

High performance of privacy-preserving acute myocardial infarction auxiliary diagnosis based on federated learning: a multicenter retrospective study

Jie Xu^{1#}, Yu Zhang^{2#^}, Huamin Yu^{3#^}, Bo Lin^{4,5,6#^}, Peng Xiang^{7^}, Te Lin¹, Huizhe Lu¹, Guiying Zhang⁶, Dejian Wang^{6*^}, Hong Yuan^{3*}, Bin Hu^{8*}, Jun Jiang^{2*^}

¹Information Center, the Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China; ²Department of Cardiology, the Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China; ³Department of Cardiovascular Medicine, the First People's Hospital of Linping, Hangzhou, China; ⁴College of Computer Science and Technology, Zhejiang University, Hangzhou, China; ⁵Innovation Centre for Information, Binjiang Institute of Zhejiang University, Hangzhou, China; ⁶Department of R&D, Hangzhou Healink Technology, Hangzhou, China; ⁷Department of Vasculocardiology, Institute of Computing Innovation, Zhejiang University, Hangzhou, China; ⁸Department of Cardiovascular Medicine, Jiande First People's Hospital, Hangzhou, China

Contributions: (I) Conception and design: J Xu, Y Zhang, H Yu, B Lin; (II) Administrative support: D Wang, H Yuan, B Hu, J Jiang; (III) Provision of study materials or patients: T Lin, Y Zhang, H Yu; (IV) Collection and assembly of data: H Lu, Y Zhang, H Yu, G Zhang; (V) Data analysis and interpretation: B Lin, P Xiang, D Wang, H Yuan, B Hu, J Jiang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

^{*}These authors contributed equally to this work.

Correspondence to: Dejian Wang. Department of R&D, Hangzhou Healink Technology, Hangzhou, China. Email: wangdejian@healink.com; Hong Yuan. Department of Cardiovascular Medicine, the First People's Hospital of Linping, Hangzhou, China. Email: Yuanhongy@163.com; Bin Hu. Department of Cardiovascular Medicine, Jiande First People's Hospital, Hangzhou, China. Email: 15990318@qq.com; Jun Jiang. Department of Cardiology, the Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China. Email: jiang-jun@zju.edu.cn.

Background: Multicenter medical research is becoming a new trend. However, interagency data privacy security protection has become a major bottleneck of multicenter research. Therefore, to overcome this privacy protection issue, the aim of the present study was to apply a self-developed privacy-preserving machine learning framework for researchers who can build models on medical data from multiple sources, while providing privacy protection for both sensitive data and the learned model.

Methods: Based on Arya, a novel privacy computing platform developed by Healink, we constructed a privacy-preserving federated learning (FL) model using the fully connected neural network with datasets from 2–3 individual medical institutions. In the dataset, 80% of records were used for joint modeling on acute myocardial infarction (AMI) diagnosis. Modeling efficacy was evaluated with the remaining 20% of records. As the control, 1,500 medical records from 1 medical institution were used for single-center modeling and efficacy evaluation. During the process, the original data were still kept in individual hospital without moving or transferring out of the hospital. The diagnostic efficacy (sensitivity, positive predictive value, and accuracy) was evaluated.

Results: Our privacy-preserving FL model gives reliable AMI diagnostic efficacy. Three-center modeling (79% sensitivity, 88% positive predictive value, and 82.3% accuracy) and two-center modeling (77.8% or 77.6% sensitivity, 86.7% or 85.30% positive predictive value, and 81% or 79.7% accuracy) achieved relative high diagnostic efficacy; and single-center modeling achieved relative low diagnostic efficacy (76% sensitivity, 84.7% positive predictive value, and 79% accuracy).

Conclusions: The Arya privacy computing platform is efficient and practical for the FL model, which

[^] ORCID: Yu Zhang, 0000-0002-6688-5927; Huamin Yu, 0000-0003-3934-6095; Bo Lin, 0000-0001-5682-2140; Peng Xiang, 0000-0001-9617-597X; Dejian Wang, 0000-0002-6724-1859; Jun Jiang, 0000-0001-6926-9516.

could promote multicenter medical research securely without sacrificing diagnostic efficacy.

