Research Proposal

1. TITLE

A Machine Learning Approach of Logistic Organ Dysfunction Score Estimation with Data Acquired from Bedside in ICU

2. BACKGROUND AND RATIONALE

Logistic Organ Dysfunction Score (LODS), designed in 1996 by J.R.Le Gall for ICU, is a organ dysfunction score system with weighted variables with the worst values in first 24h that reflect severity level within an organ system and among organ systems. It is considered as an outcome risk prediction scoring system since an equation is included that converts LOD score into probability of mortality. In the later research, it is confirmed to be usable for the first 7 days in ICU, and is widely used in other departments as well. However, it is much less used than the similar score, Sequential Organ Failure Assessment (SOFA), as SOFA can be used everyday and has almost same accuracy. Moreover, there are already some outcome risk prediction scores. So, although LODS is a hybrid score system, it is used as widely as SOFA.

Medical score system needs accuracy to reflect patients' condition, and ease of use to reduce clinician time in using and learning. Only if a score system is substantially better than another one, proved by a large number of research, it will be used in practice. So improving an existed score system is more feasible than replacing, and change should be small but useful.

Designing score system is like a trade-off between accuracy and usability. Excellent accuracy always work with poor usability. Moreover, bad usability affects clinicians using the score more than accuracy. There are many researches target to accuracy that will save more patients' life. However, only a good usability will let clinicians start using the score system. So this research target to usability, so that more clinicians will start to use LODS.

3. RESEARCH QUESTION(S)

According to the characteristics of score system, there are two things to improve, accuracy and usability. This research will concentrate on the usability. LODS derives 12 variables which depend heavily on laboratory measurements. To this end, this research will find an computational solution to estimate LODS with values that can be acquired from bedside in the first 7 days in ICU.

Usability and accuracy are always coupled. One is improved, then another one is reduced. Is it possible to substantially improve usability and affect accuracy a little? This research will target to this and propose a solution with help of machine learning.

The traditional score system select laboratory results to check organ function, for instance bilirubin amount is used to check liver function. Body is a system which one organ dysfunction will affect others. So is it possible to use some vital signs or other easy-acquired values to check organ situation? After all, the methods to get the sample for laboratory check need specific

equipment and environment, and cost time, which is expensive. This research will find some metrics to reflect organ function without the expensive cost.

4. RESEARCH METHODOLOGY

Based on the development of machine learning and the increasing usage of electric health record (EHR) system, using machine learning algorithms to analytic EHR data could find potentiality and assist clinicians. There have been researches to use bedside value to estimate patients' severity in first 24h in ICU, use vital signs to monitor specific disease severity by estimating SOFA and use deep learning algorithms to improve score system. This research will combine several algorithms and use a big medical dataset to propose a model to improve the LODS.

A subset of Medical Information Mart in Intensive Care (MIMIC) IV dataset is used in this experiment. The dataset consists a large number of ICU data from Beth Israel Deaconess Medical Center (BIDMC) in Boston between 2008 and 2019. It has pre-existing institutional review board (IRB) approval.

This research only works on adults' (>18 years) ICU records, and more than 24h ICU stays will be included. So it will only improve LODS usage in adult-patients in ICU. The usage in other departments and paediatric will be excluded.

Usability will be compared between new solution and current methods by the amount of variables, the difficulty to get values, and the complexity to calculate the solution. Brier Score will be used to measure the calibration. As the essential and most used metrics for regression analysis, Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) will be used to assess the prediction performance. Hosmer-Lemeshow chi-square statistic will be used to check the solution fit. SOFA value of each case can also be used as a standard to compare the result.

5. PLAN OF WORK & TIME SCHEDULE

This research consists data preparation and apply algorithms to analytic data. In each part, tools and algorithms selection are necessary.

- More literature review (3.4.2023-23.4.2023)
- Data preparation (24.4.2023-28.5.2023)
  - Acquire necessary data
  - Remove data out of scope
  - Process missing data, insert or remove (includes algorithm selection)
- Use algorithms to select variables (29.5.2023-2.7.2023)
- Prepare data with selected variables (3.7.2023-16.7.2023)
- Apply algorithms on processed data to propose the model (17.7.2023-3.9.2023)
- Analyse and interpret results (4.9.2023-17.9.2023)
- Finish 1st version thesis (1.4.2023-1.10.2023) With the first version, literature review, methods, algorithms, results and discussion will be included, with most of the experiment details.
- Finish final version (15.11.2023) All experiment details will be included. If the research goes well, an academic article will also be finished at this time.

6. BIBLIOGRAPHY

Le Gall JR, Klar J, Lemeshow S, Saulnier F, Alberti C, Artigas A, Teres D. The Logistic Organ Dysfunction system. A new way to assess organ dysfunction in the intensive care unit. ICU Scoring Group. JAMA. 1996 Sep 11;276(10):802-10. doi: 10.1001/jama.276.10.802. PMID: 8769590.

Johnson AE, Kramer AA, Clifford GD. A new severity of illness scale using a subset of Acute Physiology And Chronic Health Evaluation data elements shows comparable predictive accuracy. Crit Care Med. 2013 Jul;41(7):1711-8. doi: 10.1097/CCM.0b013e31828a24fe. PMID: 23660729.

Tunç Aşuroğlu, Hasan Oğul, A deep learning approach for sepsis monitoring via severity score estimation, Computer Methods and Programs in Biomedicine, Volume 198, 2021, 105816, ISSN 0169-2607

Timsit J-F, Fosse J-P, Troché G, et al.Calibration and discrimination by daily Logistic Organ Dysfunction scoring comparatively with daily Sequential Organ Failure Assessment scoring for predicting hospital mortality in critically ill patients. Critical care medicine. 2002;30(9):2003– 2013. [PubMed: 12352033]

Pettilä V, Pettilä M, Sarna S, Voutilainen P, Takkunen O. Comparison of multiple organ dysfunction scores in the prediction of hospital mortality in the critically ill. Critical Care Medicine. 2002;30(8):1705–1711. [PubMed: 12163780]

Vincent J-L, Moreno R, Takala J, et al.The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. In: Springer-Verlag; 1996.