Federated Learning for Healthcare Domain - Pipeline, Applications and Challenges

MADHURA JOSHI*, ANKIT PAL*, and MALAIKANNAN SANKARASUBBU,
Saama AI Research, India

Federated learning is the process of developing machine learning models over datasets distributed across data centers such as hospitals, clinical research labs, and mobile devices while preventing data leakage. This survey examines previous research and studies on federated learning in the healthcare sector across a range of use cases and applications. Our survey shows what challenges, methods, and applications a practitioner should be aware of in the topic of federated learning. This paper aims to lay out existing research and list the possibilities of federated learning for healthcare industries.

CCS Concepts: • **Security and privacy**; • **Computing methodologies** → **Artificial intelligence**;;

Additional Key Words and Phrases: Federated learning, GDPR, transfer learning

ACM Reference format:
Madhura Joshi, Ankit Pal, and Malaikannan Sankarasubbu. 2022. Federated Learning for Healthcare Domain - Pipeline, Applications and Challenges. *ACM Trans. Comput. Healthcare* 3, 4, Article 40 (October 2022), 36 pages.
https://doi.org/10.1145/3533708

40

1 INTRODUCTION

In the last few years, digital healthcare data has grown significantly. At the same time, recent breakthroughs in **deep learning (DL)** have been used in a variety of current medical data processes, including automatic disease diagnosis [72, 91], classification, biomedical data analysis, Question Answering in the medical domain [83], and segmentation [67, 76]. These methods hold immense promise and innovation in this field. In the coming future, the advancement of these methods will refine health care systems and improve medical practices worldwide.

Diagnostic tools, **machine learning (ML)** based healthcare solutions, and models must be exposed to a wide variety of cases and data that cover a full range of possible anatomies to capture more informative patterns in the medical data. It is well known that data from a single source can be significantly biased by the equipment, demographics, and acquisition protocol. Therefore, training a model on data from a single source would skew its prediction performance towards the population. Moreover, it is computationally expensive and time-consuming.

Training models in a parallelized manner [30] and within small batches can mitigate a few of these challenges. However, though this approach addresses computation challenges, it does not necessarily preserve the privacy of data. Clinical research often involves studies from a large amount of data collected from various sources. Health institutions, individuals, insurance companies, and the pharmaceutical industry all have access to medical data.

*Madhura Joshi and Ankit Pal equally contributed to this research.

Authors' address: A. K. Pal, Saama AI Research Lab A-3,4, Olympia National Tower, Smartworks, Guindy Industrial Estate, Guindy, Chennai, Tamil Nadu, India 600032; email: m.joshi@saama.com; M. Joshi, Saama AI Research Lab A-3,4, Olympia National Tower, Smartworks, Guindy Industrial Estate, Guindy, Chennai, Tamil Nadu, India 600032; email: ankit.pal@saama.com; M. Sankarasubbu, Saama AI Research Lab, 900 East Hamilton Avenue, Suite 200 Campbell, CA 95008; email: malaikannan.sankarasubbu@saama.com.

2637-8051/2022/10-ART40
https://doi.org/10.1145/3533708

Furthermore, each institution may be linked to a unique collection of stakeholders. These data are often sensitive and cannot be aggregated or accessed.

Access to a large amount of high-quality medical data is possibly the most crucial factor for enhancing Machine Learning (ML) applications in the healthcare domain. However, security and privacy issues of healthcare data have raised broad ethical and legal concerns in recent years, given the sensitive nature of health information [1].

The assembly and transmission of these datasets are ethically and legally required to protect patient privacy. Most healthcare centers, laws at the country level, and regulatory bodies, e.g., the **General Data Protection Regulation (GDPR)** and the **Health Insurance Portability and Accountability Act (HIPAA)**, have passed new laws that control sharing data while preserving user security and privacy [69, 125]. Moreover, information and control about medical data storage, transmission, and usage are central to patient rights.

The sensitive and distributed nature of **EHR (Electronic Health Records)** in real-world scenarios simulates a need for an effective mechanism to learn from data residing in health-related institutions, hospitals while accounting for data privacy. This motivates us to examine the potential and value of federated learning for the healthcare domain. Federated Learning is an advanced distributed learning technique that leverages datasets from various universities without explicitly centralizing or sharing the training data [66, 124].

Federated learning provides many advantages as compared to centralized learning. It enables training a global model from distributed data. This method also focuses on preserving data privacy by only sharing mathematical parameters and metadata while keeping the actual data as secure as possible and preventing attacks and tracebacks.

The global ML model is distributed to each client site, where an instance is trained locally. The updates from locally trained instances are then aggregated at regular intervals to improve the global model. The updated global model is then sent back to the local devices, where the learning continues. These steps are repeated until a particular convergence threshold is satisfied or lasts for a long time to improve the deep learning model continuously.

The parameters and metadata sharing depend on many factors such as use cases, data management and regulation, business agreement and protection, pipeline, and infrastructure. In this article, we provide an overview of the federated learning approach in the healthcare domain. Our contribution is as follows:

- We demonstrate the components of the federated learning setup and discuss the communication architecture and building blocks of a federated learning system.
- We examine the various challenges that a federated learning setup faces in terms of privacy, data, and communication in the healthcare system.
- We survey existing works on federated learning in the health sector and propose a comprehensive list of applications classified into prognosis, diagnosis, and clinical workflow.

The structure of the paper is as follows: the first section describes the components of a federated learning setup as well as a federated learning pipeline. Section 2 discusses the challenges and concerns of federated learning in the healthcare area. Section 3 concludes with a study of federated learning applications in the health care area and the tools required for implementation.

2 PIPELINE

2.1 Architecture of Communication

2.1.1 Centralized. Centralized federated learning architecture is the most commonly used architecture. In this setting, data flows in an asymmetric nature. A central server is responsible for aggregating the information (e.g., local models) from the data owners, coordinating with all the participating client devices, and sending back the model updates. Client devices usually communicate only with the server; hence the server acts as a system bottleneck. However, Single-point failure is one of the drawbacks in a centralized setting.

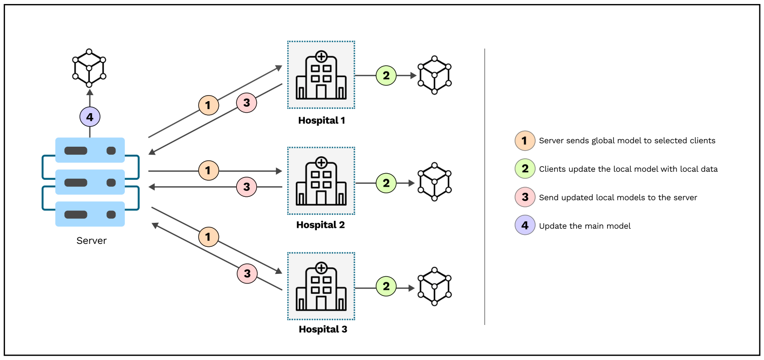

Fig. 1. Illustration of centralized architecture in federated learning.

The following is the architecture of centralized federated learning. First, a global model is sent to client devices. Second, the global model is updated with client data. Third, the revised model is sent back to the server. Finally, local models are averaged on the server to create a new global model. Furthermore, this process is repeated. The architecture of centralized federated learning is shown in Figure 1.

Google uses the centralized architecture design in their Android keyboard setting. The central server collects the local model information from users' mobile devices and updates the global model, and later the updates are sent back to the users for inferences. FedAvg is one such example of a centralized framework [75].

2.1.2 Decentralized. In the decentralized, federated learning architecture, the client devices can communicate to train a global model and update it directly without any central server. A decentralized setup prevents the single point failure issue. SimFL is a federated, decentralized framework [66].

The following is the architecture of decentralized, federated learning. First, the local gradients are updated. Second, these gradients are sent to selected parties. Third, the model is updated with local data and gradients. Lastly, the updated model is sent to other parties, and the process continues. The architecture of decentralized, federated learning is shown in Figure 2. This method is usually preferred because it involves the exchanging of local models which are aggregations of a large quantity of data. However, decentralized federated learning works on the concept of mutual trust between the users which leads to a few drawbacks. The question arises when one has to use DFL in a single sided trust environment. A decentralized learning algorithm called **Online Push Sum (OPS)** [46] addresses this challenge. It involves a rigorous regret analysis, tested and compared with other algorithms like **Decentralized Online Gradient (DOG)** and **Centralized Online Gradient (COG)**.

Blockchain is an excellent example of a decentralized platform [131]. Another example of a similar design is the decentralized cancer diagnostic system amongst hospitals/medical institutions. Each hospital distributes the local model that has been trained with patient data and acquires the global model for future diagnosis [97].

2.2 Scale of Federated Learning system

2.2.1 Cross-Silo. In a Cross-Silo federated learning setting, small numbers of organizations or data centers (e.g., medical or financial) train a global model with a large amount of data along with computation power. In an environment where organizations are prohibited from sharing their data due to data privacy and security reasons, cross-silo federated learning comes into the picture. The computation and storage capacity of the cross-silo setting is high on a relatively small scale, while the stability is low compared to cross-device.

2.2.2 Cross-Device. In a Cross-Device federated learning setting, there are a scalable number of clients with small amounts of data compared to a cross-silo setting where there are only a few clients with large amounts of data. As a result, the cross-device system develops models for large-scale dispersed data inside the same application [73]; the clients can be enabled and disabled as per the requirement and are usually mobile and IoT

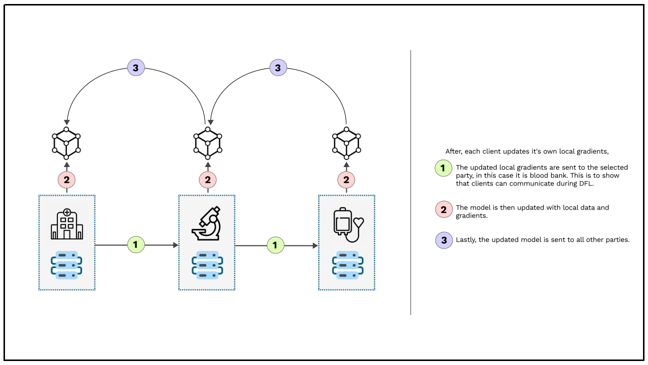

Fig. 2. Illustration of decentralized architecture in federated learning.

devices. The system should be powerful enough to manage many devices and handle common failures such as unstable connections, power failure, etc.

2.3 Federated Server

The central server serves as a manager in a cross-device environment. It oversees the communication between the client devices, the server, and the global machine learning model. At the beginning of the federated learning training process, the client starts a server training service; a login request is required from each client-side to join the server training session.

The server will check the credentials to validate the request and allow the clients to join the session only in a successful authentication. The server session decides the minimum and a maximum number of client devices that need to be added to start the training process.

In a decentralized environment, all devices interact directly with one another and participate equally in global model training. The server, in this case, is all of the local devices. Practical Federated Gradient Boosting Decision Trees [68] is an example of such a framework, in which each device trains decision trees sequentially, and the final model is the sum of all the trees. Designing a completely decentralized FLS with appropriate communication overhead is difficult.

Each local model training happens on the client's side, so the server does not need access to the model training data. Moreover, local models only share the model updates instead of the actual copy of data. The client's devices decide the number of epochs to run during each round of training. Typically, the computing is used for model training and aggregation, while the communication is used to exchange model parameters.

Once the server receives all the model updates from all the participating devices, it performs the model aggregation based on the selected aggregation algorithm and gets the overall updated global model. This process completes the single round of the Federated Learning session. Furthermore, the next round starts and training continues until the maximum number of rounds on the server is completed. The aggregation technique aids in the gathering of model updates from local models of the devices. Listed below are some popular application-oriented aggregation algorithms in detail.

- **FedAvg Algorithm.** FedAvg is the first and most widely used Federated learning algorithm proposed by Google [75]. It distributes training data across mobile devices and trains a shared model by combining locally calculated updates. The FedAvg algorithm is a robust approach experimented on five different architectures, four kinds of datasets, and combined stochastic gradient descent on each client with a server that performs model averaging. The algorithm can be written as

$$f(w) = \sum_{k=1}^{K} \frac{n_k}{n} F_k(w) \quad \text{where} \quad F_k(w) = \frac{1}{n_k} \sum_{i \in \mathcal{P}_k} f_i(w) \tag{1}$$

Where $\mathcal{P}_k$ the set of indexes of data points on client k, distributed uniformly at random, with $n_k = |Pk|$ and w are model parameters. The focus of the proposed research is on the optimization's non-IID data distributions and imbalanced properties, as well as the nature of the communication constraints.

- **FedMa Algorithm.** FedMA [113] is a federated learning method for current neural network designs such as **Convolutional Neural Networks (CNNs)** and LSTMs. FedMA builds the shared global model layer by layer by matching and averaging hidden components with comparable feature extraction signatures (i.e., channels for convolution layers; hidden states for LSTM; neurons for fully connected layers).
- **FedNAS Algorithm.** One of the major issues on working with Non-I.I.D data is that it requires the developers to design and choose multiple architectures that tunes the hyper parameters and fits the scattered data. This design process is expensive due to a large number of communication rounds and computational burden. FedNAS - Federated Neural Architecture Search [43] is an algorithm that improves the model accuracy by automating the manual design process. The algorithm helps scattered workers to collaboratively search for a better architecture.
- **FedGKT Algorithm.** The fact that federated learning ensures privacy and confidentiality is one of its most appealing features. There is, however, a limit to the computation capabilities of the edge nodes. Convolutional neural networks are used to boost the model's performance and accuracy, however when employed with FL restrictions, large models might burden edge nodes and increase communication costs. A group knowledge transfer training technique intends to reduce the need for edge compute while keeping edge training inexpensive by lowering communication bandwidths for training with CNNs at the edge. FedGKT [44] requires 9 to 17 times fewer computing resources on edge devices than FedAvg. In comparison to FedAvg, the authors claim that this algorithm can achieve the same or slightly higher accuracy.
- **Secured Weighted Aggregation.** The proposed algorithm [41] relies on homomorphic encryption - cryptosystem for calculating the client's weights in a privacy-preserving manner. The algorithm adopts a **Zero-Knowledge Proof (ZKP)** based verification scheme to prevent the central servers and clients from receiving fraudulent messages from each other. This is the first aggregation algorithm that deals with data disparity and fraudulent messages
- **Secure Federated matrix factorization.** The authors of this research offer a Federated Matrix Factorization Framework to address the problem of disclosing data through gradients FedMF [17]. Distributed machine learning and homomorphic encryption techniques are used in the system. FedMF is a user-level distributed matrix factorization framework that allows the model to be learned by uploading the gradient information to the server rather than the raw preference information. The authors utilize homomorphic encryption to improve distributed matrix factorization and boost security since gradient information can leak user information to some extent.
- **Inprivate Digging.** The authors concentrate on a tree-based data mining concept and provide privacy-preserving approaches for two of the most popular tasks: Regression and Binary Classification, in which individual data owners can train locally in a differentially private way [130]. They created and implemented a privacy-preserving **Gradient Boosting Decision Tree (GBDT)** system that allows several regression trees trained by distinct data owners to be safely aggregated into an ensemble without the need of a third party.
- **Federated Forest.** The authors provided a system that allows random forests to be trained in a vertical FL scenario [72]. Throughout the development of each node, the party with the related split feature is in charge of splitting the samples and exchanging the results. To safeguard privacy, they encrypt the data that is sent. Their method is equally accurate as of the non-federated one. Both classification and regression tasks are supported by the model.

2.3.1 Model Evaluation. In traditional centralized machine learning, evaluation metrics are used to assess the performance and measure the quality of the model. Evaluation metrics mainly include accuracy, precision, and recall, etc. Accuracy is a fraction of the correct samples to all samples. Precision is the fraction of actual positive samples among the positive samples, while recall is the fraction of actual positive samples among the samples from true positive or false negative.

In a federated learning setup, the above metrics are insufficient to evaluate the performance of the models. In this setting, it is intended to identify the evaluation metrics for both qualitative and quantitative methods. Training speed, performance, and the quantity of data transferred are utilized as assessment measures in existing federated learning studies. An FL model is evaluated by examining the aggregated model after being assigned to each clients' local evaluation dataset. Following then, the server shares each client's performance with it, aggregating the local results to produce global assessment metrics.

FedEval [18] is an evaluation framework for FL systems. It introduces the "ACTPR" model, i.e., using accuracy, communication, time consumption, privacy, and robustness as its evaluation targets.

2.4 Federated Client

In a federated learning setting, the client devices are the hardware components that perform the local model training and loss minimization. The main objective is to train local device models on their private data and share the updates with the federated server where the aggregation is performed. In a decentralized federated learning system, the client devices communicate among themselves without the interference of the central server. Typical clients in federated learning settings could be smartphones, IoT devices, or organizations such as hospitals. Each client has its specific private training dataset and its local model.

A study in [73] describes the client management patterns that control the local devices' information and their connection with the central server. The client registry is the initial component, and it controls the information about the participating devices. The second component, the client selector, selects the devices for the federated training task. To improve the training model's performance and efficiency, the third component, client cluster, clusters client devices depending on particular parameters (e.g., data distribution, available resources).

2.4.1 Data Partitioning. Federated learning systems are commonly classified as horizontal, vertical, or hybrid based on how data is spread through the sample and feature spaces.