Keywords: Arya privacy computing platform; acute myocardial infarction; multicenter modeling; federated learning

Submitted Aug 12, 2022. Accepted for publication Sep 15, 2022.

doi: 10.21037/atm-22-4331

View this article at: <https://dx.doi.org/10.21037/atm-22-4331>

Introduction

With the development of medical research, there is a strong requirement for the quality of multicenter medical research, which has many significant advantages over single-center medical research, including sufficient data size, improved generalizability, and reproducibility of the research outcomes (1). Multicenter medical research aims to strengthen collaborations among institutions, promote new discoveries with pooled dataset from multiple sources, and accelerate the translation of research outcomes into clinical practice (2).

However, many medical institutions are unwilling to share their data, because data sharing could increase privacy security risks. The leakage of sensitive medical information to researchers, other institutions, and unauthorized users can result in financial loss, social discrimination, and data abuse. All these concerns hinder the collaborative benefits of multicenter medical research. Therefore, there is an urgent need to develop a framework to support multicenter medical research efficiently, by technically breaking data islands and encouraging joint model without moving or transferring original data. To solve this problem, several privacy-enhancing technologies have been developed, including privacy-preserving machine learning, secure multiparty computation (MPC), homomorphic encryption, differential privacy, and federated learning (FL) (3-7).

MPC is a collaborative computing technology to protect privacy among a group of untrusted participants. MPC should ensure the independence of input, correctness of calculation, decentralization, and other characteristics, and does not disclose each input value to other members participating in the calculation. It is mainly aimed at the problem of how to safely calculate a convention function without a trusted third party (8,9). At the same time, each participant cannot get any input information of other entities except the calculation results.

FL is a machine learning framework that was first proposed by Google in 2016, and is expected to become the

basis of next-generation artificial intelligence collaborative algorithms and networks. FL can effectively determine data use and machine learning modeling among multiple participants or computing nodes on the premise of ensuring information security during big data exchange, protecting the privacy of terminal data and personal data, and ensuring legal compliance (10,11).

Blockchain (BT) is a new application mode of computer technology and includes distributed data storage, point-to-point transmission, consensus mechanisms, and encryption algorithms, characterized by decentralization and transparency so that database records can be shared safely (12,13). At present, it is being studied all over the world and can be widely used in several fields, including the medical scene.

The Arya privacy computing platform, which was developed by Hangzhou Healink Technology in 2020, integrates the latest technologies such as FL, secure MPC, distributed machine learning, and BT. In this current study, we present a privacy-preserving FL model with medical data from 3 individual hospitals based on the Arya privacy computing platform, and then evaluate model efficacy based on acute myocardial infarction (AMI) diagnosis. The AMI diagnostic model would benefit doctors and people in countryside as an auxiliary diagnosis tool. We present the following article in accordance with the STARD reporting checklist (available at <https://atm.amegroups.com/article/view/10.21037/atm-22-4331/rc>).

Methods

Datasets

The dataset was provided by 3 individual medical institutions (the Second Affiliated Hospital, Zhejiang University School of Medicine; Jiande First People's Hospital; and the First People's Hospital of Linping). The dataset from the Second Affiliated Hospital, Zhejiang University School of Medicine consisted of 1500 records

