Anesthesiology and Pain Medicine
Volume 12 · Issue 4 · e127140

OPEN PEER REVIEW

Prediction of Acute Kidney Injury After Cardiac Surgery Using Interpretable Machine Learning

Azar Ejmalian, Atefe Aghaei, Shahabedin Nabavi, Maryam Abedzadeh Darabad, Ardeshir Tajbakhsh, Ahmad Ali Abin, Mohsen Ebrahimi Moghaddam, Ali Dabbagh, Alireza Jahangirifard, Elham Memary, Shahram Sayyadi

SUBMITTED	REVISED	ACCEPTED	DOI
23 Apr 2022	16 Aug 2022	4 Sep 2022	10.5812/aapm-127140

About this report

This open peer review (OPR) report shows the public comments made on this article during peer review, by the reviewers, the associate editor and the editor-in-chief, for every revision round, together with the authors' replies. Comments appear only with the consent of the people who wrote them. Reviewers who chose to sign their report are named; the others are shown as "Referee". Comments are edited for language only, and confidential remarks to the editors are not included.

More about open peer review: brieflands.com/knowledgebase/opr

Contents

1 **Original submission** Referee 1 2
2 **Revision 1** 3
Authors' reply to reviewers 4
Reviewer and editor comments 8

1 Original submission

1 comment · 4 Jul 2022

Referee 1 Reviewer

MINOR REVISION 4 Jul 2022

Dear author, the following items need to be revised:

- English revision in terms of grammar and text clarification is required. For example, it cannot be claimed so definitively that “Furthermore, more than 25% of patients who developed severe AKI will die within a year.” Such sentences need to be revised.
- Abbreviations should be defined at first mention in the following sections of the manuscript: title, abstract, text, each figure and table legend. AUC was not defined in the abstract. It also doesn’t match its definition (i.e., operating characteristics curve) in the body of the manuscript.
- The aim of the study is somewhat unclear: is it to determine the reliability of interpretable machine learning models to predict AKI after cardiac surgery, to compare several interpretable machine learning models in predicting AKI after cardiac surgery, or both?
- The last two sentences in the methods section, “AUC denotes the overall success of an experiment where Pr[TP] is a function of v = Pr[FP]. According to this, $AUC \approx 0.5$ and $AUC \approx 1$ reflect poor and good results, respectively”, would be more understandable if an acceptable range were specified, as none of the AUCs in the results were 1.
- Section 3-1- Experimental Setup is not related to the results and should be included in the methods.
- The first three sentences in the conclusion are not related to this section and need to be revised.

2

Revision 1

AUTHORS' RESPONSE TO REVIEWERS

Authors' reply to reviewers

The next pages contain the authors' point-by-point reply to the comments of this revision round, as submitted with the revised manuscript. It is published as the authors wrote it.

Response to reviews:

We wish to thank all reviewers for the time and effort they have put into thoughtfully considering our manuscript. All reviewer recommendations have been taken into account in this revised manuscript. Responses to specific reviewer comments are listed below.

Reviewer 1:
Thank you for your valuable manuscription.
My comment :
1-In methods you didn't specify exclusion criteria such as age<18 ,….
In many articles like your work, they excluded who were dialysis dependent at base line.
If you had exclusion criteria, please mention how many patients did exclude and why?

Response:

Done. The only exclusion criteria for this study is that the patients are over 18 years old. Considering the importance of balancing the number of data of different classes in machine learning techniques, the number of patients of both classes is almost balanced. Relevant changes were applied in the text.

2-There are many important risk factors such as nephrotoxic drugs, intra operative hypotention, post op sepsis. what is your idea about them?

Response:

Nephrotoxic drugs are included in the data as PDH. In other cases, the information that could be obtained retrogradely from the patients' files was collected. Things like the need to use inotropic drugs in POST OPERATIVE data were collected to separate hypotensive from normotensive patients. Finally, due to a large number of data variables in the time and place conditions available for data collection, it was impossible to add more variables. Also, the excessive increase of the investigated features causes machine learning methods to suffer decision-making problems due to the curse of dimensionality. Therefore, in this case, it has been acted based on literature and expert analysis.

3-In table 1 page 6 LAB; Are these variables about pre operation?
Some of them are unusual range, such as:
Cr =12 Bs =589 Hct=15

Response:

These values were included in the patients' files, and trustworthiness in transferring content has been considered in data collection. Therefore, although these such values appear as outliers, they have been included in the data collection based on extracting patients' files.

4-Why did you check Cr only in first and 7th day? During this period the patients may have raise in Cr.