- **Horizontal Federated Learning**
  - **Concept.** Horizontal federated learning involves data that is shared horizontally which involves datasets sharing the same feature space but being different in samples [124]. Each group in horizontal FL has access to the entire feature set and labels, allowing them to train their local model using their dataset. After that, all of the parties exchange their model updates with an aggregator and then generate a global model by combining, for example, the model weights obtained from different parties. [120] An example of horizontal data split in federated learning setup is shown in Figure 3. Assume there is a blood bank laboratory that has details of the patients stored as groups in a database collected under Name, Age, and Blood Person features. In Figure 3, persons 1, 2, and 3 represent groups of patients' data samples sharing the same features - Name, Age and Blood Group. Suppose a Horizontal Federated learning method is to be used. In that case, the machine learning model will run on each of these samples (person) individually, allowing them to train the local model and finally exchange and aggregate model weights to generate a global model.
  - **Applications.** The FedAvg algorithm proposed in [101] is the best example of a typical horizontal federated learning setup. Although horizontal federated learning protects user privacy and aims to control the communication cost, it poses challenges compared to distributed learning. Keeping in mind the protocols of federated learning and applications of neural networks, a new algorithm titled "Federated neural network" was proposed for single and multi-objective search purposes [133]. In comparison to Secure Multiparty Computation and Homomorphic Encryption, this study also analyzes how Differential Privacy may be employed best in horizontal federated learning.

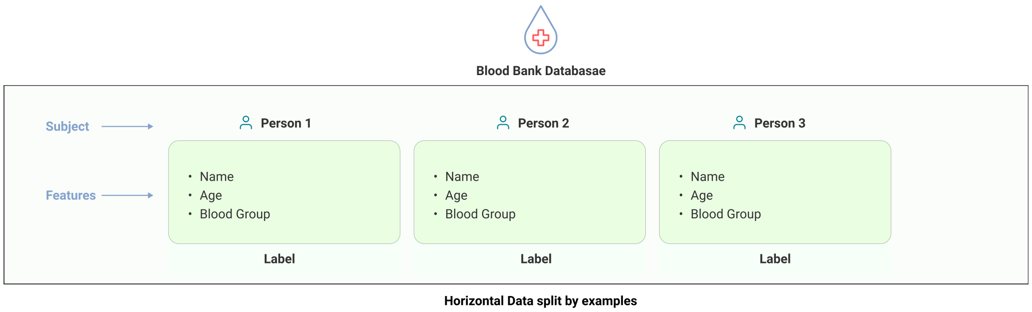

Fig. 3. Horizontal Data split by examples.

EHRs have proven to be an aid for the development of scientific evidence for improving the quality of healthcare systems and they are also used as a data resource for laboratory-based public health surveillance. Techniques have been created for computing statistics on distributed databases without disclosing any personal details other than the statistical results. In a distributed dataset, duplicate records can lead to incorrect statistical findings. As a result, safe deduplication is an effective preprocessing step for improving the precision of statistical analysis of a distributed dataset. A stable protocol uses a deterministic record linking algorithm for deduplication of horizontally partitioned datasets [128]. They introduced a new robust and scalable protocol for privacy-preserving deduplication of a horizontally partitioned dataset based on Bloom filters. According to the results of the experiments, within 45 seconds, one million virtual documents were deduplicated across 20 data custodians. The protocol is proven to be more effective and flexible than previous protocols for the same problem.

- **Vertical Federated Learning**
  - **Concept.** Vertical federated learning can also be called feature-based federated learning. This is used when the data shared among them contains different features but similar samples. This requires a different kind of training architecture when compared to horizontal federated learning. It may or may not involve a central server or a third neutral party. **Vertical FL (VFL)** applies to collective situations in which individual parties do not have access to the whole collection of features and labels and thus are unable to train a model locally using their datasets. Parties' datasets, in particular, must be synchronized to construct the full function vector without revealing their respective training data, and model training must be performed in a privacy-preserving manner [120]. An example of vertical data split by examples is shown in Figure 4. Assume a blood bank laboratory in a hospital has data from the same group of patients. As seen in Figure 4, the blood bank database stores the details of the patient groups under Name, Age, and Blood group. At the same time, the hospital database stores the details of the same patient groups under Name, Age, Date of Birth, Blood group, and Medical History. If a Vertical Federated learning method is to be used, the first dataset for the machine learning model would be the data sample from Person 1 of the Blood Bank Database and Hospital Database. Furthermore, the second dataset will be data samples from Person 2 of the Blood Bank database and Hospital Database, respectively.
  - **Applications.** FedAI uses horizontal federated learning for improving an anti-money laundering model and the use of vertical federated learning to obtain a better risk management model. Homomorphic Encryption is the privacy-preserving technique that is usually adopted when the data is distributed vertically. Recently, a two-party design has been proposed by eliminating the trusted coordinator, which considerably decreases the system's complexity [125].

In the healthcare sector, information sharing is becoming increasingly relevant. First and foremost for medical reasons, such as the sharing of treatment records between healthcare providers, but also for secondary purposes, such as the implementation of value-based healthcare and processes. While these

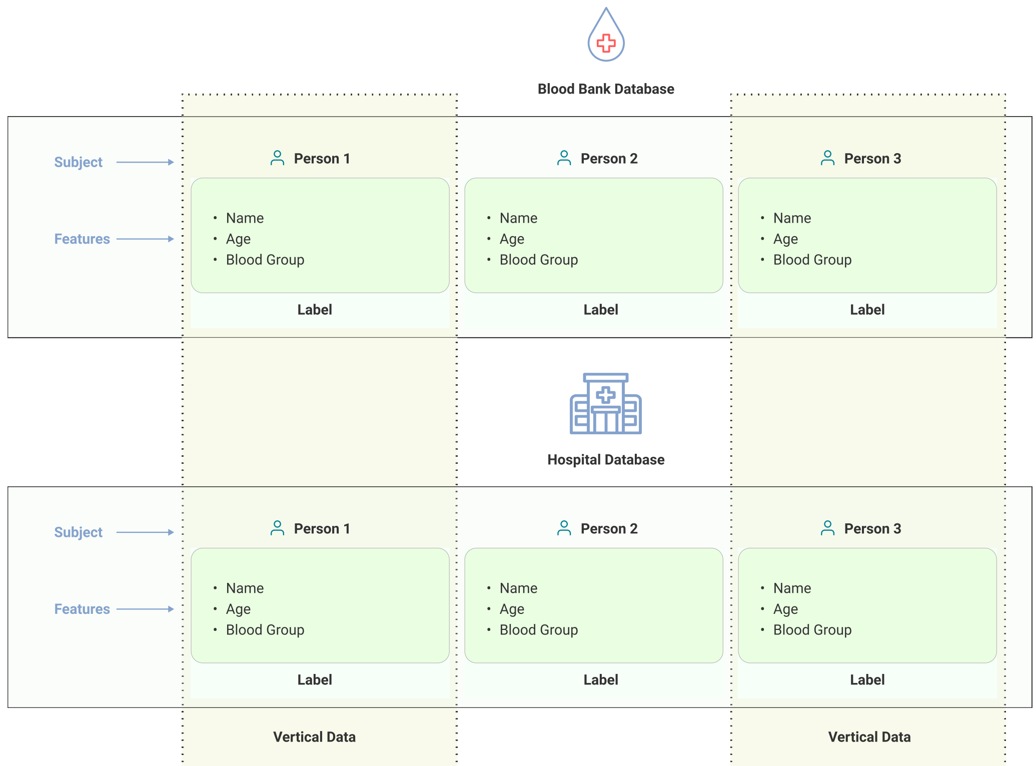

Fig. 4. Vertical Data split by examples.

requirements allow information to be transferred, they also pose concerns regarding maintainability and possession, as well as protection and privacy. Provenance and permission become more complicated as data is shared by various health care providers. Sending programs containing queries and algorithms to the data source is one of the alternatives to data transfer. To address these problems, the authors in [103] provide an architecture to enable algorithm conversion and execution, and use this infrastructure in a proof-of-concept setup. The **proof-of-concept (PoC)** focuses on processing data that has been vertically partitioned from two institutes. In a population cohort analysis, the PoC is used as a baseline to look at the causes of diabetes initiation and development, including social and environmental conditions.

Since full collections of labels and functions are not controlled by one person, privacy-preserving vertical FL is difficult. Existing vertical FL approaches necessitate multiple peer-to-peer interactions among parties, resulting in longer training periods, and are limited to (roughly) linear models and only two parties. To bridge this void, the authors of [120] suggest FedV, a framework for safe gradient computing in vertical settings for a variety of commonly used machine learning models. Linear simulations, logistic regression, and support vectors are examples of these types of models. FedV uses functional encryption mechanisms to eliminate the need for peer-to-peer correspondence between parties. FedV can attain shorter preparation times as a result of this. It's also successful for wider and ever-changing groups of parties.

- **Hybrid/Transfer Federated Learning**
  - **Concept.** Federated transfer learning [124] was introduced to tackle the challenge of integrating the scattered data and improve statistical modeling while performing data federation. It is also studied that this framework requires minimal modifications to the model structure and the results produced are as

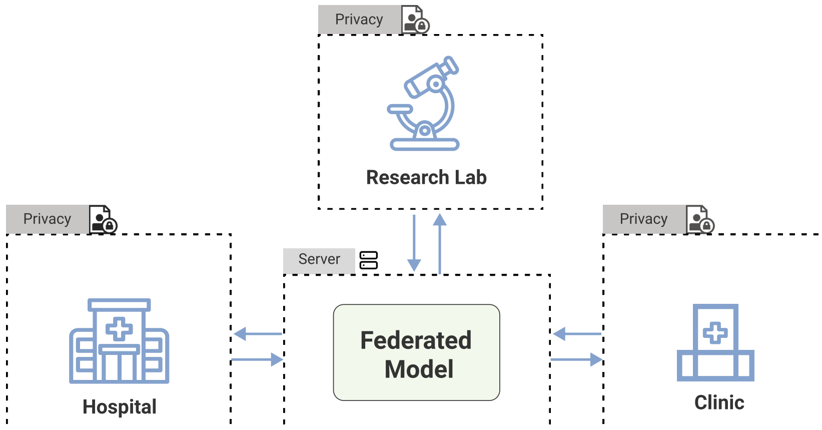

Fig. 5. Transfer learning in federated learning.

efficient as the non-privacy preserving transfer learning. Federated transfer learning does not depend on any requirement such as common feature space or common sample space and it supports transfer learning in providing solutions for the entire sample and feature space while data federation is taking place.

Federated transfer learning is a term that recognizes difficult situations in which data parties only have partial overlap in the user or function space and uses current transfer learning strategies to collaboratively construct models. The current formulation is only good for two clients [58]. An illustration of hybrid/transfer federated learning is shown in Figure 5.

– **Applications.** FedHealth is one such algorithm that uses the concept of federated transfer learning in smart wearable healthcare devices [57]. The algorithm performs data aggregation through federated learning and builds personalized models by transfer learning without compromising on the privacy and security of the model and data.

  The shortage of massive datasets has hampered the progress of deep learning (DL) approaches in the field of **Brain-Computer Interfaces (BCI)** for the classification of **electroencephalographic (EEG)** recordings [57]. The ability to create a large EEG BCI dataset by combining several small ones for jointly training machine learning models is limited due to privacy issues associated with EEG signals. Addressing this issue is a novel privacy-preserving deep learning architecture centered on the federated learning system, for EEG classification called Federated Transfer Learning (FTL) [57]. As a result, in a subject-adaptive study, the FTL method had a 2% better classification accuracy.

  Although current **federated learning systems (FLSs)** primarily focus on one type of partition, data partitioning among parties in many other applications can be a combination of horizontal and vertical partitioning. As an example, consider a cancer detection method. A consortium of hospitals needs to build a FLS for cancer diagnosis, but each hospital has separate patients and medical test outcomes. In such cases, transfer learning could be an option. [101] suggest a stable federated transfer learning method that can learn a representation of a party's features using similar instances. This framework involves only minor changes to the original model layout and achieves the same degree of consistency as non-privacy-preserving transfer learning. It is adaptable and can be used for a variety of stable multi-party machine learning activities.

2.4.2 Data Pre-processing and Feature Engineering. Data Preprocessing is an essential and challenging task in the FL training process due to the sensitive and distributed nature of the data. Moreover, it takes a long time to adapt centralized preprocessing approaches to federated data. Data preprocessing involves data cleaning, missing values imputation, and other steps to maintain consistency between distributed datasets of different clients without knowing the underlying distribution.

A few scenarios require domain knowledge to clean, structure the data, and extract data features (properties, characteristics, attributes) from raw data. [57] used medical knowledge and AI techniques to clean and normalize the raw patient data collected from hospitals. The data from the hospital includes outpatient/in-patient prescriptions, outpatient/in-patient EHRs, and other metadata.

2.4.3 Privacy Preservation in FL. Federated learning involves the participation of thousands or millions of devices [126] such as phones, cars, medical institutions, etc. The federated learning environment, where the model is learned locally without disclosing to any clients the input data or the output of the model, avoids direct leakage when training or using the model, assuring the client's (health systems from medical institutions in this case) dataset is kept as private as possible. This is demonstrated by a model developed using patient Electronic Health Records (EHR) and a consumer-based application such as screening atrial fibrillation with electrocardiograms obtained by smartwatches [85]. Sensitive patient data is kept in local institutions or with individual consumers rather than being exposed to the federated model learning process, ensuring the patient's privacy. The 'No peek' rule [111] refers to techniques of distributed deep learning models that do not look at raw data once it leaves the clients. A good example explaining this would be hospitals or healthcare systems at institutions that are not allowed to share data with for-profit entities due to trust issues. Such institutions are also restricted from sharing it with outside entities due to the consent of patients and various regulations like HIPAA [111]. Listed below are a few methods of privacy-preserving in FL.

Secure Privacy-preserving learning on medical data mainly involves protecting the collection of personal data used in ML model training and inference in such a way that it does not reveal any additional information about the subjects. Cryptographic techniques and differential privacy techniques are the most extensively utilized methods for privacy-preserving ML [4].

- **Cryptographic methods** are used when encrypted data is required during testing and training phases. The widely used method under this is Homomorphic encryption and Secret sharing. Homomorphic encryption involves the encryption of data using ciphertext and public keys. It enables computations like addition and multiplication, which is a base for other complex functions, on encrypted data.
- **Secret Sharing.** The technique of transmitting secrets to various parties is known as secret sharing while keeping a "share" of the secret. Only when all individual shares are merged will the secret be reconstructed. The secret is reconstructed in some configurations that won't need all shares to be merged. A privacy-preserving system for emotion detection is introduced in [48]. The authors used a multi-secret sharing method to transfer audio-visual data gathered from users to the cloud using edge devices where a CNN and sparse auto-encoder were applied for the extraction of features. The **vector machine (SVM)** was used for emotion and support identification.
- **Differential privacy. Differential Privacy (DP)** is another technique to protect the privacy of individual data which has been used in areas that use algorithms like boosting [34], principal component analysis [19], and support vector machines. Differential privacy involves the addition of noise to model updates as they communicate with servers and clients [33]. Differential privacy is known to shield users from data leakage to a limited level and may lower prediction accuracy performance [22]. Hence, many researchers combine DP with the Secured Multiparty computations technique to decrease the growth of noise injection and data poisoning attacks [5]. While there is an exchange in convergence rates at the participant level - Differential Privacy (DP) may not help much in the protection of users and also may affect the accuracy of the model. Participant level DP is good when a large pool of devices are participating in the process. The ability of participant-level DP for a small pool of participants is yet to be improved and focused on working for a better convergence at this small level.

Table 1. Summary of Existing Studies on FL in Healthcare Since 2015

Title	Ref.	Year	Category	Content
A Federated Network for Translational Cancer Research Using Clinical Data and Biospecimens	[52]	2015	Learning Systems	This report describes a fully functional federated data and biospecimen sharing network for cross-institutional cancer research collaboration
Privacy-preserving GWAS analysis on federated genomic datasets	[27]	2015	Framework	On federated genomic datasets, this research proposes a privacy-preserving GWAS methodology
Privacy-Preserving Integration of Medical Data	[78]	2017	Protocol	This work presents a safe and privacy-preserving method for searching and integrating health care data from diverse sources
LoAdaBoost: Loss-based AdaBoost federated machine learning with reduced computational complexity on IID and non-IID intensive care data	[50]	2018	Learning Systems	LoAdaBoost, a methodology for increasing the efficiency of federated machine learning, was suggested in this research, and the algorithm was evaluated using data from intensive care units in hospitals
Federated learning of predictive models from federated electronic health records	[13]	2018	Framework	A novel FL framework is presented that can train predictive models through peer-to-peer cooperation instead of exchanging raw EHR data
FADL: Federated-Autonomous Deep Learning for Distributed Electronic Health Record	[70]	2018	Learning Systems	By presenting a novel approach called Federated-Autonomous Deep Learning, this study illustrates the efficacy of FL by using ICU data from 58 different hospitals to predict patient mortality can be trained quickly without transferring health data out of their silos under FL environment (FADL)
Patient Clustering Improves Efficiency of Federated Machine Learning to predict mortality and hospital stay time using distributed Electronic Medical Records	[49]	2019	FL in Biomedical	The community-based federated machine learning (CBFL) technique is described in this research, and it is tested on non-IID ICU EMRs
FedHealth: A Federated Transfer Learning Framework for Wearable Healthcare	[20]	2019	FL in Healthcare IoT	To address data privacy concerns, this paper proposes a federated transfer learning system for wearable healthcare
Communication-Efficient Federated Deep Learning with Asynchronous Model Update and Temporally Weighted Aggregation	[20]	2019	Learning Systems	This paper presents a synchronous learning strategy for FL clients
Federated deep learning for detecting COVID-19 lung abnormalities in CT: A privacy-preserving multinational validation study	[32]	2019	Diagnosis	With external validation on patients from a global cohort, this report reveals the efficiency of an FL system for identifying COVID-19 associated CT anomalies
Federated Learning for Healthcare Informatics	[119]	2019	Survey	This survey study provides an overview of federated learning systems, focusing on biomedical applications
Federated electronic health records research technology to support clinical trial protocol optimization: Evidence from EHR4CR and the InSite platform	[25]	2019	FL in EHR data	This paper determines if inclusion/exclusion (I/E) criteria of clinical trial protocols can be represented as structured queries along with those executed using a secure federated research platform (InSite) on hospital electronic health records (EHR)
Predicting Adverse Drug Reactions on Distributed Health Data using Federated Learning	[24]	2019	Framework	To increase the global model's predictive power, this research proposes two unique approaches to local model aggregation