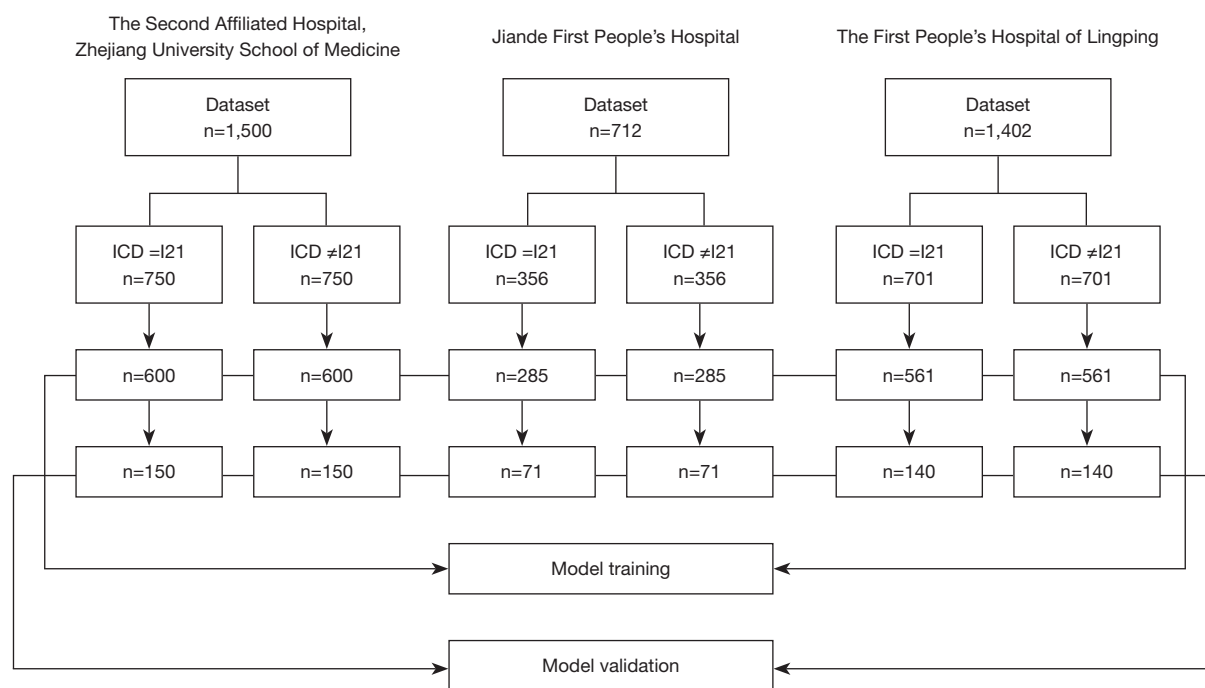


Figure 1 Data enrollment for model training and validation. ICD, International Coding Diagnosis.

(each record corresponding to a patient), including 750 AMI [International Coding Diagnosis (ICD) code =I21] patients (denoted by D_{pos1}) and 750 non-AMI (ICD code \neq I21) patients (denoted by D_{neg1}). The dataset from Jiande First People's Hospital consisted of 712 records, including 356 AMI (ICD code =I21) patients (denoted by D_{pos2}) and 356 non-AMI (ICD code \neq I21) patients (denoted by D_{neg2}). The dataset from the First People's Hospital of Linping consisted of 1,402 records, including 701 AMI (ICD code =I21) patients (denoted by D_{pos3}) and 701 non-AMI (ICD code \neq I21) patients (denoted by D_{neg3}). All the data were not merged and were stored in the information center of each hospital separately. A total of 12 biochemical indicators (troponin-I, troponin-T, creatine kinase isoenzymes, aspartate aminotransferase, homocysteine, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, total cholesterol, triglycerides, total bilirubin, indirect bilirubin, and direct bilirubin) were included in the study. The dataset was split into the training set and validation set in both a class-level and institution-level stratified manner. In total, 80% of records were used in the model training phase to compute a model for AMI prediction with 12 covariates, and the remaining 20% of records were used in the testing phase to assess the accuracy of the model with the 12 given covariates. As shown in

Figure 1, 600 randomly selected AMI patients and 600 non-AMI patients from the Second Affiliated Hospital, Zhejiang University School of Medicine, 285 AMI patients and 285 non-AMI patients from the Jiande First People's Hospital, and 561 AMI patients and 561 non-AMI patients from the First People's Hospital of Linping were included in the training set; the remaining patients were included in the validation set. The process of data enrollment for model training and validation is shown in Figure 1. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the ethics committee of the Second Affiliated Hospital, Zhejiang University School of Medicine (No. 2022-0695). The other 2 hospitals (Jiande First People's Hospital, the First People's Hospital of Linping) were also informed and agreed the study. And written informed consent was waived, as the data were analyzed retrospectively in a privacy-preserving manner.

