MEDINFO23

8 – 12 JULY 2023 | SYDNEY, AUSTRALIA

A hemodialysis mortality prediction model based on active contrastive learning

Yu WANG

Associate researcher

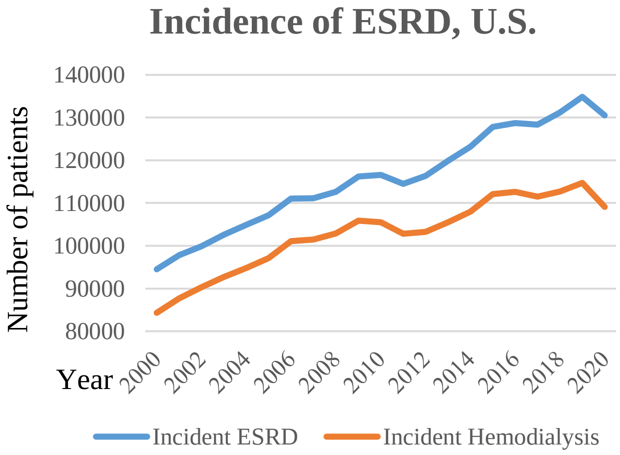
Zhejiang Lab

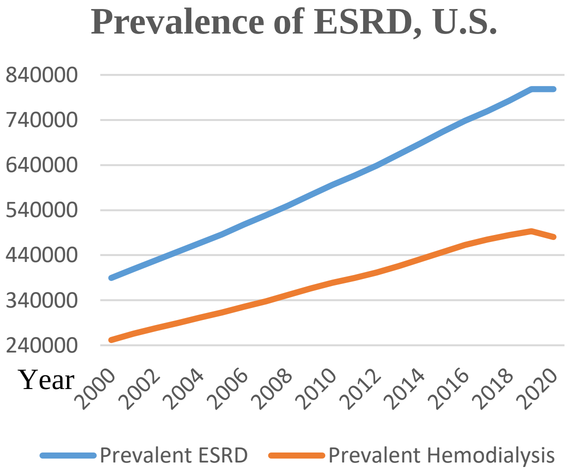
CONTENTS

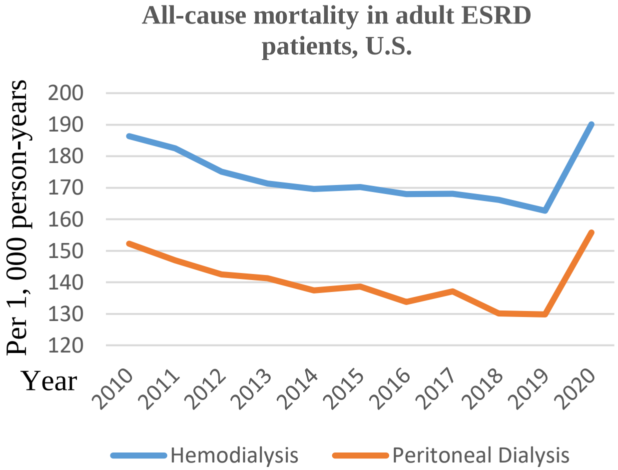
- 1. Introduction
- 2. Methods
- 3. Results
- 4. Discussion and conclusions
- 5. Acknowledgements and references

1. Introduction

- Hemodialysis (HD) is the main treatment for end-stage renal disease with high mortality and heavy economic burdens.



