2017 in the First Affiliated Hospital of Nanjing Medical University and Xiamen Cardiovascular Hospital Xiamen University will be collected retrospectively and followed-up prospectively. About 500 retrospective cases meet the inclusion criteria according to preliminary estimation. The prospective recruitment has started in the above 14 hospitals since January 2018. The retrospective cases and the first 1000 prospective cases will be used to develop the prediction models. And the next 1000 prospective cases will be used for model validation. The flow diagram of the progress is illustrated in Figure 1. (page 7, line 5-12)

The retrospective data collection in the two sub-centers started in March 2017, and prospective enrollment in all 14 sub-centers has started in January 2018. The follow-up period is scheduled to end in December 2019. The major part of data analysis will be performed from January to March 2020. The study framework and process is summarized in Figure 3. (page 14, line 17-21, Study timeframe section)

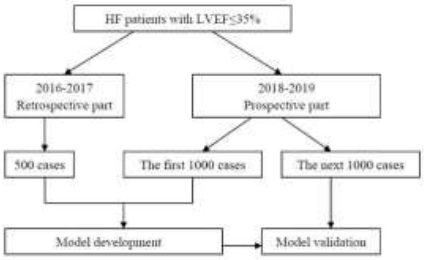

Figure 1 Flow diagram of progress

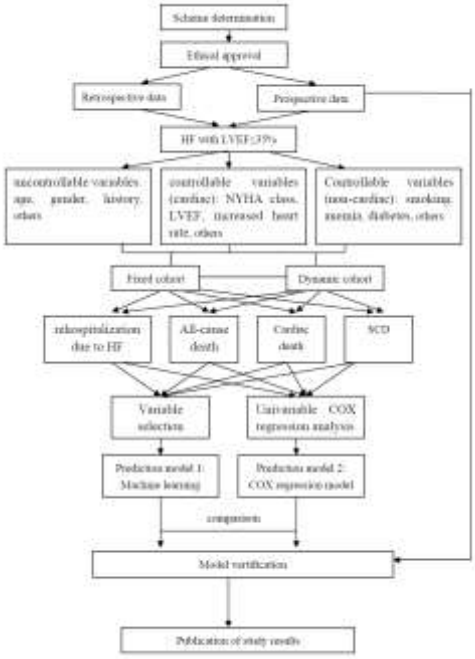

Figure 3 Study framework and process

VERSION 2 – REVIEW

REVIEWER	Avinainder Singh Brigham and Women's Hospital, Boston, MA
REVIEW RETURNED	31-Oct-2018

GENERAL COMMENTS	Revisions are acceptable. No further comments.

REVIEWER	Dimitrios I. Fotiadis University of Ioannina
REVIEW RETURNED	09-Nov-2018

GENERAL COMMENTS	The authors addressed my comments.

	I recommend them to transfer in the manuscript the information/explanations provided to comments 12 and 15 (List_of_Responses.pdf) of my previous review.

REVIEWER	Maryam Panahiazar University of California San Francisco, USA
REVIEW RETURNED	08-Dec-2018

GENERAL COMMENTS	Prediction of sudden cardiac death in heart failure patients with low left ventricular ejection fraction is a very well defined problem in the field and your work will have a high impact. Thanks for considering my suggestions and update your protocol based on current dates. And thank you for making your references and plan up to date as well. In term of machine learning approached please consider some previous works. I am not agreed with "However, ML has not been reported to be applied to SCD risk prediction based on large HF population". They are several studies in term of using machine learning to SDC risk predictions which I believe you need to consider them, I will be more than happy to share it. In some of them, they have not used the term ML per se but they have used ML methods as you mentioned. I appreciate this work because of the large-scale study cohort but if you want to claim ML, please make sure you will compare your models with other models or/and other studies in this field to choose the best model. Even though it is not exactly for the same target predictions. But you will need the comparison and optimization. You mentioned "The major part of data analysis will be performed from January to March 2020. " it is kind of nonrealistic to do all data analysis in 3 months. Please reconsider that in your study timeframe (as a suggestion). There are a few limitations of predictors of SCD in elderly population which I would like you to consider in your study, such as a diastolic function that could be difficult to assess in the elderly population. Also, HF signs and comorbidities sometimes are mimic. Please reconsider risk predictors. There is a new study came out early 2018 with Ayesta et al. and they discussed these risk predictors and limitations as well. This is a good work and I wish you all the best in this study.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Avinainder Singh

Institution and Country: Brigham and Women's Hospital, Boston, MA

Comments:

Revisions are acceptable. No further comments.

Response: Thank you very much for your approval.

Reviewer: 2

Reviewer Name: Dimitrios I. Fotiadis

Institution and Country: University of Ioannina

Comments:

The authors addressed my comments.

I recommend them to transfer in the manuscript the information/explanations provided to comments 12 and 15 (List_of_Responses.pdf) of my previous review.

Response: We greatly appreciate your suggestion and have revised the relevant content. For your convenience, the related statements are pasted here:

They employ computer algorithms to identify patterns in large datasets with a large number of variables, analyze rules automatically and build both linear and non-linear models in order to make data-driven predictions or decisions.[31] (page 6, line 4-7)

[31] Quinlan J R. Induction of decision trees. Machine learning 1.1[M]. 1986.