Data preprocessing

Not all patients had all 12 biochemical indicators, which causes the unfilled items, or missing values, in the dataset. Missing values in D_{pos1} , D_{neg1} , D_{pos2} , D_{neg2} , D_{pos3} , D_{neg3} were handled separately. The missing value v_m^i of an indicator i

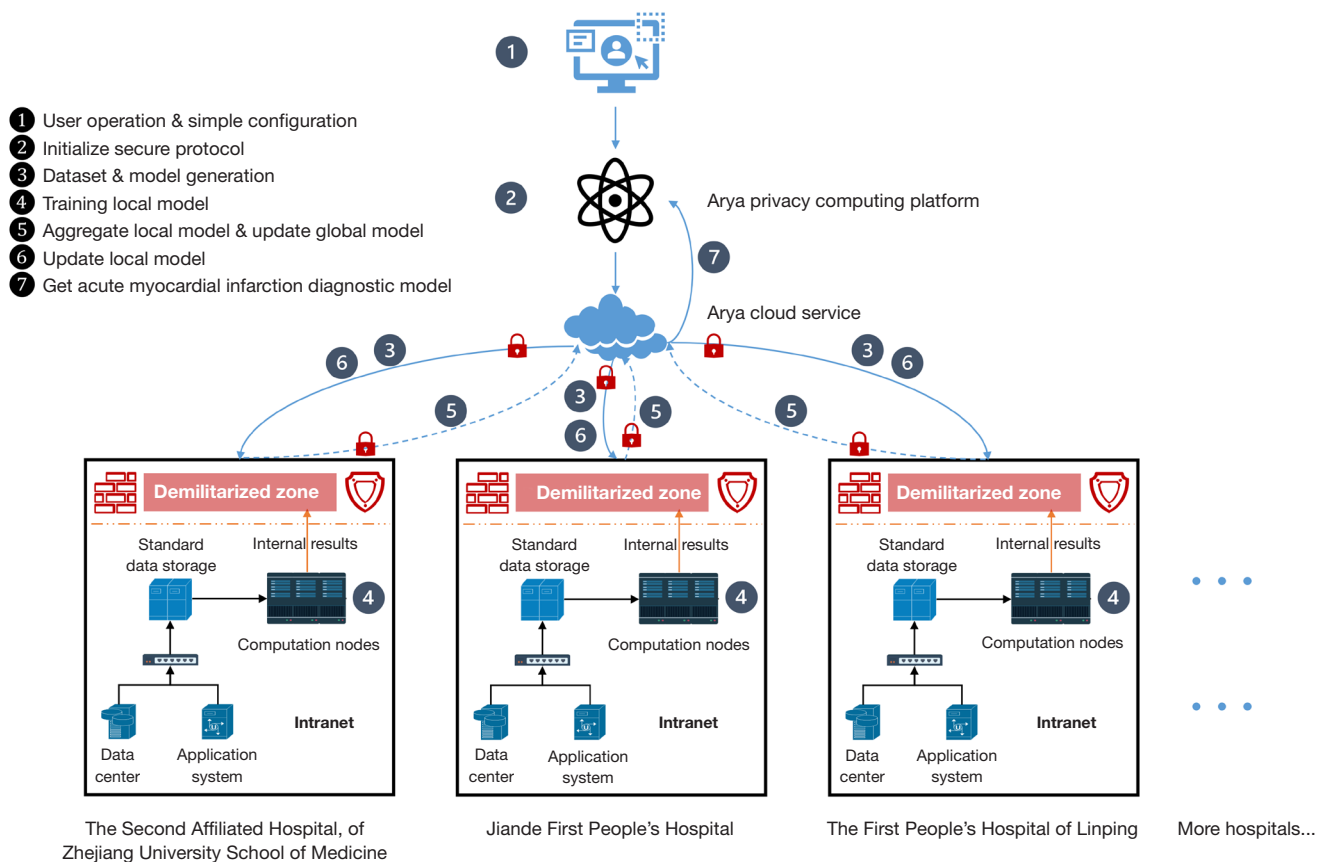


Figure 2 Construction process of the federation model.

is processed by filling with the mean value of the column as follows:

$$v_m^i \leftarrow \frac{\sum V_n^i}{|V_n^i|} \quad [1]$$

where v_n^i is the set of non-missing values of indicator i , and $|\cdot|$ is the number of a set. Each column of indicators in the data subset is processed separately based on this method.

Data quality assessment

Data quality was evaluated and divided into low-quality data (the number of missing values in a record >2) and high-quality data (the number of missing values in a record <2).