(Continued)

Table 1. Continued

Title	Ref.	Year	Category	Content
Privacy-preserving Federated Brain Tumour Segmentation	[68]	2019	FL in Biomedical	Adopting the BraTS dataset for brain tumor segmentation, this research investigates the possibility of using differential-privacy approaches to secure patient data in a federated learning context
Multi-site fMRI Analysis Using Privacy-preserving Federated Learning and Domain Adaptation: ABIDE Results	[69]	2020	Medical Image Analysis	This work proposes a privacy-preserving multi-site fMRI classification that ensures that private information cannot be retrieved from model gradients or weights
Stochastic Channel-Based Federated Learning With Neural Network Pruning for Medical Data Privacy Preservation: Model Development and Experimental Validation	[99]	2020	Learning Systems	For the study of distributed medical data, this research proposes a privacy-preserving approach called stochastic channel-based federated learning (SCBFL)
Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data	[100]	2020	Medicine	This research shows that utilizing data from ten universities, the models achieve 99 percent model quality and discuss the impact of data distribution across participating institutions
FedHome: Cloud-Edge based Personalized Federated Learning for In-Home Health Monitoring	[115]	2020	FL in Healthcare IoT	FedHome, a cloud-edge-based federated learning architecture for in-home health monitoring, is proposed in this research
The future of digital health with federated learning	[93]	2020	Survey	This survey report looks at how FL could help with the future of digital health, as well as the obstacles
Federated Learning on Clinical Benchmark Data: Performance Assessment	[63]	2020	Benchmark	The research uses three benchmark datasets, including a clinical benchmark dataset, to assess the reliability and performance of FL
FedMed: A Federated Learning Framework for Language Modeling	[116]	2020	Framework	To address model aggregation and communication costs in the FL environment, this study provides a unique Federated Mediation (FedMed) framework with adaptive aggregation, mediation incentive scheme, and topK method
Federated Learning for Breast Density Classification: A Real-World Implementation	[94]	2020	Medical Image Analysis	This article demonstrates the efficacy of FL by training a model for breast density categorization based on Breast Imaging, Reporting, and Data systems utilizing data from seven clinical institutions across the world (BI-RADS)
Federated Transfer Learning for EEG Signal Classification	[56]	2020	Learning Systems	This work proposes a unique privacy-preserving DL architecture called federated transfer learning that uses the FL in EEG classification (FTL)
COVID-19 detection using federated machine learning	[96]	2021	Diagnosis	To determine which parameters impact model prediction accuracy and loss, this study employed a descriptive dataset and chest x-ray (CXR) images from COVID-19 patients in an FL context
Implementing Vertical Federated Learning Using Autoencoders: Practical Application, Generalizability, and Utility Study	[16]	2021	Learning Systems	Without revealing the raw data, this research shows that FL on vertically partitioned data may perform equivalent to centralized models
Federated Learning Meets Human Emotions: A Decentralized Framework for Human–Computer Interaction for IoT Applications	[23]	2021	FL in Biomedical	This article combines facial expression and voice inputs to construct an emotion monitoring & analysis system using FL
FeARH: Federated machine learning with anonymous random hybridization on electronic medical records	[28]	2021	Learning Systems	This research study suggests a novel FL method to deal with untrustworthy conditions

(Continued)

Table 1. Continued

Title	Ref.	Year	Category	Content
Federated Learning for Thyroid Ultrasound Image Analysis to Protect Personal Information: Validation Study in a Real Health Care Environment	[64]	2021	Diagnosis	The purpose of this research is to see if FL's performance is equivalent to that of traditional deep learning
Learning From Others Without Sacrificing Privacy: Simulation Comparing Centralized and Federated Machine Learning on Mobile Health Data	[71]	2021	FL in Healthcare IoT	The research explores FL use cases in a mHealth environment and uses an mHealth data set to simulate federated learning
Cloud-Based Federated Learning Implementation Across Medical Centers	[92]	2021	FL in EHR data	This research mimics an FL environment in order to investigate multiple federated learning implementations and apply FL algorithms to data from two academic medical facilities' electronic health records
Federated learning improves site performance in multicenter deep learning without data sharing	[97]	2021	FL in EHR data	This study demonstrates how to provide multi-institutional training in an FL environment without centralization
A Resource-Constrained and Privacy-Preserving Edge-Computing-Enabled Clinical Decision System: A Federated Reinforcement Learning Approach	[121]	2021	FL in EHR data	This article combines mobile-edge computing (MEC) with software-defined networking to make use of the processing and storage capabilities available among edge nodes (ENs) (i.e., MEC servers) in the FL environment
Variation-Aware Federated Learning with Multi-Source Decentralized Medical Image Data	[122]	2021	Learning Systems	Variation-aware federated learning (VAFL) is a methodology proposed in this research for minimising client variations by transforming all clients' images into a shared image space
Federated Learning in a Medical Context: A Systematic Literature Review	[87]	2021	Survey	This survey article examines federated learning and its relevance to sensitive healthcare data
Federated Learning for Smart Healthcare: A Survey	[80]	2021	Survey	The application of FL in smart healthcare and IoT devices are reviewed and surveyed in this survey report

- **SMPC - Secure MultiParty Computation.** Secured MultiParty Computation approach provides clear security guarantees even though it is, in general, a less efficient approach. **Privacy-Preserving Record Linkage (PPRL)** [61] is one such example where secured multi-party computation is being used. SMC was also introduced to **Quantitative Structure-Activity Relationship (QSAR)** and drug-target Interaction prediction for drug discovery [61]. SMC applies the concept of three distinct responsibilities for parties. There are input parties who contribute data to the calculation that must be protected. Computation parties carry out the privacy-preserving computation on the data given. Apart from what is revealed through the architecture of the application, the computation parties should not learn anything new from the execution. Finally, there are result parties who are given the computation's results.

2.5 Machine Learning Pipeline

The life cycle of a federated learning system consists of eight primary steps which include task initialisation, selection, configuration, model training, client server communication, scheduling and optimisation, versioning testing deployment, and termination. The steps are illustrated in the diagram shown below (Figure 6).

The Table 2 shows the Summary of Existing Studies on Deployment of ML and FL in Healthcare Since 2016. In FL settings, three major machine learning steps include model selection, model training, and hyperparameter tuning.

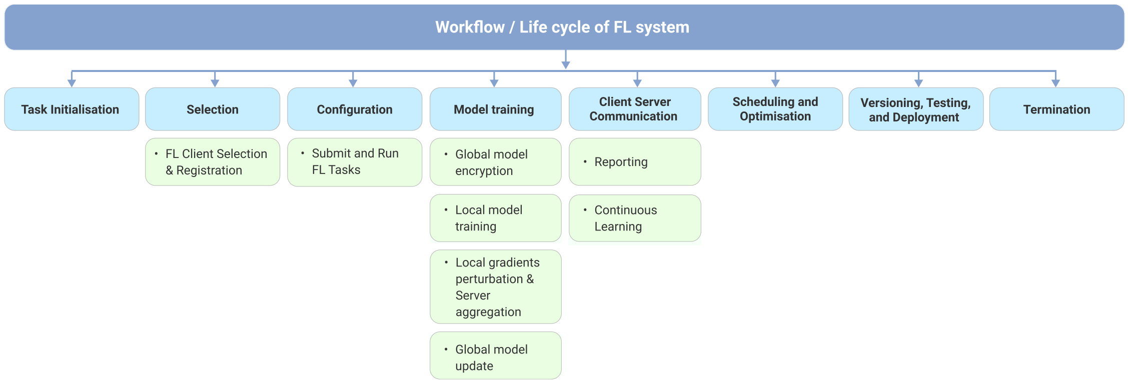

Fig. 6. Machine Learning pipeline.

2.5.1 Model Selection. The model selection stage includes the process of selecting the optimal model for federated data. Model selection depends on the type of task as well as size, quality, and type of the federated data. For example, if the problem is classification on diagnostic images, a convolution neural network is the best choice for this task. There have been several attempts in recent years to propose and create new models for federated settings. However, most FL tasks consider state-of-the-art, widely used models in their setting. **Neural networks (NN)** are the most popular and state of the art in many machine learning tasks in the FL setting. For example, [42] uses a variant of LSTM called **Coupled Input and Forget Gate (CIFG)** to predict the next word in the mobile keyboard. [21] uses a specific Convolutional Neural Network (CNN) network with a transfer learning method for the classification task. A tree-based FLS has shown impressive performance on classification and regression tasks. [65] used the FLSs for **Gradient boosting decision trees (GBDTs)** on horizontally and vertically federated data, respectively. Most Federated Learning frameworks use stochastic gradient descent methods to optimize machine learning models such as neural networks and logistic regression. Besides Neural Networks and tree-based networks, linear models (e.g., linear regression, support vector machines) are standard and easy-to-use methods.

2.5.2 Model Training. Following the selection of a model, the next stage in the ML pipeline is model training. Different variants of the model and a combination of optimizing hyperparameters are used to decide the final model with good accuracy. This process [9] involves the following steps:

- Training the local models on their local training dataset
- Sharing of the local parameters to the server
- Aggregation of local models' parameters on the server using the aggregation operator and
- Updating the local models with the aggregated global model
- Repeat the loop

2.5.3 Model Parameters. The parameters are learned from federated data, and hyper-parameters are utilized to fine-tune the output for the best match. The hyper-parameters such as learning rate, number of training epochs, mini-batch size, and optimizer have to be tuned based on the constraints of the ML application (e.g., available computing power, available memory, bandwidth). [114] demonstrated how local data could be used to fine-tune federated models. They provided methods for determining the appropriate hyper-parameters for fine-tuning and proved that it enhances models' next word prediction in mobile keyboards. [129] offered numerous variations of the fine-tuning strategy to improve the local adaptation. The network architecture is designed manually, which takes a significant amount of effort and experience in that field. **Neural Architecture Search (NAS)** is an algorithmic-based method to search the neural network design utilizing optimization algorithms. In

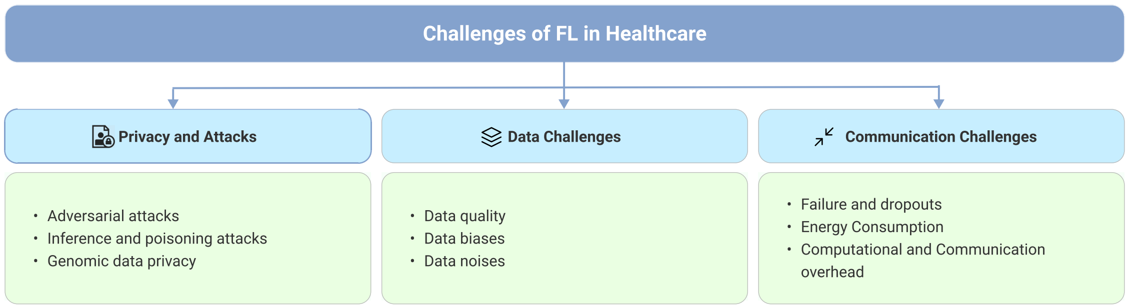

Fig. 7. Challenges of Federated Learning in healthcare.

recent years there has been progress in federated neural architecture search. It aims to optimize the design of models in the federated learning environment. The offline federated NAS framework proposed by [132] uses a multi-objective evolutionary algorithm to design the optimal network.

3 CHALLENGES & ISSUES

Some various challenges and issues can be found at each step in the implementation of a federated learning system. In this paper, we classify the challenges and issues into three major subcategories. The first subcategory involves all issues related to the privacy of the model and the attacks that happen on the system. The second subcategory consists of papers that describe the issues related to data such as limited data, data bias, or data poisoning.

The kinds of communication that can take place in a federated learning setup are between a server and an edge device, between two edge devices, and/or between more than one server. The main challenge here is to answer some of the following questions: How effective can this communication be without having to lose the participating devices? How can the number of communication rounds be reduced to avoid large usage of battery or net on edge devices? And, how can all of this be done without affecting the model and its result metrics? This subsection brings down papers addressing these major issues. Hence, the third subsection speaks about communication challenges. An illustration of this section is presented in Figure 7.

3.1 Privacy and Attacks

- **Adversarial Attacks.** Improper or inadequate learning refers to cases where improper hyperparameters are learned in the ML/DL model, e.g., learning rate, epochs, and batch size. In a predictive healthcare setting, machine learning models are created using previous patient data and then evaluated on new patients, raising concerns about the validity of the predictions. This is because of distribution shifts, and such differences can be exploited for generating adversarial examples [84] which is now a huge concern. In addition, ML/DL versions are strictly susceptible to various risks to the protection and privacy, such as adversarial attacks [11, 106].
- **Inference and poisoning attacks.** Federated learning systems are yet to be aware of the future federated learning algorithm design on privacy preservation [74]. This survey on threats to federated learning provides a concise introduction to the topic of federated learning and the two major federated learning methods are exposed to inference attacks and poisoning attacks. It also summarizes the kinds of threat models a system can be prone to and highlights the key techniques as well as fundamental assumptions adopted by various attackers.
- **Evasion Attack.** Evasion/exploratory attacks are among the most common adversarial attacks and are carried out during the inference time. This setting does not imply any influence over the training data.

In this attack, an adversary attempts to evade the deployed model by feeding malicious data samples or collecting evidence about the model features during the inference phase. The amount of knowledge available to the adversary about the model determines the efficiency of such attacks. Various methods have been presented to make the FL models more robust against evasion attacks.

- **Inadequate model training.** Machine learning models while training may involve flaws. The model training flaws include improper or inadequate training, violations of privacy, model poisoning, or theft. Improper or inadequate learning refers to cases where improper parameters are learned in the ML/DL model, e.g., learning rate, epochs, and batch size. The deployment of the model involving ML/DL techniques mainly revolves around human-centric decisions. Hence, considering fairness and accountability while ensuring the robustness of the system is crucial at this stage. The deployment stage can involve Evasion attacks, System Disruption, Network Issues, etc.
- **Gradient Inversion Attack.** This attack strategy implies that it is possible to recover and reconstruct input data from the gradient knowledge of trained and untrained model parameters. [38] investigates the impact of design and settings on the difficulty of reconstructing an input image from grading data and demonstrates that any input to a fully connected layer could be reconstructed analytically without regard to the rest of the model design.
- **Privacy in Genomic Data.** The human genome is a complete set of genetic information of a human living organism composed of four different bases (A, T, G, C) and can provide a treasure of highly sensitive and personal information of an individual. With the innovation in the next-generation technologies, a whole new genome complex can be determined of an organism. The use of genome sequencing is used for purposes like personalized genomic medicine, disease diagnosis, and preventive treatment. [2] discusses the various privacy issues in genomic data processing such as querying on genomic data and carrying alignment processes on commercial public clouds in a privacy-preserving and effective manner. The paper also concludes that despite the innovations and study in medicines and health science, the use of sequencing technology still lacks privacy preservation of genomic data leading to a lot of genomic data leakage.

3.2 Challenges Related to Data

- **Noises in different kinds of data.** The healthcare industries face a huge threat when it comes to data collection. Large amounts of clinical data are collected in the form of EHRs, medical images, radiology reports, etc., which requires a lot of human effort and is time-consuming. Multishot MRI, one of the widely used imaging modalities used to acquire high-resolution medical images can involve Instrumental and Environmental Noise due to some undesirable artifacts in the resulting image [90]. A variety of health data including patient data from multi-omic approaches, as well as clinical, behavioral, environmental, and drug data, can be analyzed using the AI technologies being developed at the present. [112] mentions five major types of data used in AI for health such as multi-omics data, clinical data, behavioral/wellness data, and environmental data, as well as research and development data. Each data type has its challenges, implications, and future directions.
- **Improper annotation of data.** Healthcare datasets involve annotation of data samples which is a crucial step in the process of healthcare applications. Hence, they should be performed by legally keeping privacy concerns in mind and with proper guidelines. The inability to perform labeling rightly leads to improper annotations and many efficiency challenges like an imbalanced dataset, class imbalance, and Bias and data sparsity [90].
- **Data biases.** Data biases [58] act as the main driver of unfairness in machine learning models. It can result in high risks when used in federated learning systems. Moreover, biases can end up affecting areas like training data, cognitive sampling, reporting, and confirmation biases. **Artificial intelligence (AI).** models often need a vast volume of high-quality training data, which contrasts sharply with the existing drug development pipelines' limited and skewed data. And, this decentralized machine learning paradigm

has a high scope in contributing to the improvement of AI-Based drug discovery [118]. The superiority of the federated learning mechanism is demonstrated on pooled datasets with seven aqueous solubility datasets including high and low biases. The authors also address the small data and biased data dilemma in drug discovery and prove the promising role of federated learning.

- **False positives and false negatives.** The healthcare industry deals with a large amount of patient data and this involves high chances of containing missing observations or variables. However, ignoring these values during analysis by knowing their relationships with already observed and unobserved data is the simplest way to avoid them. On the other hand, using these missing observations leads to well-known problems like false positives and false negatives, and thus there is a need for complete and compact healthcare data. Issues like false positives and false negatives not only are caused due to the use of missing values but also due to incomplete training or inefficient training of the model. The root cause behind inefficient training is the incomplete data being fed for inference. ML-powered healthcare demands cautious applications of analytical methods [88]. And that is why along with quantity, quality of data also plays a major role.
- **Data Quality [15].** A survey paper on applicable machine learning algorithms in a federated environment mentions the importance of **Independent and Identically Distributed (IID)** data points and their contribution to the factor of lowering the chances of class imbalance. Many times the local data collected can lack in size and major data points, which may obstruct representing the underlying structure of a federated learning system. Using such a small-sized imbalance dataset has a high chance of hindering the global model and its accuracy. [54] has trained the global model on historical patient data of **Radiation Therapy (RT)**; because RT technique is constantly evolving, such models provide little value to clinical practice today. In addition, Phase III clinical trials give high-quality evidence, but they have the drawback of taking a long time to complete. However, including more recent patient data and updating the model with the most up-to-date practice insights might help alleviate the problem. According to the author of the [127], federated learning for an institution/hospital takes several months to develop. Furthermore, obtaining data from electronic health records (EHR) remains challenging since data is typically dispersed across several databases and apps.
- **Data Leakage.** Federated learning provides a solution that enhances data privacy, which is a crucial concern of FL training. However, some work has proposed showing the leakage in federated learning representing a still unexplored area of research due to several factors that may result in security issues. A quantifiable method of measuring privacy would enable better choices about the minimum privacy settings required to sustain clinically acceptable results [37].
- **Patient Variability.** Care varies widely between locations, with significant variation in performance across indicators. Hospital levels could explain more of the observed diversity than clinician levels.