Response:

Usually, the acute period for postoperative kidney injury is seven days. During these seven days, the amount of Cr may increase, go through an increasing process and then decrease or remain high. Since it

was not possible to check the amount of Cr every day, we took help from studying this amount on the first and seventh postoperative days to know whether the patient experienced AKI in these seven days or not.

5-what is importance of nephrologist consult in table? Did you use it?

Response:

Knowing whether the patient underwent a nephrology consultation or not would only help us to understand how serious the patient's kidney injury was. This postoperative data was not included in the input data of the machine learning methods.

6-what precent overall developed AKI?

Response:

The statistics related to the number of patients in each category, including AKI+ and AKI-, are fully described in section 3-2. It should be noted that the percentage of patients with AKI is 46.8 on the first day after surgery and 28.8 on the seventh day. It has been tried to maintain the balance of both classes as much as possible in data collection, which is essential in machine learning algorithms.

7-There isn't any other discussion about similar research in discussion.

Response:

A comparison with the study [10] is included in the discussion, which is the closest study in terms of purpose and results.

8-what is your limitation in this research?

Response:

Retrograde access to the data was one of the main limitations of this research. Unfortunately, many critical data in the patients' files that must be completed during the operation have not been recorded due to the lack of proper supervision of cardiac anesthesiologists in completing the files. This limitation is mentioned at the end of the discussion section.

Reviewer 2:
1.please rewrite abstract and more clear

Response:

Done. The abstract was again revised and improved.

2. retrospectively or prospectively
Response:

In conducting this study, data were collected both retrospectively and prospectively. In section 2-1, the amount of data collected for both groups has been mentioned.

3.please confirm how do you predict AKI after cardiac surgery, and explain about that

Response:

As the aim of this study shows, multiple machine learning techniques have been used to predict AKI after cardiac surgery. These machine learning techniques are mentioned in the text, and appropriate references are cited for further study.

4.How increase risk of AKI?

Response:

According to the literature, a set of risk factors can increase the probability of AKI after surgery, especially cardiac surgery. In our study, these risk factors have been analyzed based on their impact on the occurrence of AKI and are shown in Figures 2 and 3.

5.What's your manuscript novelty?

Response:

Using most of the machine learning techniques to predict AKI after cardiac surgery, choosing the best algorithm, collecting high-volume dataset, and using interpretable approaches that are among the latest methods to confirm the results of the predictions obtained from the algorithms are the main contributions of this study. Specifically, interpretability can be the main novelty of this study, which shows the value of each feature in prediction. Therefore, by examining the interpretability results, the therapy team can ensure the results of the predictions.

Reviewer 3:

Dear author the following items needs to be revised:

- English revision in terms of grammar and text clarification is required. For example, it cannot sharply be claimed that "Furthermore, more than 25% of patients who developed severe AKI will die within a year." , and such sentences are needed to be revised.

Response:

Done. The review of the text in terms of writing and grammatical errors was done by a native reviewer.

It should be noted that the sentence mentioned by the honorable reviewer was quoted from a reference, which is cited at the end of the following sentence.

- Abbreviations should be defined at first mention in the following sections of the manuscript: title, abstract, text, each figure and table legend. AUC was not defined in the abstract. Even it doesn't match its definition (i.e., operating characteristics curve) in the body of the

manuscript.

Response:

Done. Revisions were made in this regard.

- The aim of the study is somehow not clear; to determine the reliability of interpretable machine learning models to predict AKI after cardiac surgery, or to compare several interpretable machine learning models in predicting AKI after cardiac surgery, or both?

Response:

In this study, the prediction of AKI after cardiac surgeries has been done using several machine learning techniques. The results of these techniques have been compared to choose the best machine learning method, and the best technique has been selected. Then, using three methods to control the interpretability of machine learning methods, the interpretability of the results of the predictions was checked.

- The last two sentences in the methods section "AUC denotes the overall success of an experiment where Pr[$\mathcal{TT}$] is a function of v = Pr[$\mathcal{FT}$]. According to this, $\mathcal{AUC} \approx 0.5$ and $\mathcal{AUC} \approx 1$ reflect poor and good results, respectively", it will be more understandable if an acceptable range was determined, as none of the AUCs in the results were 1.

Response:

Done. The mentioned sentence meant the range of possible values. Therefore, this sentence was corrected because it was misleading.

- The section 3-1- Experimental Setup is not related to results and should be included in methods.

Response:

Done.

- The first three sentences in conclusion is not related to this section and needs to be revised.

Response:

Done.

2 Revision 1

0 comments

No public comments in this round.