FL workflow on Arya

The FL workflow on Arya used in this study is shown in Figure 2. First, the user defined the dataset and model

configurations, including features, labels and hyper-parameters, and selected collaborative hospitals on Arya. Then, the user initiated the multi-center collaborative modeling task when the basic model was configured. Second, the Arya cloud service exchanged the keys of collaborative hospitals using the Diffie–Hellman algorithm, and adopted the secure aggregation algorithm and secure transport protocol to protect data security and user privacy during distributed collaborative computing in the modeling process. Third, the Arya cloud service sent the configuration of the dataset and model to the servers at the hospitals. Based on the dataset configuration, the servers selected matched the dataset from their own databases separately. Then, the computation nodes of the hospitals received the model configuration and built their own local artificial intelligence models. In the present study, the local model applied to the aforementioned fully connected neural network. Fourth, after initializing local models and other environment settings, the computation nodes of

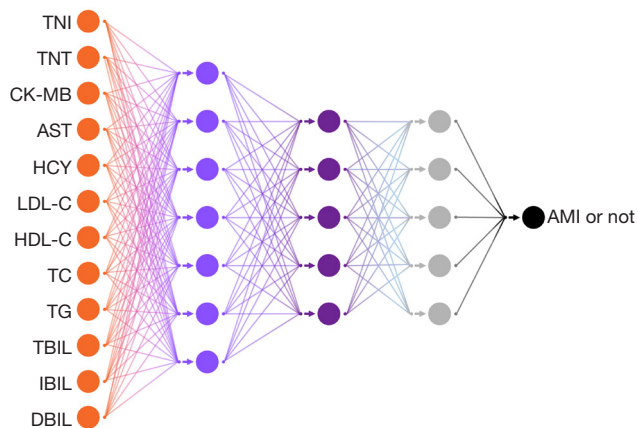


Figure 3 Architecture of fully connected neural network used as the basic model. AMI, acute myocardial infarction; TNI, troponin-I; TNT, troponin-T; CK-MB, creatine kinase isoenzymes; AST, aspartate aminotransferase; HCY, homocysteine; LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol; TC, total cholesterol; TG, triglycerides; TBIL, total bilirubin; IBIL, indirect bilirubin; DBIL, direct bilirubin.

the hospitals start model training with selected local data. Then, the network weight of the local model was added with random mask and sent to the Arya cloud service. It is worth noting that the original data were still stored in the hospital without any transmission during this process. Fifth, the Arya cloud service received the network weight and calculated the weighted mean based on the proportion of respective sample size in order to merge the local models. The example of 2-center collaborative modeling is shown as follows:

$$w_g \leftarrow \frac{n_i}{n_i + n_j} w_i + \frac{n_j}{n_i + n_j} w_j \quad [2]$$

where w_i and w_j are the network weight of the local models, n_i and n_j are the numbers of matched data in the hospitals, and w_g is the network weight of the global federation model updated in this iteration. The secure aggregation algorithm assigned random masks to each computation node to avoid leakage of w_i and w_j during the transmission, and also ensures that the weighted sum of each collaborator's random mask is 0 when calculating w_g . Sixth, the Arya cloud service sent the latest network weight of the global model to the computation nodes of the hospitals and updated their local models. Seventh, steps 4–6 were repeated k times, that is, the number of the global iteration was set to k , to get the

final trained multicenter AMI diagnostic model.

Local model settings

In FL, the computation node at each institution was asked to build a local model that learns knowledge only from internal data. We chose the fully connected neural network as a basic model. As shown in *Figure 3*, the local model consists of 1 input layer, 3 hidden layers, and 1 output layer. The input layer contains 12 neurons that receives 12 biochemical indicators as input. Information is then forwarded to the next 3 hidden layers with 7–5–5 neurons, respectively. The output layer then gives the signal of whether a patient has AMI or not. Values of each neuron in 1 layer are linear combined and activated by rectified linear unit (ReLU) in the input layer as follows:

$$ReLU(x) = \max(0, x) \quad [3]$$

and the hidden layer, or sigmoid as follows in the output layer:

$$\sigma(x) = \frac{1}{1 + e^{(-x)}} \quad [4]$$

The local model applied binary cross-entropy as the loss function, and was optimized with Adam algorithm with a learning rate of 0.05.