  In a study, more than 7,000 cardiovascular patients were treated with adult stem cells [36]. However, the data so far has revealed neutral or minor advantages. One of the most important questions to explore when thinking about designing better trials is why there is so much variation both across and within trials. Patients in the FOCUS-CCTRN trial [86] received autologous bone marrow cells to treat chronic ischemic heart disease. Several cardiac clinical outcomes showed no improvement when compared to placebo. On the other hand, individual patient outcomes were highly varied [10].

  Care variability or uneven care practices has an impact on many elements of healthcare delivery. It can assist in improving patient safety, reducing healthcare expenditures, and improving **key performance indicators (KPIs)** for both people and the healthcare system by minimizing variability of care.

3.3 Communication Challenges

- **Failure or dropouts during communication.** While distributed learning conjointly aims at training one model on multiple servers, a standard underlying assumption is that the native datasets are Identically Distributed (IID) and roughly have constant size. None of those hypotheses are created for federated learning;

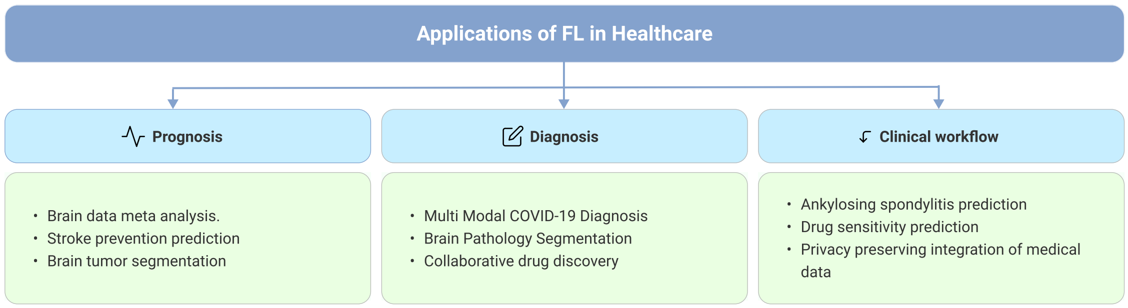

Fig. 8. Applications of federated learning in healthcare.

instead, the datasets are usually heterogeneous and their sizes could span over many orders of magnitude. Moreover, the clients concerned in federated learning could also be unreliable as they're subject to a lot of failures or drop out. Some clients' models do not get included in every FL round due to interrupted connectivity or slow internet conditions. Those devices or models are called stragglers [37]. Since the clients usually trust less powerful communication media (i.e., Wi-Fi) and powered systems (i.e., smartphones and IoT devices) compared to distributed learning where nodes are usually data centers that have powerful procedure capabilities and are connected to at least one another with quick networks.

For an Edge AI task to be accomplished there are multiple communication rounds between edge nodes. Edge nodes are usually smartphones or devices. Small access to training data is available at every edge node that is participating in the training. These nodes then perform edge training which results in high communication costs, especially in cases of limited bandwidth [130]. The aim of communication in every round is to compute a certain function value concerning intermediate value at edge devices. The paper points out a challenge of alleviating communication overheads under privacy and resource constraints and the need of reducing the communication rounds for training and inference. The paper also introduces communication efficient methods to achieve efficient results. Communication methods for edge AI at the algorithmic level, zeroth-order method, first-order method, second-order method, and federated optimizations have been well illustrated and explained.

- **Computational and Communication Overhead.** Communication has become one of the primary challenges for federated learning as the wireless networks and end-user internet connections can potentially become expensive and unreliable soon. [58] mentions how federated averaging and sparsification and/or quantization of model updates have demonstrated a significant reduction in communication cost with minimal impact on training accuracy. [27] demonstrates that employing secure MPC experiments to do privacy-preserving federated genomic data analysis is more costly than performing the exact computation in a centralized non-encrypted environment.
- **Energy consumption while communication.** IoT (Internet of Things) involves widespread use of mobile devices involving computing and sensing capabilities which involves a collection of data at a societal scale. Valuable data collection and maintenance backed with centralized machine learning models entail security and privacy issues leading to less participation of devices in smart city methods. While the mechanics of privacy preservation in federated learning ensures the privacy of data is preserved, mobile crowdsensing involves a huge amount of energy consumption. Federated learning involves on-device training in turn earning leading to high consumption of batteries from local devices which in turn deters users from participating in the process. For instance, this [53] survey paper presents the potential of federated learning along with the overview of challenges and issues faced while incorporating federated learning into smart city sensing.

4 APPLICATION

In this section, we present a comprehensive of survey existing works on federated learning in the health sector and propose a list of applications classified into prognosis, diagnosis, and clinical workflow. The Table 1 shows the summary of Existing Studies on FL in Healthcare Since 2015, and Figure 8 illustrates the classification of Applications of federated learning in healthcare.

4.1 Prognosis

Prognosis is a diagnostic term that refers to assessing the possible or potential progression of an illness, such as whether the signs and effects will change or worsen (and how quickly) or stay constant over time. The natural course of the diagnosed illness, the individual's physical and emotional health, the current medications, and other considerations, are used to make a prognosis. Listed below are healthcare-related applications helping to perform a prognosis on a particular disease or type of data.

Privacy preserving stroke prevention. An investigation into Facebook's data protection breaches in processing consumer information for uninformed usage has sparked recent data privacy issues. As a result, legislation such as the **General Data Protection Regulation (GDPR)** [Regulation, 2016] has been proposed to prevent organizations from sharing data without prior user consent. To fix data privacy, they use a new and efficient methodology called federated learning platform, which allows them to jointly train a machine learning model using data from multiple clients without directly exchanging data between them. Tencent and WeBank collaborated to create a privacy-preserving stroke prediction technology. Stroke prevention, as well as the risk factors associated with it, has long been a public health priority around the world. The scientists and engineers suggest a privacy-preserving scheme for predicting stroke risk and intend to use cloud servers to deploy the federated prediction model [57].

Meta analysis on brain data. In reality, data sharing is limited by the need to migrate vast amounts of biomedical data, as well as the administrative workload that comes with it. Researchers sought an analysis approach in meta-analysis or federated learning paradigms as a result of this situation. The **Enhancing NeuroImaging Genetics by Meta-Analysis (ENIGMA)** consortium is one of the best examples of this kind of research method. Brain scans of previously unimaginable amounts can be found in data banks all around the world. Due to various privacy and legal considerations, separate databases held at multiple locations cannot always be exchanged directly, restricting the use of big data in the study of brain disorders. A Federated learning platform allows them to jointly train a machine learning model using data from multiple clients without directly exchanging data between them. Addressing the issue of privacy and restrictions on sharing the data, the authors propose a federated learning system for safely accessing and meta-analyzing any biomedical data while maintaining individual privacy. They demonstrate the framework by using the ENIGMA Shape platform to provide the first implementation of federated analysis that is consistent with ENIGMA's standard pipelines [102].

Brain tumor segmentation. Classification of electroencephalographic (EEG) recordings and brain tumor segmentation was difficult due to the need of large datasets [56, 68]. This lack of large datasets has led to the success of deep learning (DL) methods in the field of Brain-Computer Interfaces (BCI). A novel privacy-preserving DL architecture has been suggested for EEG classification called federated transfer learning (FTL). A test proposed architecture's success on the PhysioNet dataset for 2-class motor imagery classification [56]. Demonstrating the possibility of implementing differential-privacy strategies to secure the patient data in a federated learning setup, a brain tumor segmentation was performed on the BraTS dataset [68].

Breast density classification [94]. For the advancement of clinically relevant models, hospitals and other academic institutes often need to partner and host centralized databases. Owing to data protection and ethical issues involved with data sharing in healthcare, this overhead will easily become a technical problem and typically necessitates a time-consuming approval process. And if these issues are overcome, data is precious, and organizations may choose not to exchange complete datasets. In a real-world collaborative environment, the use of federated learning (FL) to create a medical imaging classification model was demonstrated bringing together seven health organizations from around the world to train a model for breast density classification focused on **Breast Imaging, Reporting, and Data System (BIRADS)** [94].

Multi-Disease Chest X-ray classification. Deep learning approaches can yield important results in medical imaging analysis, although they involve a large volume of high-quality data. Since a client cannot have enough data to train and construct a quality model, working with multiple clients may address data insufficiency problems in deep learning but add privacy restrictions. This paper [75] proposes a federated deep learning method for multi-disease classification from chest X-rays, with pneumonia as an example. The FDL technique measures pneumonia from a chest X-ray and distinguishes between viral and bacterial pneumonia. Clients train local models with minimal private data at the edge server and send them to the central server for global aggregation without sending the chest-X-ray images to a central server [8].

Adverse Drug Reaction prediction. Since healthcare data is scattered, collecting a sufficiently comprehensive dataset to track rare cases entails combining data from various data silos. Analyses derived from various data sources may be contradictory or inaccurate, necessitating the use of tools to properly aggregate the information. A time lag exists between the **ADR (adverse drug reaction)** case, claim filing, adjudication, and claim consolidation into a database of current claims-based systems. As a result, there is an unmet need for reliable, flexible, and effective methods for forecasting ADRs using distributed health data while maintaining patient privacy. Since healthcare data is dispersed, assembling a robust dataset to document unusual events necessitates merging data from several data silos. To address the issue a federated learning-based system has been proposed, which allows health data to be shared across several platforms. Without ever transferring the raw data from their respective pages, the architecture helps one to train a global model based on each site's local data. It is the first time federated machine learning algorithms have been used to forecast ADRs using distributed electronic health data [24].

Predictions on SARS-COV-2 chest X Rays. Confidential data all pose realistic difficulties when using electronic health data to forecast adverse drug reactions (ADR). Another example where there is a need for data collaboration is within clinical and science communities. When adapting to rapidly changing and pervasive environmental threats, the pandemic has highlighted the importance of quickly conducting data collaborations. One recent work on an AI-based SARS-COV-2 **Clinical Decision Support (CDS)** algorithm is a concrete example of these forms of partnerships. Another example is during the SARS-COV-2 pandemic, 20 institutes partnered on a healthcare FL study that used vital signs, laboratory results, and chest x-rays to predict possible oxygen needs of infected patients, resulting in the "EXAM" (EMR CXR AI Model) [37].

GWAS analysis on genomic datasets. The biomedical community benefits from the growing availability of genomic data for scientific studies, such as **Genome-Wide Association Studies (GWAS)**. However, high-quality GWAS typically necessitates a large number of tests, which may exceed a single institution's ability. Concerns regarding patient safety and clinical knowledge security arise from federated genomic data analysis (as data are being exchanged across institutional boundaries). On federated genomic databases, a privacy-preserving GWAS architecture. has been proposed where computations are layered on top of stable **multi-party computation (MPC)** structures [27].

NeuroLOG. A creation of a framework OntoNeuroLOG was implemented by federating five neuroimaging data repositories in Paris, Rennes, Grenoble, and Sophia Antipolis. The creation of the framework OntoNeuroLOG and its use to confirm heterogeneous data to a standard model are the main features of this work. The project focuses on research implementations that need a multi-center, multi-disciplinary approach: (1) epilepsy (surgical treatment of drug-resistant epilepsy) and (2) neurodegenerative disorders (Alzheimer's disease) [39].

4.2 Diagnosis

The description of the type and origin of a certain phenomenon is referred to as diagnosis. Diagnosis is used in many different fields to assess "cause and effect," with differences in the application of logic, analytics, and practice. Listed below are healthcare-related applications helping to perform a diagnosis on a particular disease or type of data such as drug discovery, COVID-19 prediction at the edge, and the use of EHRs to improve mortality prediction.

COVID-19 diagnosis at the edge. COVID-19 has been a major focus of research in 2020, especially after the **World Health Organization (WHO)** declared it a pandemic in March, with various activities focused on diagnosis, prevention, and the production of a possible vaccine. Risk identification [82], touch monitoring, false news identification, emotion analysis, and screening and diagnosis are some of the main applications of data science approaches, especially machine learning and data visualization techniques, in the international response to the COVID-19 pandemic. Despite major advancements in recent years, cloud-based healthcare systems appear to be underutilized due to their shortcomings in meeting rigorous protection, privacy, and quality of service criteria (such as low latency). The authors take advantage of edge computing's capabilities in medicine by exploring and testing the ability of intelligent analysis of clinical visual data linked to COVID19 at the edge. They also enable remote healthcare centers to benefit from a multi-modal shared learning model without having to share any knowledge about the local data's modality or the data itself. The authors suggest a CFL-based collective learning system for the role of COVID19 diagnosis with visual evidence such as X-rays, Ultrasound images, and CT Scans, based on an emerging idea of **clustered federated learning (CFL)** [89].

Estimation of blood pressure. According to the World Health Organization (WHO), chronic heart disease was the leading cause of death from 2000 to 2019, accounting for 16% of all global deaths in 2019. During this time, most deaths have been caused by heart disease. This not only has a significant impact on the lives of those involved but also on public healthcare services. **Electrocardiogram (ECG)** and **blood pressure (BP)** readings are widely used by clinicians to consider the dynamics between the healthy and dysfunctional core. These methods are also very invasive, particularly when continuous **arterial blood pressure (ABP)** readings are taken, and they are often quite expensive. To address the problem, the authors present a decentralized learning approach to continuous ABP calculation that is capable of large-scale real-world deployment while protecting patient privacy. This architecture, to their knowledge, is the first example of a GAN capable of continuous ABP generation from an input PPG signal and using a federated learning methodology [14].

Data Variability in Medical Imaging. Deep learning has made rapid strides in image recognition and target detection in recent years. These advancements have also led to improvements in automating clinical processes within medical imaging, thanks to Convolutional Neural Networks (CNNs) pattern recognition abilities. Deep CNNs, for example, have paved the way for breakthroughs in retinopathy diagnosis, lung nodule identification, and brain tumor segmentation. Insufficient patient data makes it difficult to train deep learning models for medical applications, particularly for rare diseases. Efforts to share patient data are often hampered by legal, technological, and privacy issues. The current implementation of CWT has a major flaw in that it isn't designed to accommodate differences in sample sizes, mark distributions, resolution, and acquisition settings in training data through organizations. The authors [7] present CWT modifications to reduce output losses caused by heterogeneity in training sample sizes and mark ranges through academic training splits, and test the effectiveness of their changes on virtual dispersed tasks for (DR) identification and irregular chest radiograph classification. This is the first research to show that data heterogeneity in training sample sizes and mark distributions across institutions can cause distributed learning models for medical imaging to perform poorly [7].

A federated network for translational cancer research. Obtaining adequate quantities of annotated human tissues remains a major barrier to translational cancer science, which is needed to move cancer treatment closer to precision medicine. Major new bridging infrastructures, including more functional biorepositories that connect human tissue to clinical phenotypes and outcomes, are required for advancements in cancer science and personalized medicine. Cancer researchers have been at the forefront of creating biomedical data and resource sharing consortia, but they have traditionally relied on centralized structures in which a single organization serves as a middleman between requesting researchers and participating institutions. The downside of centralization is that as the number of organizations grows, it becomes precarious. The **TIES (Text Information**

Extraction System) Cancer Research Network was established by four cancer centers as a federated network that allows member organizations to share data and biospecimen. [52] mentions pathology data that has been de-identified and analyzed with the TIES natural language processing framework can be accessed by member sites, resulting in a pool of rich phenotype data linked to clinical biospecimens. The possible effect of federated quests around the network on translational science can be seen in studies involving rare diseases, uncommon phenotypes, and complex biological behaviors. The network meets many main criteria, including local data and credentialing power, the inclusion of rich phenotype data, and applicability to a wide range of study goals [52].

Multi-site fMRI analysis. Data has a "non-rivalrous" value, which means it can be used by several parties at the same time to produce new data items or services, according to economics literature, Data pooling would have a synergistic impact. Sharing vast volumes of medical data is critical for precision medicine, with **functional MRI (fMRI)** data relating to certain neurological conditions or disorders being an interesting example. Deep learning models have shown to be useful in a variety of functions, including neuroimage processing. However, to successfully train a high-quality deep learning algorithm, a large volume of patient data must be gathered. The time and cost of acquiring and annotating massive fMRI datasets, for example, make it impossible to obtain a large number at a single location. The authors of the paper [69] use a privacy-preserving approach to solve the issue of multi-site fMRI classification. They suggest a federated learning approach to solve the problem, in which a decentralized iterative optimization algorithm is used and mutual local model weights are changed by a randomization mechanism. Overall, the findings show that using multi-site data without exchanging data can improve neuroimage analysis accuracy and help discover accurate disease-related biomarkers [69].