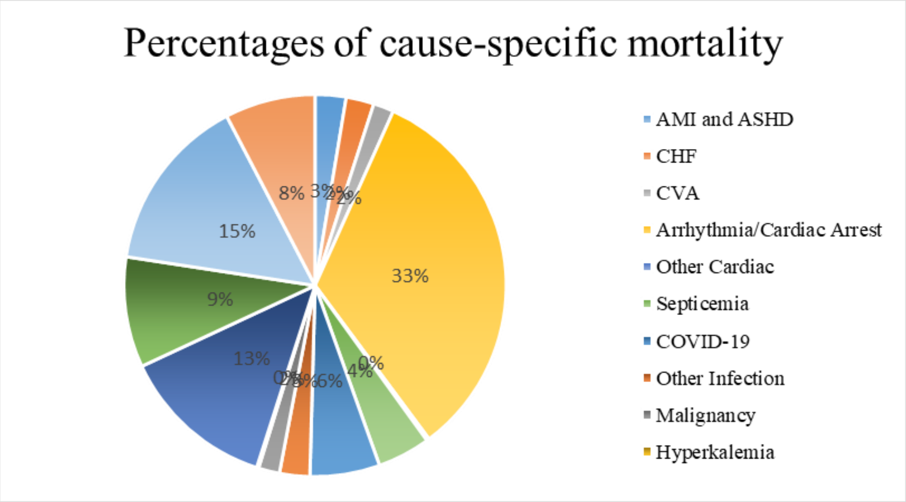




Data from USRDS Annual Data Report: 2022

1. Introduction

- Controlling the mortality and identifying risk factors of patients undergoing maintenance HD are of great clinical significance.



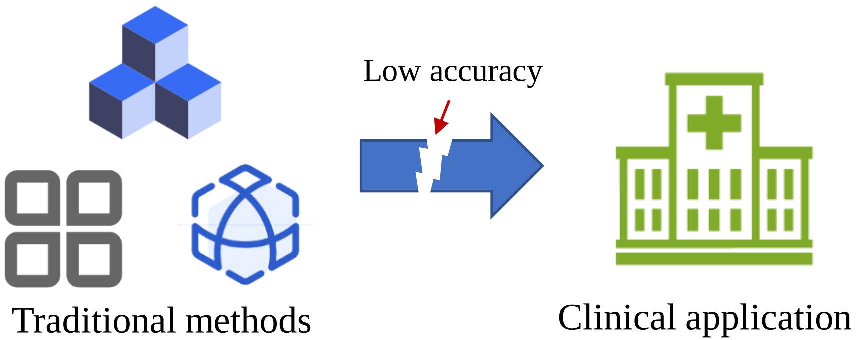
Mortality in patients with ESRD receiving hemodialysis

cause-specific mortality	**Percentages**
AMI and ASHD	2.6%
CHF	2.4%
CVA	1.7%
Arrhythmia/Cardiac Arrest	33.1%
Other Cardiac	0.2%
Septicemia	4.5%
COVID-19	5.9%
Other Infection	2.6%
Malignancy	1.8%
Hyperkalemia	0.2%
Withdrawal	13.1%
All Other Causes	9.4%
Unknown Causes	14.9%
Missing	7.7%

Data from USRDS Annual Data Report: 2022

1. Introduction

- This study aimed to develop and validate a more accurate and generalized AI model for predicting first-year mortality in patients undergoing maintenance HD.



More accurate and generalized AI model !

Traditional methods are not generalizable or accurate enough.

2. Methods - Study population

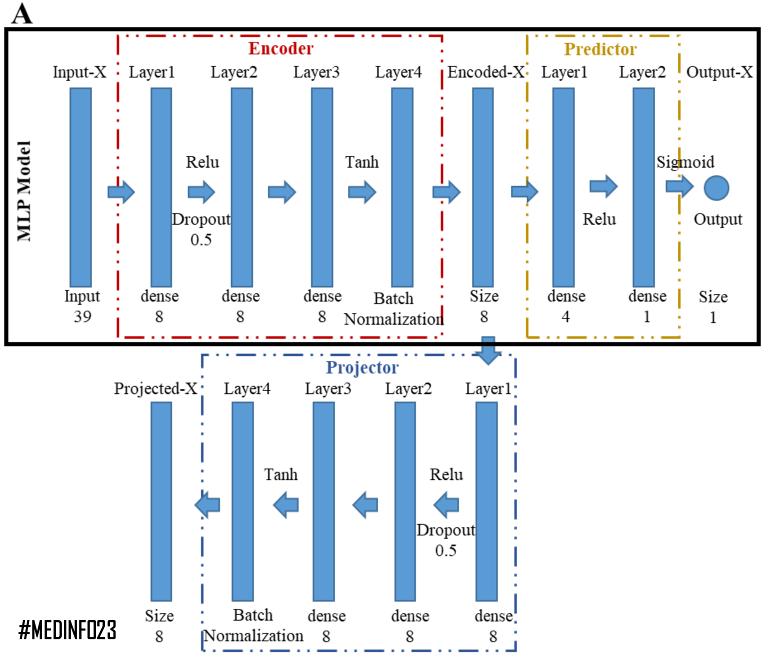
Data Source	EHR system of the First Affiliated Hospital of Zhejiang University (FAHZJU), PRC
Time	Patients received HD treatment between January 1, 2000, and August 20, 2016
Cutoff time of observation	20-Aug-17
Exclusion criterion	*According to the definition of maintenance HD, patients who died within 3 months after starting dialysis were excluded.