Global model settings

In this study, we chose $k=100$, i.e., the number of the global iteration was set to 100, to get the final trained multicenter AMI diagnostic model.

Results

Data quality

A total of 1,168 (77.9%) records were defined as high-quality data for the Second Affiliated Hospital, Zhejiang University School of Medicine, 224 (31.5%) records were defined as high-quality data for Jiande First People's Hospital, and no record was defined as high-quality data for the First People's Hospital of Linping because 2 columns of values were missing.

Based on the data quality, the dataset from the Second Affiliated Hospital, Zhejiang University School of Medicine was applicable for single-center modeling and privacy-preserving multicenter modeling. The datasets from Jiande First People's Hospital and the First People's Hospital of Linping could not be used for single-center modeling due

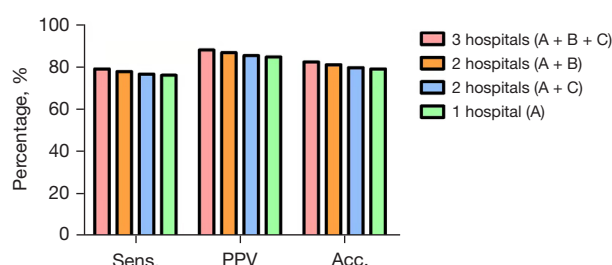


Figure 4 Modeling efficacy evaluation. (A) The Second Affiliated Hospital, Zhejiang University School of Medicine; (B) Jiande First People's Hospital; (C) the First People's Hospital of Linping. Acc, accuracy; PPV, positive predictive value; Sens, sensitivity.

to too many missing values that were difficult to repair. In the present study, we tried to use the datasets from Jiande First People's Hospital and the First People's Hospital of Linping for 2- and 3-center modeling to investigate the possibility of performance improvement.

Performance of 3-center modeling by FL

The results showed that the model trained on 3-center data achieved 79% sensitivity, 88% positive predictive value, and 82.3% accuracy for AMI prediction (Figure 4).

Performance of 2-center modeling by FL

The results showed that the model trained on 2-center data (the Second Affiliated Hospital, Zhejiang University School of Medicine and Jiande First People's Hospital) achieved 77.8% sensitivity, 86.7% positive predictive value, and 81% accuracy for AMI prediction, and the model trained on 2-center data (the Second Affiliated Hospital, Zhejiang University School of Medicine and the First People's Hospital of Linping) achieved 76.6% sensitivity, 85.3% positive predictive value, and 79.7% accuracy for AMI prediction (Figure 4).

Performance of single-center modeling

Due to the low-quality dataset from Jiande First People's Hospital and the First People's Hospital of Linping, we only used datasets from the Second Affiliated Hospital, Zhejiang University School of Medicine for single-center modeling as a comparison. The fully connected neural network build in the single-center scenario followed the same settings as the multicenter modeling. The results showed that the model

trained on single-center data achieved 76% sensitivity, 84.7% positive predictive value, and 79% accuracy for AMI prediction (Figure 4).

Our privacy-preserving FL model gave reliable AMI diagnostic efficacy. Three-center modeling achieved 79% sensitivity, 88% positive predictive value, and 82.3% accuracy. Two-center modeling achieved 77.8% or 77.6% sensitivity, 86.7% or 85.3% positive predictive value, and 81% or 79.7% accuracy. Single-center modeling achieved 76% sensitivity, 84.7% positive predictive value, and 79% accuracy.

Discussion

AMI is a leading cause of mortality, despite recent advances in percutaneous coronary intervention and pharmacotherapy. In the recent years, the incidence of AMI has been gradually increasing, especially in those under the age of 45 years (14). Therefore, early prediction and diagnosis is helpful for prevention and maintenance, and reducing AMI morbidity and mortality.