Patch-Based Surface Morphometry for Alzheimer's Disease. In hospitals and academic centers, unprecedented rates of brain **magnetic resonance imaging (MRI)** currently exist. Simultaneously, the rapid advancement of software and hardware has made it scientifically possible to extract useful knowledge about the underpinnings of brain diseases such as Alzheimer's disease from these combined databases (AD). Researchers have faced significant challenges in obtaining or sharing these details due to patient safety issues, data limitations, and legal complications. To address this issue, large-scale collaborative networks, such as the ENIGMA Consortium1, were established, which used secure meta-analyses to study data from hundreds of institutions around the world without sharing patients' scans or protected information. Seeking major causal factors that may predict/relate to health conditions or cognitive function, such as finding anatomically irregular regions in the brains of **Alzheimer's Disease (AD)** patients, is more interesting in brain imaging studies. As a result, the authors suggest a novel federated feature selection scheme based on group lasso regression using patch-based surface morphometry features from T1-weighted brain MRI images of AD, **mild cognitive impairment (MCI)**, and stable elderly test subjects. By deliberately choosing (and visualizing) core functions, their work generalizes and enriches federated learning science. The method can discover new important features to be used as imaging biomarkers of MCI and AD by expanding access to information from large imaging datasets [116].

FADL: Federated-Autonomous Deep Learning for Distributed Electronic Health Record. Data from electronic health records (EHRs), patient-generated health data from mobile devices, and other health-related information are useful for optimizing health outcomes, especially in precision medicine. Healthcare records are kept in various locations and data silos, such as clinics, pharmacies, payors, and mobile computers. Traditionally, healthcare data are disseminated through several locations and consolidated in a network for review. However, due to stringent rules and the sensitivity of the results, healthcare data transfers are complicated. These impediments not only make data usage costly, but also slow down knowledge delivery in healthcare, where timely changes are often needed. Using ICU data from 58 independent hospitals, the authors [70] demonstrate that machine learning models used to forecast patient mortality can be trained effectively without taking health data out of silos using a distributed machine learning approach. They suggest a new approach called **Federated-Autonomous Deep Learning (FADL)**, which trains a portion of the algorithm using data from all data sources in a distributed manner and another portion using data from individual data sources [70].

Clinical trial protocol optimization. Clinical science is a time-consuming, labor-intensive, and expensive undertaking. Specific challenges associated with these bottlenecks include problems assessing patient demographics, determining eligible patients for enrollment, optimizing procedures, manual and inefficient data collection, data source reliability, and the difficulty of recognizing and monitoring infrequent adverse incidents. Furthermore, there are workflow problems and bottlenecks that hinder clinical trial behavior, such as suboptimal research design, slow and lengthy patient registration, site selection, and procedure optimization, all of which contribute to time and cost requirements. Some of these problems, such as protocol optimization and patient selection, can be mitigated by prudent re-use of data found in Electronic Health Records (EHRs). The growing use of EHRs in Europe and elsewhere provides a large, rich, and highly important pool of health data that has the potential to enhance clinical trial delivery. This data may be used to assess clinical trial viability using computable representations of the parameters, improve patient identification, clinical trial execution, and adverse effect monitoring. The authors in [25] investigated the concerns by examining the **Inclusion and Exclusion (I/E)** criteria of 23 completed trials in a variety of therapeutic areas that were sponsored by seven pharmaceutical companies to determine the proportion of I/E criteria that could be represented in a computable format and the right to query hospital EHRs to determine the number of currently qualifying patients correctly [25].

Heart Disease predictions from Electronic Health Records. In the age of "big data," computationally efficient and privacy-aware solutions for large-scale machine learning problems are critical, especially in the healthcare context, where large volumes of data are processed in several locations and owned by various institutions. The authors here discuss three issues concerning healthcare data: (1) data exist in several places (e.g., clinics, physicians' offices, home-based computers, patients' smartphones); (2) data access is increasing, necessitating the use of scalable frameworks; and (3) aggregating data in a centralized database is infeasible or impractical due to size and/or data protection issues. Based on their medical histories as outlined in their Electronic Health Records, the authors create a distributed (federated) approach to forecast hospitalizations for patients with heart diseases within a target year (EHRs). They devise a federated optimization scheme (cPDS) to address the sparse Support Vector Machine problem. They apply their latest approach to a dataset of de-identified Electronic Cardiac Records from the Boston Medical Center, which includes patients with heart disorders [13].

4.3 Clinical Workflow

Clinical workflow is often conducted to optimize consistency in the workflow, analyze current frameworks, research a process and its implementation, and so on. The health care applications mentioned below conduct or include a clinical workflow on a specific disease, analysis on drug sensitivity, an EHR linking platform, and cloud-based output of federated learning on EHR's obtained from two healthcare systems to predict the risks of diseases linked to tobacco and radon.

FedMed: A Federated Learning Framework for Language Modelling. Society is entering a smart age, with the latest breakthroughs of the modern technological revolution-Industry 4.0 and **Internet of Things (IoT)** technologies-where all items are enclosed with a network of interconnectivity and automation through intelligent digital technique. Meanwhile, edging systems are flooded with heterogeneous data, ranging from real-time sensor activity logs to consumer data. During the migration process, however, data is quickly attacked and poisoned. This increases the difficulty of machine learning. The emergence of federated learning techniques, as well as the obstacles and risks it poses, has occurred in recent years. Traditional FL strategies depend on averaging aggregation or don't take into account connectivity costs. The authors suggest a novel **Federated Mediation (FedMed)** paradigm with adaptive aggregation, mediation reward system, and topK strategy to solve concept aggregation and coordination costs in federated language modeling. Perplexity and contact rounds are used to test the results. Three datasets are used in the experiments (i.e., Penn Treebank, WikiText-2, and Yelp) [116].

Federated learning on clinical benchmark data. FL may be used to address privacy concerns and reduce the possibility of a data violation in clinical records so data transfer and centralization are not necessary. Since medical data is among the most vulnerable forms of personal data, privacy protection is especially important for medical data processing. De-identification techniques have traditionally been used to protect patients' privacy. A performance evaluation using federated learning on clinical benchmark data was performed. The **Modified National Institute of Standards and Technology (MNIST)** dataset, **Medical Information Mart for Intensive Care-III (MIMIC-III)** dataset, and **PhysioNet Electrocardiogram (PECG)** dataset were used in a federated learning analysis. By changing the MNIST, MIMIC-III, and ECG datasets, they also validated FL in environments that simulate real-world data distributions [63].

Identifying potential risk variants in ankylosing spondylitis. Genome-wide association studies (GWAS) have been common for detecting possible risk variants in a variety of diseases. Large sample size is usually needed for a statistically significant GWAS to identify disease-associated **single nucleotide polymorphisms (SNPs)**. A single institution, on the other hand, normally only has a small number of samples. As a result, cross-institutional collaboration is expected to maximize sample size and statistical capacity. However, cross-institutional collaborations present serious problems, one of which is data protection. While data sharing within a broad network can greatly support biomedical science, it also poses potential threats to data privacy due to the exchanging of personal information about individuals. The consequences of patient private information leaks include, but are not limited to, workplace discrimination, denial of benefits, higher insurance premiums, and so on. To address the problems of data privacy due to the exchange of personal information, the authors suggest a novel privacy-preserving federated GWAS architecture (iPRIVATES). iPRIVATES, which includes privacy-preserving federated processing, allows various organizations to collaborate on GWAS analysis without disclosing patient-level genotyping results [117].

Drug sensitivity prediction. Users with a personalized recommendation framework face a conundrum: learning from data will boost recommendations, but only when other users can share their private anonymization. Good personalized forecasts are critical in precision medicine, but the genetic knowledge from which the predictions are based is often especially vulnerable since it clearly describes the patients and therefore cannot be conveniently anonymized. Genomics is a critical domain for privacy-aware modeling, especially in precision medicine. Many people like to keep their and their descendants' genomes secret, but basic anonymization is insufficient since a genome is inherently recognizable. As a result, the hospital or clinic that holds the genomic data must be very vigilant about privacy concerns when disclosing the genomic data, even though the data is required and helpful for potential diagnosis and care decisions. Even with moderate-sized data, the proposed differentially private regression approach [47] combines the theoretical appeal and asymptotic efficiency with good prediction accuracy. Under comparatively strict differential privacy assurances, their approach exceeds the predictive precision of state-of-the-art non-private lasso regression with just 4x more tests [47].

Integration of Medical Data. Medical evidence is often maintained by many organizations. However, comprehensive evaluations can necessitate the integration of these datasets without jeopardizing patient or commercial privacy. An effective privacy-preserving algorithm, **Multiparty Private Set Intersection (MPSI)**, computes the intersection of several private datasets using this, the authors suggest a functional MPSI with the following characteristics: the size of each party's datasets is independent of that of the other parties, and the statistical complexity for each party is independent of the number of parties [77].

Design and implementation of EHRs. Electronic clinical evidence has increased dramatically as a result of healthcare laws and government incentives encouraging the use of Electronic Health Records (EHRs). As a result, academics and public health authorities have shown a growing interest in data linkage that can be used in cross-site health studies. Linking EHR data through healthcare agencies, on the other hand, necessitates striking a balance between data accessibility and privacy. Some regions have adopted **health information exchange (HIE)** programs to provide healthcare facilities with up-to-date clinical data on patients through hospitals for better and more organized treatment, as part of inter-institutional arrangements for the sharing of PHI. However, owing to concerns about reliability, anonymity, and protection, as well as other problems, organizational HIEs exchanging patient-level identifiers, are still not commonly used. The authors define a real-world implementation of a software framework **(Distributed Common Identity for the Integration of Regional Health Data–DCIFIRHD)** that uses a structured and distributed encryption algorithm to conduct stable, cross-site aggregation and linking of EHR data for analysis. As part of the HealthLNK study initiative, they applied the application in a major metropolitan area (Chicago, IL, USA), aggregating over 5 million patients' clinical records through six healthcare institutions [60].

Cloud-based FL. From diagnosis to medication decisions, machine learning (ML) models can improve health care. However, ML model generalizability is jeopardized due to a lack of adequate heterogeneous data due to patient privacy concerns. When compared to a simulation in a single organization, heterogeneous data from multiple centers increased model accuracy. Furthermore, cloud systems come equipped with the required tools and security measures to support federated learning deployments. The amount and quality of data used to train an ML algorithm are extremely important, particularly for more complex models. The availability of diverse multidimensional patient data sets in the age of precision medicine necessitates greater population surveys for generalization. Furthermore, data shortage of underrepresented communities may contribute to prejudices if training data does not adequately portray these populations' characteristics. Quality of healthcare data and algorithmic problems are also well-known ML roadblocks. To assess the federated machine learning solution, the authors use electronic health record data from two academic medical centers on a Microsoft Azure Cloud Databricks network to test various federated learning applications in both a virtual and real-world setting. Using data from two healthcare systems' electronic health records (EHRs), they trained machine learning models to forecast the risks of diseases linked to tobacco and radon [92].

Clinical decision support system. Clinical decision support systems (CDSS) are computer-based applications designed to improve patient care and healthcare delivery by assisting clinicians in analyzing health information better and enhancing medical decisions [59, 105]. A CDSS uses knowledge management to get clinical insights from the patient and health-related data based on multiple factors. Clinical decision support systems assist the clinicians at the time of care and help to make a better care plan [3, 12].

CDSS comprises three main components, including Base, inference engine, and Communication Mechanism. The base can be classified into two types of systems called knowledge-based and non-knowledge-based.

Knowledge-based support systems are defined by well-established rules (IF-THEN statements) that determine 'what is true?' Support systems without a knowledge base, on the other hand, answer the question "what to do?" advising on the next steps for treatments, which drug to prescribe, etc. These systems still require a data source and use artificial intelligence (AI), statistical pattern recognition, or machine learning (ML), to make medical decisions and provide recommendations [6, 40, 110].

A large amount of data is often required to train an AI system, but clinical data's varied and sensitive aspects present a barrier to the standard centralized network. The author of the paper [108] proposed a deep learning-based clinical decision support system trained and managed under a federated learning paradigm to solve the data and privacy challenges. The paper utilized a novel strategy to ensure the safety of patient privacy and overcome the risk of cyberattacks while enabling large-scale clinical data mining.

5 TOOLS

This section presents the most commonly used tools to implement federated learning applications in the health care area. Figure 6 illustrates the Federated learning frameworks.

5.1 TensorFlow Federated

TFF [107] is an open-source framework to perform machine learning and computations on localized information. TFF allows developers to simulate federated learning algorithms on the models and data. TFF's interfaces are unionized in two layers. **Federate Core (FC)** and **Federate Learning (FL)** APIs. Federated Core API is the foundation layer upon which is built the Federated Learning API. This is a strong functional programming

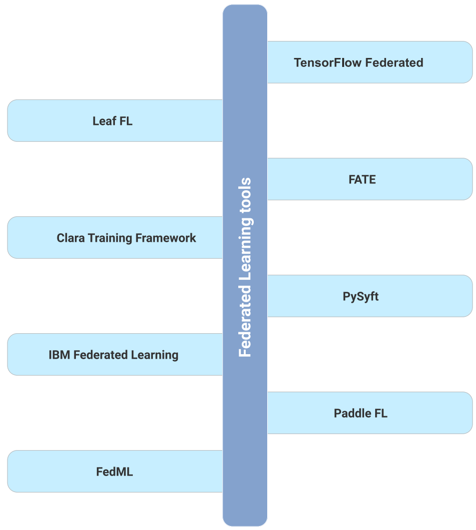

Fig. 9. Tools of federated learning.

environment that involves a combination of novel federated learning algorithms and TensorFlow with distributing higher-revelations operators at the core of the system. Whereas Federated Learning API is at a higher-level interface, this API allows developers to use the existing FL algorithms (training and evaluation) in their models. Developers can incorporate their functions and interfaces within the federated heart. FC, as a Python package, includes Python interfaces that can be used to create new Python features. It supports several styles, including Tensor types, sequence types, tuple types, and function types, to make it simple to use, particularly for developers familiar with TensorFlow. It supports a wide range of federated operators, including federated sum, federated minimize, and federated broadcast.

TFF currently only supports FedAvg and does not provide any privacy mechanisms. It can now only be deployed on a single computer, with the federated setting applied by simulation.

5.2 Federated AI Technology Enabler (FATE)

Webank's AI Department initiated an open-source project called FATE [35]. This framework provides a secure computing framework supporting the AI Ecosystem. It uses privacy-preserving techniques like Homomorphic encryption and Secured Multi-Party Computation. It also supports ML algorithms like Logistic regression, Deep learning, tree-based algorithms, and transfer learning. EggRoll, FederatedML, FATE-Flow, FATE-Serving, FATE-Board, and KubeFATE are its six main modules.

The federated algorithms and stable protocols are used in FederatedML. It currently supports training a variety of machine learning models, including NNs, GBDTs, and logistic regression, in both horizontal and vertical federated settings. To ensure anonymity, it also incorporates safe multi-party computing and homomorphic encryption. To run an FL algorithm, users simply set the parameters. FATE also includes comprehensive documentation of how to deploy and use it.

Practitioners must change FATE's source code to incorporate their federated algorithms because FATE has algorithm-level interfaces. Non-expert consumers would find this difficult.

Table 2. Summary of Existing Studies on Deployment of ML and FL in Healthcare Since 2016

Title	Ref.	ML Model	Data Type	Year	#Hospitals/clients	#Samples
Distributed learning: Developing a predictive model based on data from multiple hospitals without data leaving the hospital – A real life proof of concept	[55]	Bayesian network	EHR	2016	5	287
Developing and Validating a Survival Prediction Model for NSCLC Patients Through Distributed Learning Across 3 Countries	[54]	SVM	EHR	2017	3	894
Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records	[49]	k-means	EHR	2019	208	200,859
Distributed learning on 20,000+ lung cancer patients – The Personal Health Train	[31]	Logistic regression	EHR	2020	8	23,203
Federated Learning of Electronic Health Records Improves Mortality Prediction in Patients Hospitalized with COVID-19	[109]	Federated MLP,LASSO	EHR	2020	5	4029
Joint Imaging Platform for Federated Clinical Data Analytics	[98]	CNN based organ segmentation	Images	2020	10	-
Stochastic Channel-Based Federated Learning With Neural Network Pruning for Medical Data Privacy Preservation: Model Development and Experimental Validation	[99]	MLP	EHR	2020	5	30,760
Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data	[100]	U-Net	Images	2020	13	352
Federated semi-supervised learning for COVID region segmentation in chest CT using multi-national data from China, Italy, Japan	[123]	Fully Convolutional Network (FCN)	Images	2021	3	1704
Real-Time Electronic Health Record Mortality Prediction During the COVID-19 Pandemic: A Prospective Cohort Study	[104]	Stacked regression model	EHR	2021	12	28,538
Federated Learning used for predicting outcomes in SARS-COV-2 patients	[37]	ResNet-34	Images	2021	20	16,148
Cloud-Based Federated Learning Implementation Across Medical Centers	[92]	ANN/Logistic regression	EHR	2021	2	10,000
Federated learning improves site performance in multicenter deep learning without data sharing	[97]	3D Anisotropic Hybrid Network	Images	2021	3	300

5.3 PySyft

PySyft [95] is a secure and private deep learning library. It is a library for answering questions using data you cannot see. It uses privacy-preserving techniques like Differential Privacy, Homomorphic Encryption, and multiparty computation with deep learning frameworks like Pytorch and Tensorflow.

Both PyTorch and TensorFlow can be used with PySyft. It can be installed on a single computer or several computers, with the WebSocket API used to communicate between clients.

Though PySyft offers several tutorials, there is no comprehensive documentation on the system's interfaces or architecture. PySyft does not support a diversified computing paradigm like on-device training on Mobile or IoT.

5.4 Leaf

LEAF [62] is a benchmarking framework for studying in federated settings, with programs consisting of federated learning, multi-task mastering, meta-learning, and on-tool learning.

It includes six databases that cover a variety of topics, such as image recognition, emotion analysis, and next-character prediction. A collection of utilities is given to split datasets into separate parties in an IID or non-IID manner. A reference implementation is also presented for each dataset to explain how to use the dataset in a training phase.