Features	**Count**
Demographic information	9
Diagnoses	9
Laboratory test results	12
Drugs	9
Total	39

Ultimately, we included 1,229 ESRD patients (survival: 1193; death: 36) on maintenance HD.

2. Methods - Model architectures

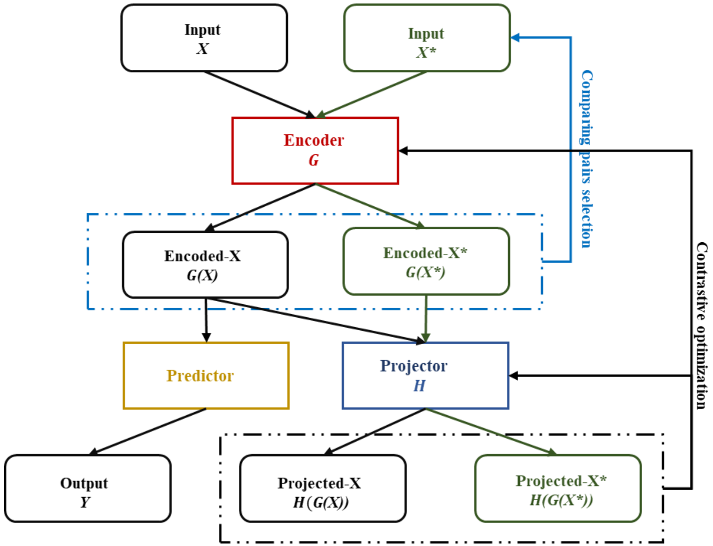
ACL model architecture (A) and flowchart (B)



- The MLP model consists of an Encoder and a Predictor.
- The ACL model consists of MLP model and a Projector.

2. Methods - Training and evaluation

B Training flowchart (B)



- Training
  - First stage – Train MLP model with cross-entropy loss
  - Second stage – Actively selected the most representative comparing sample pairs, the ones with the largest, smallest and random distances. And calculated a contrastive loss to optimize the Encoder and the Projector.
- Evaluation
  - 5-fold cross validation (80% for training, 20% for validation)
  - verified F1-score and AUROC across MLP model, ACL model, SVM and RF

3. Results

In total, we verified F1-score and AUROC across MLP model, ACL model, SVM and RF by 5-fold cross validation (80% for training, 20% for validation). The comparing results is shown in Table1. We used the t-SNE algorithm to visualize the distribution of features of Input layer, Encoded-X layer and Projected-X layer across training dataset and validation dataset in Figure 2.

Table 1. Model performance across the 5 folds

Model	F1-score [95% CI]	AUROC [95% CI]
SVM	0.667 [0.667-0.667]	0.561 [0.522-0.600]
RF	0.702 [0.663-0.741]	0.736 [0.680-0.792]
MLP	0.797 [0.752-0.843]	0.830 [0.794-0.866]
ACL	0.816 [0.777-0.855]	0.851 [0.819-0.883]

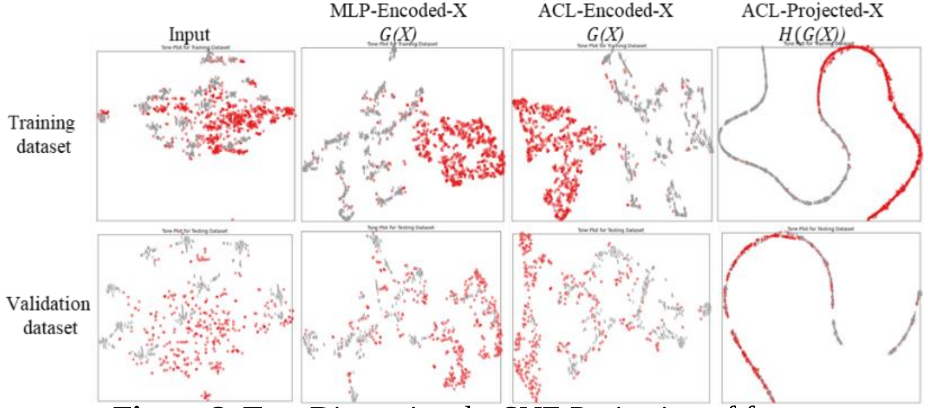


Figure 2. Two-Dimensional t-SNE Projection of features at different layers in ACL model. (Red: Survival, Grey: Death)