Several diagnostic models have been developed based on artificial intelligence (15-18). Du *et al.* reported that a system model based on improved machine learning showed a certain effect in their clinical analysis of AMI (17). Wang *et al.* established a multitask interactive attention learning model and found it could improve the diagnostic efficiency of AMI (18). However, privacy security risks have hindered the development of artificial intelligence-based diagnostic models in AMI and other diseases. The Arya privacy computing platform was developed by Healink is to solve the dilemma between data value and privacy protection. It is based on collaborative computing and joint modeling of multiparty data, so as to help medical institutions better apply data to carry out medical research on the basis of ensuring the security of original data.

In the current study, we explored the possibilities of medical data collaboration on National Regional Medical Center for Cardiovascular Diseases constructed by the Second Affiliated Hospital, Zhejiang University School of Medicine. We first established a FL model for AMI diagnosis with medical data from 3 individual medical institutions based on Arya. During the process, Arya was used to build a single data platform based on privacy and security computing for each hospital as the core, and then to build a data collaboration alliance network that could provide 2-way promotion and improve the service level. In addition, as a distributed machine learning paradigm, FL can effectively solve the problem of data islands, so that each participant can

only exchange encrypted model parameters and complete the establishment of the model without exchanging original data. The machine learning algorithms that can be used in FL include neural networks, which were used in the current study, logistic regression, and a tree-based model (19,20), which is suitable for complex iterative computing scenarios of big data modeling and predictive analysis. The technical characteristics of FL include the following: (I) the data of all parties shall be kept locally without disclosing privacy or violating laws and regulations; (II) under the FL system, each participant has the same identity and status; (III) the modeling effect of FL is the same as that of modeling the whole dataset in 1 place, or there is little difference; and (IV) all participants combine data to establish a virtual common model and a system for mutual benefit (21).

The results demonstrated that our privacy-preserving FL model gives reliable diagnostic efficacy for AMI. Three-center modeling and two-center modeling achieved relative high. And single-center modeling achieved relative low diagnostic efficacy. These findings indicate that multicenter medical research favors to achieve reliable diagnostic efficacy, and Arya is efficient and practical for real-world applications, and to promote multicenter medical research in a secure manner without sacrificing diagnostic efficacy.

Valuable data with scientific research value have been shelved by data managers for risk control due to the lack of privacy and security computing power of past collaboration platforms, creating further obstacles for the development of multicenter medical research. Therefore, it is important to establish a platform for big medical data use and exchange based on privacy security computing technology. The findings in the current study provide a good example for multicenter medical research based on a privacy computing platform, and achieved good results. In the future, privacy security computing technology will continue to play a leading role in medical research, especially multicenter medical research, with better data value sharing.

Acknowledgments

Funding: This work was supported by the National Natural Science Foundation of China (No. 82170332) and the Key Research and Development Program of Zhejiang Province (No. 2022C01145).

Footnote

Reporting Checklist: The authors have completed the STARD

reporting checklist. Available at <https://atm.amegroups.com/article/view/10.21037/atm-22-4331/rc>

Data Sharing Statement: Available at <https://atm.amegroups.com/article/view/10.21037/atm-22-4331/dss>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://atm.amegroups.com/article/view/10.21037/atm-22-4331/coif>). BL, GZ and DW are employed by Hangzhou Healink Technology Co., Ltd. One of their jobs is to provide technical support of Arya platform and assist data processing and analysis. They promise to avoid conflicts of interest with the company, its shareholders and its customers. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the ethics committee of the Second Affiliated Hospital, Zhejiang University School of Medicine (No. 2022-0695). The other 2 hospitals (Jiande First People's Hospital, the First People's Hospital of Linping) were also informed and agreed the study. And written informed consent was waived, as the data were analyzed retrospectively in a privacy-preserving manner.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Son Y, Han K, Lee YS, et al. Privacy-preserving breast cancer recurrence prediction based on homomorphic encryption and secure two party computation. *PLoS One* 2021;16:e0260681.
2. Yuan J, Malin B, Modave F, et al. Towards a privacy preserving cohort discovery framework for clinical research networks. *J Biomed Inform* 2017;66:42-51.