Leaf databases have enough clients to model cross-device FL scenarios, but they may be too limited for questions where size is especially relevant. It only supports standardized algorithm implementations such as Fed Avg and does not support decentralized federated learning, split learning, and vertical federated learning.

5.5 Paddle FL

Paddle FL [81] provides applications in Natural Language Processing, Computer Vision, and recommendation systems. Paddle FL allowing the deployment of federated learning systems in the form of distributed clusters at a large scale has been proven to be of great benefit to developers. PaddleFL is an open-source framework based on PaddlePaddle.

FL methods, user-specified models and algorithms, distributed training setup, and FL task generator are all included in the compile time. The horizontal FL algorithms, such as FedAvg, are among the FL techniques. Users are allowed to create their models and training algorithms in addition to the FL techniques offered.

PaddleFL is still in its early stages of production, and the documentation and samples are lacking. It also lacks in performance with standardized benchmarks such as Model DNN (e.g., ResNet) and vertical federated learning. Paddle FL does not involve a flexible and generic API design with topology customization and flexible message flow.

5.6 Clara Training Framework

Clara Train SDK [26] is a domain-optimized developer software framework that consists of APIs for AI-Assisted Annotation. This allows any clinical viewer to be AI successful and allows a TensorFlow-based framework with pre-skilled models to begin AI improvement with strategies including Transfer Learning, Federated Learning, and AutoML.

Developers may use the Clara Train SDK's configurable MMAR (Medical Model ARchive) function to carry their models and components to conduct Federated Learning, as well as monitor whether the local training is run on a single GPU or multiple GPUs.

5.7 Fed ML

FedML [45] serves as a tool for federated learning as well as a forum for FL benchmarking. Its core structure is split into two layers as an FL system. On-device training for IoT and mobile computers, distributed computing, and single-machine emulation are all supported. FedML also embraces a variety of algorithms, prototypes, and databases for study variability (e.g., decentralized learning, vertical FL, and split learning).

FedML offers a simulation environment for a wide range of hardware specifications while supporting three computing paradigms [45]: standalone simulation, distributed computing, and on-device training.

5.8 IBM Federated Learning

A Python framework library for distributed machine learning processes in an enterprise environment [51]. IBM Federated Learning focuses on enterprise environments where safe rollout, failure tolerance, and fast model specification are critical; these must make use of existing machine learning libraries that enable enterprise users to access a robust collection of state-of-the-art algorithms without learning new languages. Enterprise professionals will be able to easily implement federated learning using IBM Federated Learning.

Apart from neural networks and decision trees, IBM Federated Learning facilitates the learning of a variety of machine learning models such as multi-class classification, regression, linear classifiers, and adaptation of XGBoost. IBM Federated Learning also gives you the ability to apply differential privacy to a variety of models, from basic Naive Bayes to more sophisticated differential privacy structures like those used in neural networks.

6 CONCLUSION & FUTURE DIRECTIONS

Federated learning is a learning paradigm where machine learning models are trained at the edge. It was originally designed for a variety of domains, including mobile and edge device use cases, but it has recently acquired popularity in healthcare applications.

The development of federated learning systems for healthcare has sparked a lot of interest from both industry and academics. As a result, a comprehensive overview and summary of existing FLSs in the healthcare domain is required.

Not all technological concerns have been solved, and FL will undoubtedly be a focus of study in the coming decade. Even though 5G networks are not yet widely available and commercially implemented worldwide, significant research and development efforts have been directed toward future 6G wireless systems [29].

Future research will focus on incorporating FL functionalities into future 5G/6G medical devices, how to use 6G devices, such as intelligent implants and wearables, for large-scale FL-based healthcare, and what new healthcare services 6G enables. Future e-health services, for example, will be enhanced by AI and FL capabilities, improving patient quality of life and lowering hospitalizations [79].

In this article, we presented an overview of federated learning in the healthcare industry. We conducted a comprehensive survey of recent work in the healthcare sector, with a focus on federated settings. We also discussed how to employ federated learning in healthcare, as well as the methods, applications, and issues that come with it.

REFERENCES

[1] Nabil Adam, Tom White, Basit Shafiq, Jaideep Vaidya, and Xiaoyun He. 2007. Privacy preserving integration of health care data. *AMIA Annual Symposium Proceedings. AMIA Symposium* (10 2007), 1-5. https://europepmc.org/articles/PMC2655922.

[2] Mete Akgün, A. Osman Bayrak, Bugra Ozer, and M. Şamil Sağıroğlu. 2015. Privacy preserving processing of genomic data: A survey. *Journal of Biomedical Informatics* 56 (2015), 103–111. https://doi.org/10.1016/j.jbi.2015.05.022

[3] Abeer Y. Al-Hyari, Ahmad M. Al-Taee, and Majid A. Al-Taee. 2013. Clinical decision support system for diagnosis and management of chronic renal failure. In *2013 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT)*. IEEE, 1–6.

[4] Mohammad Al-Rubaie and J. Morris Chang. 2019. Privacy-preserving machine learning: Threats and solutions. *IEEE Security & Privacy* 17, 2 (3 2019). https://doi.org/10.1109/MSEC.2018.2888775

[5] Scott Alfeld, Xiaojin Zhu, and Paul Barford. 2016. Data poisoning attacks against autoregressive models. *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, 1452–1458.

[6] Mirza Mansoor Baig, Hamid Gholam Hosseini, and Maria Lindén. 2016. Machine learning-based clinical decision support system for early diagnosis from real-time physiological data. In *2016 IEEE Region 10 Conference (TENCON)*. IEEE, 2943–2946.

[7] N. Balachandar, Ken Chang, J. Kalpathy-Cramer, and D. Rubin. 2020. Accounting for data variability in multi-institutional distributed deep learning for medical imaging. *Journal of the American Medical Informatics Association: JAMIA* (2020).

[8] Sourasekhar Banerjee, Rajiv Misra, Mukesh Prasad, Erik Elmroth, and Monowar H. Bhuyan. 2020. Multi-diseases classification from chest-x-ray: A federated deep learning approach. (2020). https://doi.org/10.1007/978-3-030-64984-5_1

[9] Nuria Rodríguez Barroso, Goran Stipcich, Daniel Jiménez-López, José Antonio Ruiz-Millán, Eugenio Martínez-Cámara, Gerardo González-Seco, M. Victoria Luzón, Miguel Ángel Veganzones, and Francisco Herrera. 2020. Federated learning and differential privacy: Software tools analysis, the sherpa.ai FL framework and methodological guidelines for preserving data privacy. *CoRR* abs/2007.00914 (2020). arXiv:2007.00914. https://arxiv.org/abs/2007.00914.

[10] Sarah Beachy, Theresa M. Wizemann, and Meredith Hackmann. 2019. Exploring sources of variability related to the clinical translation of regenerative engineering products.

[11] Battista Biggio, Blaine Nelson, and Pavel Laskov. 2013. Poisoning Attacks against Support Vector Machines. arXiv:1206.6389 [cs.LG].

[12] Tiffani J. Bright, Anthony Wong, Ravi Dhurjati, Erin Bristow, Lori Bastian, Remy R. Coeytaux, Gregory Samsa, Vic Hasselblad, John W. Williams, Michael D. Musty, et al. 2012. Effect of clinical decision-support systems: A systematic review. *Annals of Internal Medicine* 157, 1 (2012), 29–43.

[13] Theodora S. Brisimi, R. Chen, T. Mela, A. Olshevsky, I. Paschalidis, and Wei Shi. 2018. Federated learning of predictive models from federated electronic health records. *International Journal of Medical Informatics* 112 (2018), 59–67.

[14] Eoin Brophy, M. Vos, Geraldine Boylan, and T. Ward. 2021. Estimation of continuous blood pressure from PPG via a federated learning approach. *ArXiv* abs/2102.12245 (2021).

[15] Robert Carlsson. 2020. Privacy-preserved federated learning: A survey of applicable machine learning algorithms in a federated environment.

[16] Dongchul Cha, MinDong Sung, and Yu-Rang Park. 2021. Implementing vertical federated learning using autoencoders: Practical application, generalizability, and utility study. *JMIR Medical Informatics* 9, 6 (2021). https://doi.org/10.2196/26598

[17] Di Chai, Leye Wang, Kai Chen, and Qiang Yang. 2019. Secure federated matrix factorization.

[18] Di Chai, Leye Wang, Kai Chen, and Qiang Yang. 2020. FedEval: A Benchmark System with a Comprehensive Evaluation Model for Federated Learning. arXiv:2011.09655 [cs.LG].

[19] Kamalika Chaudhuri, A. Sarwate, and Kaushik Sinha. 2013. A near-optimal algorithm for differentially-private principal components. *ArXiv* abs/1207.2812 (2013).

[20] Yang Chen, Xiaoyan Sun, and Yaochu Jin. 2020. Communication-efficient federated deep learning with layerwise asynchronous model update and temporally weighted aggregation. *IEEE Transactions on Neural Networks and Learning Systems* 31, 10 (2020). https://doi.org/10.1109/TNNLS.2019.2953131

[21] Yiqiang Chen, Jindong Wang, Chaohui Yu, Wen Gao, and Xin Qin. 2021. FedHealth: A federated transfer learning framework for wearable healthcare. (2021). arXiv:1907.09173 [cs.LG].

[22] Kewei Cheng, Tao Fan, Yilun Jin, Yang Liu, Tianjian Chen, and Qiang Yang. 2019. SecureBoost: A lossless federated learning framework. *ArXiv* abs/1901.08755 (2019).

[23] Prateek Chhikara, Prabhjot Singh, Rajkumar Tekchandani, Neeraj Kumar, and Mohsen Guizani. 2021. Federated learning meets human emotions: A decentralized framework for human-computer interaction for IoT applications. *IEEE Internet of Things Journal* 8, 8 (4 2021). https://doi.org/10.1109/JIOT.2020.3037207

[24] Olivia Choudhury, Yoonyoung Park, Theodoros Salonidis, A. Gkoulalas-Divanis, I. Sylla, and Amar K. Das. 2019. Predicting adverse drug reactions on distributed health data using federated learning. *AMIA Annual Symposium Proceedings. AMIA Symposium* 2019 (2019), 313–322.

[25] B. Claerhout, D. Kalra, C. Mueller, Gurparkash Singh, N. Ammour, L. Meloni, J. Blomster, M. Hopley, G. Kafatos, Almenia Garvey, Peter Kuhn, Martine Lewi, B. Vannieuwenhuyse, Benoît Marchal, K. Mayer-Patel, Christoph Schindler, and M. Sundgren. 2019. Federated electronic health records research technology to support clinical trial protocol optimization: Evidence from EHR4CR and the InSite platform. *Journal of Biomedical Informatics* 90 (2019), 103090.

[26] Clara. 2019. NVIDIA clara. https://docs.nvidia.com/clara/. (2019).

[27] Scott D. Constable, Y. Tang, Shuang Wang, Xiaoqian Jiang, and S. Chapin. 2015. Privacy-preserving GWAS analysis on federated genomic datasets. *BMC Medical Informatics and Decision Making* 15 (2015), S2–S2.

[28] Jianfei Cui, He Zhu, Hao Deng, Ziwei Chen, and Dianbo Liu. 2021. FeARH: Federated machine learning with anonymous random hybridization on electronic medical records. *Journal of Biomedical Informatics* 117 (5 2021). https://doi.org/10.1016/j.jbi.2021.103735

[29] Chamitha de Alwis, Anshuman Kalla, Quoc-Viet Pham, Pardeep Kumar, Kapal Dev, Won-Joo Hwang, and Madhusanka Liyanage. 2021. Survey on 6G frontiers: Trends, applications, requirements, technologies and future research. *IEEE Open Journal of the Communications Society* 2 (2021), 836–886.

[30] Jeffrey Dean, Greg Corrado, Rajat Monga, Kai Chen, Matthieu Devin, Mark Mao, Marc Aurelio Ranzato, Andrew Senior, Paul Tucker, Ke Yang, Quoc Le, and Andrew Ng. 2012. Large scale distributed deep networks, F. Pereira, C. J. C. Burges, L. Bottou,

and K. Q. Weinberger (Eds.). *Advances in Neural Information Processing Systems* 25. https://proceedings.neurips.cc/paper/2012/file/6aca97005c68f1206823815f66102863-Paper.pdf.

[31] Timo M. Deist, Frank J. W. M. Dankers, Priyanka Ojha, M. Scott Marshall, Tomas Janssen, Corinne Faivre-Finn, Carlotta Masciocchi, Vincenzo Valentini, Jiazhou Wang, Jiayan Chen, Zhen Zhang, Emiliano Spezi, Mick Button, Joost Jan Nuyttens, René Vernhout, Johan van Soest, Arthur Jochems, René Monshouwer, Johan Bussink, Gareth Price, Philippe Lambin, and Andre Dekker. 2020. Distributed learning on 20 000+ lung cancer patients - the personal health train. *Radiotherapy and Oncology* 144 (3 2020). https://doi.org/10.1016/j.radonc.2019.11.019

[32] Qi Dou, Tiffany Y. So, Meirui Jiang, Quande Liu, Varut Vardhanabhuti, Georgios Kaissis, Zeju Li, Weixin Si, Heather H. C. Lee, Kevin Yu, Zuxin Feng, Li Dong, Egon Burian, Friederike Jungmann, Rickmer Braren, Marcus Makowski, Bernhard Kainz, Daniel Rueckert, Ben Glocker, Simon C. H. Yu, and Pheng Ann Heng. 2021. Federated deep learning for detecting COVID-19 lung abnormalities in CT: A privacy-preserving multinational validation study. *npj Digital Medicine* 4, 1 (12 2021). https://doi.org/10.1038/s41746-021-00431-6

[33] Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. 2006. Our data, ourselves: Privacy via distributed noise generation. (2006). https://doi.org/10.1007/11761679_29

[34] Cynthia Dwork, Guy N. Rothblum, and Salil Vadhan. 2010. Boosting and differential privacy. *2010 IEEE 51st Annual Symposium on Foundations of Computer Science.* https://doi.org/10.1109/FOCS.2010.12

[35] Fate. 2019. Federated AI technology enabler (FATE). https://github.com/FederatedAI/FATE.

[36] Francisco Fernández-Avilés, Ricardo Sanz-Ruiz, Andreu M. Climent, Lina Badimón, R. Bolli, Dominique Charron, Valentin Fuster, Stefan Janssens, Jens Kastrup, Hyo-Soo Kim, Thomas Felix Lüscher, John F. Martin, Philippe Menasché, Robert D. Simari, Gregg W. Stone, Andre Terzic, James T. Willerson, Joseph C. Wu, Francisco Andre Fernández-Avilés Terzic, Lina Kathleen Darcy L. Stefanie Rosalinda Marc S. Mark Badimon Broughton DiFede Dimmeler Madonna Penn Sus, Kathleen M. Broughton, Darcy L. DiFede, Stefanie Dimmeler, Rosalinda Madonna, Marc S. Penn, Mark A. Sussman, Joost P. G. Sluijter, Kai C. Wollert, Wayne Roberto Steven Dominique María Eugenia Valentin Ge Balkan Bolli Chamuleau Charron Fernández-Santos Fu, Wayne Balkan, Steven A. J. Chamuleau, María Eugenia Fernández-Santos, Georg Goliasch, Mariann Gyöngyösi, Joshua M. Hare, Bryon A. Tompkins, Johannes Winkler, Antoni Timothy D. Doris A. Bayés-Genís Henry Taylor, Antoni Bayés-Genís, Timothy D. Henry, Doris A. Taylor, Andreu M. Amir Beatriz Felipe Climent Lerman Pelacho Prosper, Amir Lerman, Beatriz Pelacho, Felipe Prosper, Ricardo Emerson C. Giulio Sanz-Ruiz Perin Pompilio, Emerson C. Perin, Giulio Pompilio, Bernard Jozef Eric Péter Stefan Douglas W. Pedro L. Warren Gersh Bartunek Duckers Ferdinandy Janssens Losordo, Bernard J. Gersh, Jozef Bartunek, Eric Duckers, Péter Ferdinandy, Douglas Losordo, Pedro L. Sánchez, Warren Sherman, Wojtek Wojakowski, Andreas M. Zeiher, Jérôme Roncalli, Anthony Mathur, Filippo Domenico Thomas J. Jay Seppo Crea D'Amario Povsic Traverse Ylä-Herttuala, Filippo Crea, Domenico D'Amario, Thomas J. Povsic, Jay H. Traverse, and Seppo Ylä-Herttuala. 2017. Global position paper on cardiovascular regenerative medicine. *European Heart Journal* 38 (2017), 2532–2546.

[37] Mona Flores, I. Dayan, H. Roth, Aoxiao Zhong, A. Harouni, Amilcare Gentili, A. Abidin, Andrew Liu, A. Costa, B. Wood, Chien-Sung Tsai, Chih-Hung Wang, C. Hsu, C. K. Lee, Colleen Ruan, Daguang Xu, Dufan Wu, E. Huang, F. Kitamura, G. Lacey, G. Corradi, Hao-Hsin Shin, Hirofumi Obinata, Hui Ren, Jason Crane, Jesse Tetreault, Jiahui Guan, J. Garrett, J. Park, K. Dreyer, K. Juluru, Kristopher Kersten, Marcio A. B. C. Rockenbach, M. Linguraru, M. Haider, M. Abdelmaseeh, Nicola Rieke, P. Damasceno, Pedro Silva, Pochuan Wang, Sheng Xu, Shuichi Kawano, Sira Sriswa, S. Park, T. Grist, V. Buch, W. Jantarabenjakul, Weichung Wang, W. Tak, Xiang Li, Xihong Lin, Fred Kwon, Fiona Gilbert, J. Kaggie, Quanzheng Li, Abood Quraini, Andrew Feng, A. Priest, B. Turkbey, B. Glicksberg, B. Bizzo, B. S. Kim, Carlos Tor-Díez, Chia-Cheng Lee, Chia-Jung Hsu, Chin-Hsien Lin, C. Lai, Christopher Hess, Colin B. Compas, D. Bhatia, E. Oermann, E. Leibovitz, H. Sasaki, Hitoshi Mori, Isaac Yang, J. H. Sohn, Krishna Nand Keshava Murthy, Lijuan Fu, Matheus Ribeiro Furtado de Mendon, Mike Fralick, M. Kang, M. Adil, Natalie Gangai, P. Vateekul, P. Elnajjar, Sara Hickman, S. Majumdar, S. McLeod, Sheridan Reed Stefan Graf, S. Harmon, T. Kodama, T. Puthanakit, T. Mazzulli, Vitor Lavor, Y. Rakvongthai, Yu Rim Lee, and Yuhong Wen. 2021. Federated learning used for predicting outcomes in SARS-COV-2 patients. *Research Square* (2021).