The prediction models derived from the above classification algorithms above will be evaluated based on the accuracy, sensitivities, specificities and the area under the Receiver operating characteristic curve.(page13, line 9-11)

Reviewer: 3

Reviewer Name: Maryam Panahiazar

Institution and Country: University of California San Francisco, USA

Comments:

Prediction of sudden cardiac death in heart failure patients with low left ventricular ejection fraction is a very well defined problem in the field and your work will have a high impact.

Thanks for considering my suggestions and update your protocol based on current dates. And thank you for making your references and plan up to date as well.

In term of machine learning approached please consider some previous works. I am not agreed with "However, ML has not been reported to be applied to SCD risk prediction based on large HF population". They are several studies in term of using machine learning to SDC risk predictions which I believe you need to consider them, I will be more than happy to share it. In some of them, they have not used the term ML per se but they have used ML methods as you mentioned.

I appreciate this work because of the large-scale study cohort but if you want to claim ML, please make sure you will compare your models with other models or/and other studies in this field to choose the best model. Even though it is not exactly for the same target predictions. But you will need the comparison and optimization.

You mentioned "The major part of data analysis will be performed from January to March 2020. " it is kind of nonrealistic to do all data analysis in 3 months. Please reconsider that in your study timeframe (as a suggestion).

There are a few limitations of predictors of SCD in elderly population which I would like you to consider in your study, such as a diastolic function that could be difficult to assess in the elderly population. Also, HF signs and comorbidities sometimes are mimic. Please reconsider risk predictors. There is a new study came out early 2018 with Ayesta et al. and they discussed these risk predictors and limitations as well. This is a good work and I wish you all the best in this study.

Response:

We greatly appreciate your comment and suggestion. We did miss some recent literature in term of SCD risk prediction by machine learning. The statement of "However, ML has not been reported to be applied to SCD risk prediction based on large HF population" is not appropriate. We have revised the Introduction section. For your convenience, the related statements are pasted here:

ML algorithms also have been applied to predict SCD in some recent studies and results indicate their significant advantages for predicting SCD.[34, 35] However, more studies based on large-scale cohort are needed to evaluate ML for prediction of SCD in HF patients. (page 6 line13-16)

[34] Ebrahimzadeh E, Foroutan A, Shams M, et al. An optimal strategy for prediction of sudden cardiac death through a pioneering feature-selection approach from HRV signal[J]. Comput Methods Programs Biomed, 2019,169:19-36.

[35] Au-Yeung W M, Reinhall P G, Bardy G H, et al. Development and validation of warning system of ventricular tachyarrhythmia in patients with heart failure with heart rate variability data[J]. PLoS One, 2018,13(11):e207215.

We also appreciate your suggestion for model comparison and optimization. Both ML and traditional multivariable COX proportional hazards regression models will be developed and compared in the prediction of SCD. Moreover, the ML model will be validated in a prospective study. Figure1 and Figure 3 show the flow diagram of model development and validation.

As for “data analysis time”, thank you for your suggestion. All data analysis performed in 3 months is really difficult. So, we adjusted our study timeframe. We will use 6 months to finish all data analysis. For your convenience, the related statements are pasted here: The major part of data analysis will be performed from January to June 2020. (page 14, line 22-23)

In our study, we focus on the HF patients with LVEF ≤35%, so the diastolic dysfunction is not considered as candidate risk factor. The recommended study of Ayesta et al. shows that the combination of different risk predictors may be the best option to predict SCD, which is consistent with our study. [Ayesta A, Martinez-Selles H, Bayes D L A, et al. Prediction of sudden death in elderly patients with heart failure[J]. J Geriatr Cardiol, 2018,15(2):185-192.] We reviewed the prognostic models of HF in recent years and summarized the candidate risk predictors which have been assessed and confirmed by an expert panel of cardiologists and statisticians. For your convenience, the related statements are pasted here:

Prognostic models of HF in the last 10 years have been reviewed, and the associated risk factors have been ranked according to their corresponding hazard ratio in respective risk models (Table 1, Figure 2). Age, sex, New York Heart Association (NYHA) class, LVEF, prior HF hospitalization, course of HF, severe valvular heart disease, atrial fibrillation, prior myocardial infarction / coronary artery bypass grafting (CABG), renal dysfunction, chronic obstructive pulmonary disease (COPD), diabetes mellitus, ischemic etiology, decreased systolic pressure, low body mass index (BMI), anemia, hyponatremia, high N-terminal pro-brain natriuretic peptide (NT-proBNP), uricemia, current smoker were included. Variables which were not listed in previous models but appear relevant to higher risk of SCD in HF patients, and would therefore, merit consideration, including syncope or pre-syncope, frequent premature ventricular beat, non-sustained ventricular tachycardia, complete left bundle branch block (CLBBB), long QT interval, increased QT dispersion. In addition, self-care ability, social

support and psychological state including depression and anxiety, are also predictors for subsequent poor prognosis in HF patients. The above risk factors have been assessed and confirmed by an expert panel of cardiologists and statisticians and will be collected in this study particularly. (page 9, line 2-18)