3. Lu WJ, Yamada Y, Sakuma J. Privacy-preserving genome-wide association studies on cloud environment using fully homomorphic encryption. *BMC Med Inform Decis Mak* 2015;15 Suppl 5:S1.
4. Kim D, Son Y, Kim D, et al. Privacy-preserving approximate GWAS computation based on homomorphic encryption. *BMC Med Genomics* 2020;13:77.
5. Carpov S, Nguyen TH, Sirdey R, et al. Practical privacy-preserving medical diagnosis using homomorphic encryption. In: *IEEE 9th International Conference on Cloud Compu-ting (CLOUD). 2016 IEEE 9th International Conference on Cloud Computing (CLOUD)*, 2016; 593-9.
6. Kocabas O, Soyata T. Utilizing homomorphic encryption to implement secure and pri-vate medical cloud computing. In: *IEEE 8th International Conference on Cloud Computing. Proceedings of the 2015 IEEE 8th International Conference on Cloud Computing*, 2015; 540-7.
7. Kocabaş Ö, Soyata T. Medical data analytics in the cloud using homomorphic encryp-tion. In: *E-Health and Telemedicine: Concepts, Methodologies, Tools, and Applications*. IGI Global, 2016; 751-68.
8. Zhou J, Feng Y, Wang Z, et al. Using Secure Multi-Party Computation to Protect Privacy on a Permissioned Blockchain. *Sensors (Basel)* 2021;21:1540.
9. Dong X, Randolph DA, Weng C, et al. Developing High Performance Secure Multi-Party Computation Protocols in Healthcare: A Case Study of Patient Risk Stratification. *AMIA Jt Summits Transl Sci Proc* 2021;2021:200-9.
10. Simm J, Humbeck L, Zalewski A, et al. Splitting chemical structure data sets for federated privacy-preserving machine learning. *J Cheminform* 2021;13:96.
11. Wang R, Lai J, Zhang Z, et al. Privacy-Preserving Federated Learning for Internet of Medical Things under Edge Computing. *IEEE J Biomed Health Inform* 2022. [Epub ahead of print]. doi: 10.1109/JBHI.2022.3157725.
12. Hasselgren A, Kralevska K, Gligoroski D, et al. Blockchain in healthcare and health sciences-A scoping review. *Int J Med Inform* 2020;134:104040.
13. Chen HS, Jarrell JT, Carpenter KA, et al. Blockchain in Healthcare: A Pa-tient-Centered Model. *Biomed J Sci Tech Res* 2019;20:15017-22.
14. Eggers KM, Ellenius J, Dellborg M, et al. Artificial neural network algorithms for early diagnosis of acute myocardial infarction and prediction of infarct size in chest pain patients. *Int J Cardiol* 2007;114:366-74.
15. Liu WC, Lin CS, Tsai CS, et al. A deep learning algorithm for detecting acute myocardial infarction. *EuroIntervention* 2021;17:765-73.
16. Wang S, Li J, Sun L, et al. Application of machine learning to predict the occurrence of arrhythmia after acute myocardial infarction. *BMC Med Inform Decis Mak* 2021;21:301.
17. Du H, Feng L, Xu Y, et al. Clinical Influencing Factors of Acute Myocardial Infarction Based on Improved Machine Learning. *J Healthc Eng* 2021;2021:5569039.
18. Wang Q, Zhao C, Qiang Y, et al. Multitask Interactive Attention Learning Model Based on Hand Images for Assisting Chinese Medicine in Predicting Myocardial Infarction. *Comput Math Methods Med* 2021;2021:6046184.
19. Katzman JL, Shaham U, Cloninger A, et al. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC Med Res Methodol* 2018;18:24.
20. Ching T, Zhu X, Garmire LX. Cox-nnet: An artificial neural network method for prognosis prediction of high-throughput omics data. *PLoS Comput Biol* 2018;14:e1006076.
21. Peyvandi A, Majidi B, Peyvandi S, et al. Privacy-preserving federated learning for scalable and high data quality computational-intelligence-as-a-service in Society 5.0. *Multimed Tools Appl* 2022;81:25029-50.

(English Language Editor: R. Scott)

Cite this article as: Xu J, Zhang Y, Yu H, Lin B, Xiang P, Lin T, Lu H, Zhang G, Wang D, Yuan H, Hu B, Jiang J. High performance of privacy-preserving acute myocardial infarction auxiliary diagnosis based on federated learning: a multicenter retrospective study. *Ann Transl Med* 2022;10(18):1006. doi: 10.21037/atm-22-4331