[38] Jonas Geiping, Hartmut Bauermeister, Hannah Dröge, and Michael Moeller. 2020. Inverting Gradients – How easy is it to break privacy in federated learning? arXiv:2003.14053 [cs.CV].

[39] B. Gibaud, G. Kassel, M. Dojat, B. Batrancourt, Franck Michel, A. Gaignard, and J. Montagnat. 2011. NeuroLOG: Sharing neuroimaging data using an ontology-based federated approach. *AMIA Annual Symposium Proceedings. AMIA Symposium* 2011 (2011), 472–80.

[40] Eren Gultepe, Jeffrey P. Green, Hien Nguyen, Jason Adams, Timothy Albertson, and Ilias Tagkopoulos. 2014. From vital signs to clinical outcomes for patients with sepsis: A machine learning basis for a clinical decision support system. *Journal of the American Medical Informatics Association* 21, 2 (2014), 315–325.

[41] Jiale Guo, Ziyao Liu, Kwok-Yan Lam, Jun Zhao, Yiqiang Chen, and Chaoping Xing. 2021. Secure weighted aggregation for federated learning.

[42] Andrew Hard, Kanishka Rao, Rajiv Mathews, Swaroop Ramaswamy, Françoise Beaufays, Sean Augenstein, Hubert Eichner, Chloé Kiddon, and Daniel Ramage. 2019. Federated learning for mobile keyboard prediction. (2019). arXiv:1811.03604 [cs.CL].

[43] Chaoyang He, Murali Annavaram, and Salman Avestimehr. 2020. FedNAS: Federated deep learning via neural architecture search. *CoRR* abs/2004.08546 (2020). arXiv:2004.08546 https://arxiv.org/abs/2004.08546.

[44] Chaoyang He, Murali Annavaram, and Salman Avestimehr. 2020. Group knowledge transfer: Collaborative training of large CNNs on the edge. *CoRR* abs/2007.14513 (2020). arXiv:2007.14513. https://arxiv.org/abs/2007.14513.

[45] Chaoyang He, Songze Li, Jinhyun So, Mi Zhang, Hongyi Wang, Xiaoyang Wang, Praneeth Vepakomma, Abhishek Singh, Hang Qiu, Li Shen, Peilin Zhao, Yan Kang, Yang Liu, Ramesh Raskar, Qiang Yang, Murali Annavaram, and Salman Avestimehr. 2020. FedML: A research library and benchmark for federated machine learning. *arXiv preprint arXiv:2007.13518* (2020).

[46] Chaoyang He, Conghui Tan, Hanlin Tang, Shuang Qiu, and Ji Liu. 2020. Central server free federated learning over single-sided trust social networks. (2020). arXiv:1910.04956 [cs.LG].

[47] A. Honkela, Mrinal Das, O. Dikmen, and Samuel Kaski. 2017. Efficient differentially private learning improves drug sensitivity prediction. *Biology Direct* 13 (2017).

[48] M. Shamim Hossain and Ghulam Muhammad. 2019. Emotion recognition using secure edge and cloud computing. *Information Sciences* 504 (12 2019). https://doi.org/10.1016/j.ins.2019.07.040

[49] Li Huang, Andrew L. Shea, Huining Qian, Aditya Masurkar, Hao Deng, and Dianbo Liu. 2019. Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records. *Journal of Biomedical Informatics* 99 (11 2019). https://doi.org/10.1016/j.jbi.2019.103291

[50] Li Huang, Yifeng Yin, Zeng Fu, Shifa Zhang, Hao Deng, and Dianbo Liu. 2020. LoAdaBoost: Loss-based AdaBoost federated machine learning with reduced computational complexity on IID and non-IID intensive care data. *PLOS ONE* 15, 4 (2020). https://doi.org/10.1371/journal.pone.0230706

[51] IBMFL. 2020. IBM federated learning. https://github.com/IBM/federated-learning-lib.

[52] Rebecca S. Jacobson, M. Becich, R. Bollag, G. Chavan, Julia Corrigan, R. Dhir, M. Feldman, Carmelo Gaudioso, Elizabeth Legowski, N. Maihle, K. Mitchell, Monica Murphy, Mayurapriyan Sakthivel, Eugene Tseytlin, and J. Weaver. 2015. A federated network for translational cancer research using clinical data and biospecimens. *Cancer Research* 75 24 (2015), 5194–201.

[53] Ji Chu Jiang, B. Kantarci, S. Oktug, and T. Soyata. 2020. Federated learning in smart city sensing: Challenges and opportunities. *Sensors (Basel, Switzerland)* 20 (2020).

[54] A. Jochems, T. Deist, I. El Naqa, M. Kessler, C. Mayo, J. Reeves, S. Jolly, M. Matuszak, R. Ten Haken, J. van Soest, C. Oberije, C. Faivre-Finn, G. Price, D. D. De Ruysscher, P. Lambin, and A. Dekker. 2017. Developing and validating a survival prediction model for NSCLC patients through distributed learning across 3 countries. *International Journal of Radiation Oncology, Biology, Physics* 99 (2017), 344–352.

[55] Arthur Jochems, Timo M. Deist, Johan van Soest, Michael Eble, Paul Bulens, Philippe Coucke, Wim Dries, Philippe Lambin, and Andre Dekker. 2016. Distributed learning: Developing a predictive model based on data from multiple hospitals without data leaving the hospital - a real life proof of concept. *Radiotherapy and Oncology* 121, 3 (12 2016). https://doi.org/10.1016/j.radonc.2016.10.002

[56] Ce Ju, Dashan Gao, Ravikiran Mane, Ben Tan, Yang Liu, and Cuntai Guan. 2020. Federated transfer learning for EEG signal classification. *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (Jul. 2020). https://doi.org/10.1109/embc44109.2020.9175344

[57] Ce Ju, Ruihui Zhao, Jichao Sun, Xiguang Wei, Bo Zhao, Yang Liu, Hongshan Li, Tianjian Chen, Xinwei Zhang, Dashan Gao, Ben Tan, Han Yu, Chuning He, and Yuan Jin. 2020. Privacy-preserving technology to help millions of people: Federated prediction model for stroke prevention.

[58] Peter Kairouz, H. Brendan McMahan, Brendan Avent, Aurélien Bellet, Mehdi Bennis, Arjun Nitin Bhagoji, Kallista Bonawitz, Zachary Charles, Graham Cormode, Rachel Cummings, Rafael G. L. D'Oliveira, Hubert Eichner, Salim El Rouayheb, David Evans, Josh Gardner, Zachary Garrett, Adrià Gascón, Badih Ghazi, Phillip B. Gibbons, Marco Gruteser, Zaid Harchaoui, Chaoyang He, Lie He, Zhouyuan Huo, Ben Hutchinson, Justin Hsu, Martin Jaggi, Tara Javidi, Gauri Joshi, Mikhail Khodak, Jakub Koneèný, Aleksandra Korolova, Farinaz Koushanfar, Sanmi Koyejo, Tancrède Lepoint, Yang Liu, Prateek Mittal, Mehryar Mohri, Richard Nock, Ayfer Özgür, Rasmus Pagh, Mariana Raykova, Hang Qi, Daniel Ramage, Ramesh Raskar, Dawn Song, Weikang Song, Sebastian U. Stich, Ziteng Sun, Ananda Theertha Suresh, Florian Tramèr, Praneeth Vepakomma, Jianyu Wang, Li Xiong, Zheng Xu, Qiang Yang, Felix X. Yu, Han Yu, and Sen Zhao. 2021. Advances and open problems in federated learning.

[59] Kensaku Kawamoto, Caitlin A. Houlihan, E. Andrew Balas, and David F. Lobach. 2005. Improving clinical practice using clinical decision support systems: A systematic review of trials to identify features critical to success. *BMJ: British Medical Journal* 330 (2005), 765.

[60] A. Kho, J. Cashy, K. Jackson, A. Pah, Satyender Goel, J. Boehnke, J. Humphries, S. Kominers, B. Hota, S. Sims, B. Malin, D. French, T. Walunas, D. Meltzer, E. Kaleba, R. C. Jones, and W. Galanter. 2015. Design and implementation of a privacy preserving electronic health record linkage tool in Chicago. *Journal of the American Medical Informatics Association: JAMIA* 22 5 (2015), 1072–80.

[61] Peeter Laud and Alisa Pankova. 2018. Privacy-preserving record linkage in large databases using secure multiparty computation. *BMC Medical Genomics* 11, S4 (10 2018). https://doi.org/10.1186/s12920-018-0400-8

[62] Leaf. 2019. Benchmarking framework for studying in federated settings. https://leaf.cmu.edu/.

[63] G. Lee and S. Shin. 2020. Federated learning on clinical benchmark data: Performance assessment. *Journal of Medical Internet Research* 22 (2020).

[64] Haeyun Lee, Young Jun Chai, Hyunjin Joo, Kyungsu Lee, Jae Youn Hwang, Seok-Mo Kim, Kwangsoon Kim, Inn-Chul Nam, June Young Choi, Hyeong Won Yu, Myung-Chul Lee, Hiroo Masuoka, Akira Miyauchi, Kyu Eun Lee, Sungwan Kim, and Hyoun-Joong Kong. 2021. Federated learning for thyroid ultrasound image analysis to protect personal information: Validation study in a real health care environment. *JMIR Medical Informatics* 9, 5 (2021). https://doi.org/10.2196/25869

[65] Qinbin Li, Zhaomin Wu, Zeyi Wen, and Bingsheng He. 2021. Privacy-preserving gradient boosting decision trees. (2021). arXiv:1911.04209 [cs.LG].

[66] Tian Li, Anit Kumar Sahu, Ameet S. Talwalkar, and Virginia Smith. 2020. Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine* 37 (2020), 50–60.

[67] Wen Li, Fucang Jia, and Qingmao Hu. 2015. Automatic segmentation of liver tumor in CT images with deep convolutional neural networks. *Journal of Computer and Communications* 3, 11 (2015). https://doi.org/10.4236/jcc.2015.311023

[68] Wenqi Li, Fausto Milletarì, Daguang Xu, Nicola Rieke, Jonny Hancox, Wentao Zhu, Maximilian Baust, Yan Cheng, Sébastien Ourselin, M. Jorge Cardoso, and Andrew Feng. 2019. Privacy-preserving federated brain tumour segmentation.

[69] Xiaoxiao Li, Yufeng Gu, N. Dvornek, L. Staib, P. Ventola, and J. Duncan. 2020. Multi-site fMRI analysis using privacy-preserving federated learning and domain adaptation: ABIDE results. *Medical Image Analysis* 65 (2020), 101765.

[70] Dianbo Liu, Timothy Miller, Raheel Sayeed, and Kenneth D. Mandl. 2018. FADL:Federated-Autonomous Deep Learning for Distributed Electronic Health Record. arXiv:1811.11400 [cs.CY].

[71] Jessica Chia Liu, Jack Goetz, Srijan Sen, and Ambuj Tewari. 2021. Learning from others without sacrificing privacy: Simulation comparing centralized and federated machine learning on mobile health data. *JMIR mHealth and uHealth* 9, 3 (2021). https://doi.org/10.2196/23728

[72] Yang Liu, Yingting Liu, Zhijie Liu, Yuxuan Liang, Chuishi Meng, Junbo Zhang, and Yu Zheng. 2020. Federated forest. *IEEE Transactions on Big Data* (2020). https://doi.org/10.1109/TBDATA.2020.2992755

[73] Sin Kit Lo, Qinghua Lu, Liming Zhu, Hye Young Paik, Xiwei Xu, and Chen Wang. 2021. Architectural Patterns for the Design of Federated Learning Systems. arXiv:2101.02373 [cs.LG].

[74] Lingjuan Lyu, Han Yu, and Qiang Yang. 2020. Threats to Federated Learning: A Survey. arXiv: 2003.02133 [cs.CR].

[75] H. Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Agüera y Arcas. 2017. Communication-efficient learning of deep networks from decentralized data.

[76] Bjoern H. Menze, Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, Keyvan Farahani, Justin Kirby, Yuliya Burren, Nicole Porz, Johannes Slotboom, Roland Wiest, Levente Lanczi, Elizabeth Gerstner, Marc-Andre Weber, Tal Arbel, Brian B. Avants, Nicholas Ayache, Patricia Buendia, D. Louis Collins, Nicolas Cordier, Jason J. Corso, Antonio Criminisi, Tilak Das, Herve Delingette, Cagatay Demiralp, Christopher R. Durst, Michel Dojat, Senan Doyle, Joana Festa, Florence Forbes, Ezequiel Geremia, Ben Glocker, Polina Golland, Xiaotao Guo, Andac Hamamci, Khan M. Iftekharuddin, Raj Jena, Nigel M. John, Ender Konukoglu, Danial Lashkari, Jose Antonio Mariz, Raphael Meier, Sergio Pereira, Doina Precup, Stephen J. Price, Tammy Riklin Raviv, Syed M. S. Reza, Michael Ryan, Duygu Sarikaya, Lawrence Schwartz, Hoo-Chang Shin, Jamie Shotton, Carlos A. Silva, Nuno Sousa, Nagesh K. Subbanna, Gabor Szekely, Thomas J. Taylor, Owen M. Thomas, Nicholas J. Tustison, Gozde Unal, Flor Vasseur, Max Wintermark, Dong Hye Ye, Liang Zhao, Binsheng Zhao, Darko Zikic, Marcel Prastawa, Mauricio Reyes, and Koen Van Leemput.2015. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Transactions on Medical Imaging* 34, 10 (2015). https://doi.org/10.1109/TMI.2014.2377694

[77] A Miyaji, Kazuhisa Nakasho, and Shohei Nishida. 2016. Privacy-preserving integration of medical data. *Journal of Medical Systems* 41 (2016).

[78] Atsuko Miyaji, Kazuhisa Nakasho, and Shohei Nishida. 2017. Privacy-preserving integration of medical data. *Journal of Medical Systems* 41, 3 (2017). https://doi.org/10.1007/s10916-016-0657-4

[79] Lorenzo Mucchi, Sara Jayousi, Stefano Caputo, Elisabetta Paoletti, Paolo Zoppi, Simona Geli, and Pietro Dioniso. 2020. How 6G technology can change the future wireless healthcare. *2020 2nd 6G Wireless Summit (6G SUMMIT)* (2020), 1–6.

[80] Dinh C. Nguyen, Quoc-Viet Pham, Pubudu N. Pathirana, Ming Ding, Aruna Prasad Seneviratne, Zihuai Lin, Octavia A. Dobre, and Won Joo Hwang. 2021. Federated learning for smart healthcare: A survey. arXiv: abs/2111.08834 (2021).

[81] PaddleFL. 2019. PaddleFL. https://github.com/PaddlePaddle/PaddleFL/.

[82] Ankit Pal and Malaikannan Sankarasubbu. 2021. Pay attention to the cough. *Proceedings of the 36th Annual ACM Symposium on Applied Computing*. https://doi.org/10.1145/3412841.3441943

[83] Ankit Pal, Logesh Kumar Umapathi, and Malaikannan Sankarasubbu. 2022. MedMCQA: A large-scale multi-subject multi-choice dataset for medical domain question answering. In *Proceedings of the Conference on Health, Inference, and Learning (Proceedings of Machine Learning Research, Vol. 174)*, Gerardo Flores, George H. Chen, Tom Pollard, Joyce C. Ho, and Tristan Naumann (Eds.). PMLR, 248–260. https://proceedings.mlr.press/v174/pal22a.html.

[84] Nicolas Papernot, Patrick McDaniel, Arunesh Sinha, and Michael Wellman. 2016. Towards the Science of Security and Privacy in Machine Learning. arXiv:1611.03814 [cs.CR].

[85] Marco V. Perez, Kenneth W. Mahaffey, Haley Hedlin, John S. Rumsfeld, Ariadna Garcia, Todd Ferris, Vidhya Balasubramanian, Andrea M. Russo, Amol Rajmane, Lauren Cheung, Grace Hung, Justin Lee, Peter Kowey, Nisha Talati, Divya Nag, Santosh E. Gummidipundi, Alexis Beatty, Mellanie True Hills, Sumbul Desai, Christopher B. Granger, Manisha Desai, and Mintu P. Turakhia. 2019. Large-scale assessment of a smartwatch to identify atrial fibrillation. *New England Journal of Medicine* 381, 20 (11 2019). https://doi.org/10.1056/NEJMoa1901183

[86] Emerson C. Perin, James T. Willerson, Carl J. Pepine, Timothy D. Henry, Stephen G. Ellis, David X. M. Zhao, Guilherme V. Silva, Dejian Lai, James D. Thomas, Marvin W. Kronenberg, A. Daniel Martin, R. David Anderson, Jay H. Traverse, Marc S. Penn, Saif

Anwaruddin, Antonis K. Hatzopoulos, Adrian P. Gee, Doris A. Taylor, Christopher R. Cogle, Deirdre Smith, Lynette Westbrook, James Chen, Eileen M. Handberg, Rachel E. Olson, Carrie Geither, Sherry Bowman, Judy Francescon, Sarah Baraniuk, Linda B. Piller, Lara M. Simpson, Catalin Loghin, David Aguilar, Sara Richman, Claudia Zierold, Judy Bettencourt, Shelly L. Sayre, Rachel W. Vojvodic, Sonia I. Skarlatos, David J. Gordon, Ray Francis Ebert, Minjung Kwak, Lemuel A. Moyé, and Robert D. Simari. 2012. Effect of transendocardial delivery of autologous bone marrow mononuclear cells on functional capacity, left ventricular function, and perfusion in chronic heart failure: The FOCUS-CCTRN trial. *JAMA* 307 16 (2012), 1717–26.