4. Discussion and conclusions

- The ACL model proposed in this study outperformed other models in predicting first-year mortality.
- The two-stage protocol of ACL can steadily improve the performance of the MLP model. This study has important clinical implications for other studies.
- This work is generalizable to analyses of cross-sectional EHR data, while this two-stage approach can be applied to other diseases as well.

5. Acknowledgements and references

Acknowledgements

This work was supported by the Zhejiang Provincial Natural Science Foundation of China (No. LQ21F030006 and No. LQ21H180002), the National Natural Science Foundation of China (No. 82001925), and the Key Research Project of Zhejiang Lab (No. 2022ND0AC01).

References

[1] USRD S. 2022 USRDS Annual Data Report: 2022 Annual Data Report. Volume 2 – End-stage Renal Disease (ESRD) in the United States: Incidence, Prevalence, Patient Characteristics, and Treatment Modalities2022.

[2]Wang Y, Zhu Y, Lou G, Zhang P, Chen J, Li J. A maintenance hemodialysis mortality prediction model based on anomaly detection using longitudinal hemodialysis data. Journal of Biomedical Informatics. 2021;123:103930, doi: 10.1016/j.jbi.2021.103930.

[3]Zhou L, Zeng X-X, Fu P. Community hemodialysis in China: opportunities and challenges. Chinese Medical Journal. 2017;130(18):2143-6, doi: 10.4103/0366-6999.213961.

[4]Sheng K, Zhang P, Yao X, Li J, He Y, Chen J. Prognostic machine learning models for first-year mortality in incident hemodialysis patients: Development and validation study. JMIR medical informatics. 2020;8(10):e20578, doi: 10.2196/20578.

[5]Floege J, Gillespie IA, Kronenberg F, Anker SD, Gioni I, Richards S, et al. Development and validation of a predictive mortality risk score from a European hemodialysis cohort. Kidney international. 2015;87(5):996-1008, doi: 10.1038/ki.2014.419.

[6]Chua H-R, Lau T, Luo N, Ma V, Teo B-W, Haroon S, et al. Predicting first-year mortality in incident dialysis patients with end-stage renal disease-the UREA5 study. Blood purification. 2014;37(2):85-92, doi: 10.1159/000357640.

[7]Doi T, Yamamoto S, Morinaga T, Sada K-e, Kurita N, Onishi Y. Risk score to predict 1-year mortality after haemodialysis initiation in patients with stage 5 chronic kidney disease under predialysis nephrology care. PloS one. 2015;10(6):e0129180, doi: 10.1371/journal.pone.0129180.

[8]Ramspek CL, Voskamp PW, Van Ittersum FJ, Krediet RT, Dekker FW, Van Diepen M. Prediction models for the mortality risk in chronic dialysis patients: a systematic review and independent external validation study. Clinical epidemiology. 2017;9:451, doi: 10.2147/CLEP.S139748.

[9]Akbilgic O, Obi Y, Potukuchi PK, Karabayir I, Nguyen DV, Soohoo M, et al. Machine learning to identify dialysis patients at high death risk. Kidney international reports. 2019;4(9):1219-29, doi: 10.1016/j.ekir.2019.06.009.

[10]Hadsell R, Chopra S, LeCun Y, editors. Dimensionality reduction by learning an invariant mapping. 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06); 2006: IEEE, doi: 10.1109/CVPR.2006.100.

[11]Nasrabadi NM. Book Review: Pattern Recognition and Machine Learning. SPIE; 2007.

[12]Zeiler MD. Adadelta: an adaptive learning rate method. arXiv preprint arXiv:12125701. 2012, doi: 10.48550/arXiv.1212.5701.

[13]Chollet F. Keras. https://github.com/fchollet/keras: GitHub; 2015.

[14]Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, et al., editors. TensorFlow: a system for Large-Scale machine learning. 12th USENIX symposium on operating systems design and implementation (OSDI 16); 2016.

THE END

THANKS