[87] Bjarne Pfitzner, Nico Steckhan, and Bert Arnrich. 2021. Federated learning in a medical context: A systematic literature review. *ACM Transactions on Internet Technology* 21, 2 (6 2021). https://doi.org/10.1145/3412357

[88] Tom Pollard, Irene Chen, Jenna Wiens, Steven Horng, Danny Wong, Marzyeh Ghassemi, Heather Mattie, Emily Lindmeer, and Trishan Panch. 2019. Turning the crank for machine learning: Ease, at what expense? *The Lancet Digital Health* 1 (09 2019), e198–e199. https://doi.org/10.1016/S2589-7500(19)30112-8

[89] Adnan Qayyum, Kashif Ahmad, Muhammad Ahtazaz Ahsan, Ala Al-Fuqaha, and Junaid Qadir. 2021. Collaborative federated learning for healthcare: Multi-modal COVID-19 diagnosis at the edge. arXiv: abs/2101.07511 (2021).

[90] Adnan Qayyum, Junaid Qadir, Muhammad Bilal, and Ala Al-Fuqaha. 2020. Secure and Robust Machine Learning for Healthcare: A Survey. arXiv:2001.08103 [cs.LG].

[91] Deepta Rajan, David Beymer, Shafiqul Abedin, and Ehsan Dehghan. 2019. Pi-PE: A pipeline for pulmonary embolism detection using sparsely annotated 3D CT images.

[92] Suraj Rajendran, J. Obeid, H. Binol, Ralph D. Agostino, K. Foley, Wei Zhang, P. Austin, Joey Brakefield, M. Gurcan, and U. Topaloglu. 2021. Cloud-based federated learning implementation across medical centers. *JCO Clinical Cancer Informatics* 5 (2021), 1–11.

[93] Nicola Rieke, Jonny Hancox, Wenqi Li, Fausto Milletarì, Holger R. Roth, Shadi Albarqouni, Spyridon Bakas, Mathieu N. Galtier, Bennett A. Landman, Klaus Maier-Hein, Sébastien Ourselin, Micah Sheller, Ronald M. Summers, Andrew Trask, Daguang Xu, Maximilian Baust, and M. Jorge Cardoso. 2020. The future of digital health with federated learning. *npj Digital Medicine* 3, 1 (12 2020). https://doi.org/10.1038/s41746-020-00323-1

[94] Holger R. Roth, Ken Chang, Praveer Singh, Nir Neumark, Wenqi Li, Vikash Gupta, Sharut Gupta, Liangqiong Qu, Alvin Ihsani, Bernardo C. Bizzo, Yuhong Wen, Varun Buch, Meesam Shah, Felipe Kitamura, Matheus Mendonça, Vitor Lavor, Ahmed Harouni, Colin Compas, Jesse Tetreault, Prerna Dogra, Yan Cheng, Selnur Erdal, Richard White, Behrooz Hashemian, Thomas Schultz, Miao Zhang, Adam McCarthy, B. Min Yun, Elshaimaa Sharaf, Katharina V. Hoebel, Jay B. Patel, Bryan Chen, Sean Ko, Evan Leibovitz, Etta D. Pisano, Laura Coombs, Daguang Xu, Keith J. Dreyer, Ittai Dayan, Ram C. Naidu, Mona Flores, Daniel Rubin, and Jayashree Kalpathy-Cramer. 2020. Federated learning for breast density classification: A real-world implementation. (2020). https://doi.org/10.1007/978-3-030-60548-3_18

[95] Theo Ryffel, Andrew Trask, Morten Dahl, Bobby Wagner, Jason Mancuso, Daniel Rueckert, and Jonathan Passerat-Palmbach. 2018. A generic framework for privacy preserving deep learning. arXiv:1811.04017 [cs.LG].

[96] Mustafa Abdul Salam, Sanaa Taha, and Mohamed Ramadan. 2021. COVID-19 detection using federated machine learning. *PLOS ONE* 16, 6 (2021). https://doi.org/10.1371/journal.pone.0252573

[97] Karthik V. Sarma, Stephanie Harmon, Thomas Sanford, Holger R. Roth, Ziyue Xu, Jesse Tetreault, Daguang Xu, Mona G. Flores, Alex G. Raman, Rushikesh Kulkarni, Bradford J. Wood, Peter L. Choyke, Alan M. Priester, Leonard S. Marks, Steven S. Raman, Dieter Enzmann, Baris Turkbey, William Speier, and Corey W. Arnold. 2021. Federated learning improves site performance in multicenter deep learning without data sharing. *Journal of the American Medical Informatics Association* 28, 6 (2021). https://doi.org/10.1093/jamia/ocaa341

[98] Jonas Scherer, Marco Nolden, Jens Kleesiek, Jasmin Metzger, Klaus Kades, Verena Schneider, Michael Bach, Oliver Sedlaczek, Andreas M. Bucher, Thomas J. Vogl, Frank Grünwald, Jens-Peter Kühn, Ralf-Thorsten Hoffmann, Jörg Kotzerke, Oliver Bethge, Lars Schimmöller, Gerald Antoch, Hans-Wilhelm Müller, Andreas Daul, Konstantin Nikolaou, Christian la Fougére, Wolfgang G. Kunz, Michael Ingrisch, Balthasar Schachtner, Jens Ricke, Peter Bartenstein, Felix Nensa, Alexander Radbruch, Lale Umutlu, Michael Forsting, Robert Seifert, Ken Herrmann, Philipp Mayer, Hans-Ulrich Kauczor, Tobias Penzkofer, Bernd Hamm, Winfried Brenner, Roman Kloeckner, Christoph Düber, Mathias Schreckenberger, Rickmer Braren, Georgios Kaissis, Marcus Makowski, Matthias Eiber, Andrei Gafita, Rupert Trager, Wolfgang A. Weber, Jakob Neubauer, Marco Reisert, Michael Bock, Fabian Bamberg, Jürgen Hennig, Philipp Tobias Meyer, Juri Ruf, Uwe Haberkorn, Stefan O. Schoenberg, Tristan Kuder, Peter Neher, Ralf Floca, Heinz-Peter Schlemmer,and Klaus Maier-Hein. 2020. Joint imaging platform for federated clinical data analytics. *JCO Clinical Cancer Informatics* 4 (11 2020). https://doi.org/10.1200/CCI.20.00045

[99] Rulin Shao, Hongyu He, Ziwei Chen, Hui Liu, and Dianbo Liu. 2020. Stochastic channel-based federated learning with neural network pruning for medical data privacy preservation: Model development and experimental validation. *JMIR Formative Research* 4, 12 (2020). https://doi.org/10.2196/17265

[100] Micah J. Sheller, Brandon Edwards, G. Anthony Reina, Jason Martin, Sarthak Pati, Aikaterini Kotrotsou, Mikhail Milchenko, Weilin Xu, Daniel Marcus, Rivka R. Colen, and Spyridon Bakas. 2020. Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data. *Scientific Reports* 10, 1 (12 2020). https://doi.org/10.1038/s41598-020-69250-1

[101] Yuanming Shi, Kai Yang, Tao Jiang, Jun Zhang, and Khaled B. Letaief. 2020. Communication-Efficient Edge AI: Algorithms and Systems. arXiv:2002.09668 [cs.IT].

[102] Santiago Silva, Boris Gutman, Eduardo Romero, Paul M. Thompson, Andre Altmann, and Marco Lorenzi. 2019. Federated learning in distributed medical databases: Meta-analysis of large-scale subcortical brain data.

[103] Johan Soest, Chang Sun, Ole Mussmann, Marco Puts, Bob Berg, Alexander Malic, Claudia Oppen, David Towend, André Dekker, and Michel Dumontier. 2018. Using the personal health train for automated and privacy-preserving analytics on vertically partitioned data. *Studies in Health Technology and Informatics* 247 (7 2018), 581–585.

[104] Peter D. Sottile, David Albers, Peter E. DeWitt, Seth Russell, J. N. Stroh, David P. Kao, Bonnie Adrian, Matthew E. Levine, Ryan Mooney, Lenny Larchick, Jean S. Kutner, Matthew K. Wynia, Jeffrey J. Glasheen, and Tellen D. Bennett. 2021. Real-time electronic health record mortality prediction during the COVID-19 pandemic: A prospective cohort study. *Journal of the American Medical Informatics Association* 28, 11 (10 2021). https://doi.org/10.1093/jamia/ocab100

[105] Reed Taylor Sutton, David Pincock, Daniel C. Baumgart, Daniel C. Sadowski, Richard N. Fedorak, and Karen I. Kroeker. 2020. An overview of clinical decision support systems: Benefits, risks, and strategies for success. *NPJ Digital Medicine* 3 (2020).

[106] Christian Szegedy, Wojciech Zaremba, Ilya Sutskever, Joan Bruna, Dumitru Erhan, Ian Goodfellow, and Rob Fergus. 2014. Intriguing properties of neural networks. arXiv:1312.6199 [cs.CV].

[107] TFL. 2019. TensorFlow federated learning. https://www.tensorflow.org/federated/federatedlearning/.

[108] Chu Myaet Thwal, Kyi Thar, Ye Lin Tun, and Choong Seon Hong. 2021. Attention on personalized clinical decision support system: Federated learning approach. *2021 IEEE International Conference on Big Data and Smart Computing (BigComp)* (2021), 141–147.

[109] Akhil Vaid, Suraj K. Jaladanki, Jie Xu, Shelly Teng, Arvind Kumar, Samuel Lee, Sulaiman Somani, Ishan Paranjpe, Jessica K. De Freitas, Tingyi Wanyan, Kipp W. Johnson, Mesude Bicak, Eyal Klang, Young Joon Kwon, Anthony Costa, Shan Zhao, Riccardo Miotto, Alexander W. Charney, Erwin Böttinger, Zahi A. Fayad, Girish N. Nadkarni, Fei Wang, and Benjamin S. Glicksberg. 2021. Federated learning of electronic health records to improve mortality prediction in hospitalized patients with COVID-19: Machine learning approach. *JMIR Medical Informatics* 9, 1 (2021). https://doi.org/10.2196/24207

[110] Gilmer Valdes, Charles B. Simone II, Josephine Chen, Alexander Lin, Sue S. Yom, Adam J. Pattison, Colin M. Carpenter, and Timothy D. Solberg. 2017. Clinical decision support of radiotherapy treatment planning: A data-driven machine learning strategy for patient-specific dosimetric decision making. *Radiotherapy and Oncology* 125, 3 (2017), 392–397.

[111] Praneeth Vepakomma, Tristan Swedish, Ramesh Raskar, Otkrist Gupta, and Abhimanyu Dubey. 2018. No peek: A survey of private distributed deep learning.

[112] Fei Wang and Anita Preininger. 2019. AI in Health: State of the Art, Challenges, and Future Directions. 016–026 pages. https://doi.org/10.1055/s-0039-1677908

[113] Hongyi Wang, Mikhail Yurochkin, Yuekai Sun, Dimitris Papailiopoulos, and Yasaman Khazaeni. 2020. Federated learning with matched averaging.

[114] Kangkang Wang, Rajiv Mathews, Chloé Kiddon, Hubert Eichner, Françoise Beaufays, and Daniel Ramage. 2019. Federated evaluation of on-device personalization. (2019). arXiv:1910.10252 [cs.LG].

[115] Qiong Wu, Xu Chen, Zhi Zhou, and Junshan Zhang. 2020. FedHome: Cloud-edge based personalized federated learning for in-home health monitoring. *IEEE Transactions on Mobile Computing* (2020). https://doi.org/10.1109/TMC.2020.3045266

[116] Xing Wu, Zhaowang Liang, and Jianjia Wang. 2020. FedMed: A federated learning framework for language modeling. *Sensors (Basel, Switzerland)* 20 (2020).

[117] X. Wu, Hao Zheng, Zuochao Dou, F. Chen, Jieren Deng, Xiang Chen, Shengqian Xu, Guanmin Gao, M. Li, Z. Wang, Yuhui Xiao, Kang Xie, Shuang Wang, and Huji Xu. 2021. A novel privacy-preserving federated genome-wide association study framework and its application in identifying potential risk variants in ankylosing spondylitis. *Briefings in Bioinformatics* (2021).

[118] Zhaoping Xiong, Ziqiang Cheng, Chi Xu, Xinyuan Lin, Xiaohong Liu, Dingyan Wang, Xiaomin Luo, Y. Zhang, Nan Qiao, M. Zheng, and Hualiang Jiang. 2020. Facing small and biased data dilemma in drug discovery with federated learning. *bioRxiv* (2020).

[119] Jie Xu, Benjamin S. Glicksberg, Chang Su, Peter Walker, Jiang Bian, and Fei Wang. 2021. Federated learning for healthcare informatics. *Journal of Healthcare Informatics Research* 5, 1 (3 2021). https://doi.org/10.1007/s41666-020-00082-4

[120] Runhua Xu, Nathalie Baracaldo, Yi Zhou, Ali Anwar, James Joshi, and Heiko Ludwig. 2021. FedV: Privacy-Preserving Federated Learning over Vertically Partitioned Data. arXiv:2103.03918 [cs.LG].

[121] Zeyue Xue, Pan Zhou, Zichuan Xu, Xiumin Wang, Yulai Xie, Xiaofeng Ding, and Shiping Wen. 2021. A resource-constrained and privacy-preserving edge-computing-enabled clinical decision system: A federated reinforcement learning approach. *IEEE Internet of Things Journal* 8, 11 (6 2021). https://doi.org/10.1109/JIOT.2021.3057653

[122] Zengqiang Yan, Jeffry Wicaksana, Zhiwei Wang, Xin Yang, and Kwang-Ting Cheng. 2021. Variation-aware federated learning with multi-source decentralized medical image data. *IEEE Journal of Biomedical and Health Informatics* 25, 7 (2021). https://doi.org/10.1109/JBHI.2020.3040015

[123] Dong Yang, Ziyue Xu, Wenqi Li, Andriy Myronenko, Holger R. Roth, Stephanie Harmon, Sheng Xu, Baris Turkbey, Evrim Turkbey, Xiaosong Wang, Wentao Zhu, Gianpaolo Carrafiello, Francesca Patella, Maurizio Cariati, Hirofumi Obinata, Hitoshi Mori, Kaku Tamura, Peng An, Bradford J. Wood, and Daguang Xu. 2021. Federated semi-supervised learning for COVID region segmentation in chest CT using multi-national data from China, Italy, Japan. *Medical Image Analysis* 70 (5 2021). https://doi.org/10.1016/j.media.2021.101992

[124] Qiang Yang, Yang Liu, Tianjian Chen, and Yongxin Tong. 2019. Federated machine learning: Concept and applications. *ACM Trans. Intell. Syst. Technol.* 10, 2, Article 12 (Jan. 2019), 19 pages. https://doi.org/10.1145/3298981

[125] Shengwen Yang, Bing Ren, Xuhui Zhou, and Liping Liu. 2019. Parallel distributed logistic regression for vertical federated learning without third-party coordinator.

[126] Timothy Yang, Galen Andrew, Hubert Eichner, Haicheng Sun, Wei Li, Nicholas Kong, Daniel Ramage, and Françoise Beaufays. 2018. Applied federated learning: Improving Google keyboard query suggestions.

[127] K. Y. Yigzaw, A. Budrionis, Luis Marco-Ruiz, Torje Dahle Henriksen, P. Halvorsen, and J. Bellika. 2020. Privacy-preserving architecture for providing feedback to clinicians on their clinical performance. *BMC Medical Informatics and Decision Making* 20 (2020).

[128] Kassaye Yitbarek Yigzaw, Antonis Michalas, and Johan Gustav Bellika. 2017. Secure and scalable deduplication of horizontally partitioned health data for privacy-preserving distributed statistical computation. *BMC Medical Informatics and Decision Making* 17, 1 (12 2017). https://doi.org/10.1186/s12911-016-0389-x

[129] Tao Yu, Eugene Bagdasaryan, and Vitaly Shmatikov. 2020. Salvaging federated learning by local adaptation. (2020). arXiv:2002.04758 [cs.LG].

[130] Lingchen Zhao, Lihao Ni, Shengshan Hu, Yaniiao Chen, Pan Zhou, Fu Xiao, and Libing Wu. 2018. InPrivate digging: Enabling tree-based distributed data mining with differential privacy. *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*. https://doi.org/10.1109/INFOCOM.2018.8486352

[131] Zibin Zheng, Shaoan Xie, Hong Ning Dai, Xiangping Chen, and Huaimin Wang. 2018. Blockchain challenges and opportunities: A survey. *International Journal of Web and Grid Services* 14, 4 (2018). https://doi.org/10.1504/IJWGS.2018.095647

[132] Hangyu Zhu, Haoyu Zhang, and Yaochu Jin. 2020. From federated learning to federated neural architecture search: A survey. (2020). arXiv:2009.05868 [cs.DC].

[133] Hangyu Zhu, Haoyu Zhang, and Yaochu Jin. 2021. From federated learning to federated neural architecture search: A survey. *Complex & Intelligent Systems* 7, 2 (4 2021). https://doi.org/10.1007/s40747-020-00247-z

Received 20 August 2021; revised 6 January 2022; accepted 25 April 2022