

Evaluating Open-Source Large Language Models for Educational Applications

Bhavisha Vishalbhai Parvadiya^{1,*}, Sahidsha Abdulhamid Diwan²

Abstract

Mobile healthcare technologies generate sensitive information through electronic health records, smartphones, wearable sensors, medical imaging systems, and remote monitoring platforms. These distributed data sources can improve disease prediction, continuous monitoring, clinical decision support, and personalized healthcare, but centralized collection creates substantial privacy, security, interoperability, and data-governance concerns. Federated learning (FL) provides a distributed machine learning paradigm in which participating healthcare institutions and personal devices collaboratively train models while retaining raw data at their original locations. This paper presents a focused review of FL for privacy-preserving mobile healthcare, emphasizing system architecture, learning mechanisms, healthcare applications, privacy protection, security threats, and deployment challenges. The discussion covers electronic health records, medical imaging, wearable sensing, remote patient monitoring, and multimodal healthcare. Federated Averaging (FedAvg), differential privacy, secure aggregation, homomorphic encryption, and trusted execution environments are examined as important components of privacy-aware FL systems. Recent research from 2024–2026 indicates a shift from proof-of-concept experiments toward multimodal learning, resource-aware architectures, secure Internet of Medical Things (IoMT) integration, and clinically oriented evaluation. The review also identifies persistent limitations involving non-IID data, device heterogeneity, communication overhead, model poisoning, privacy leakage, interoperability, fairness, explainability, and clinical validation. Based on these observations, the paper outlines research directions centered on adaptive aggregation, efficient edge training, personalized models, multimodal federation, robust security, standardized evaluation, and real-world clinical deployment.

Keywords: Federated learning, mobile healthcare, privacy preservation, Internet of Medical Things, wearable devices, electronic health records, secure aggregation, differential privacy, medical imaging

INTRODUCTION

Healthcare artificial intelligence increasingly depends on data generated outside a single clinical institution. Hospitals maintain electronic health records (EHRs), diagnostic centers produce medical images, laboratories maintain test results, and patients generate continuous physiological information through smartphones and wearable devices. Remote monitoring platforms further extend this ecosystem by collecting data from patients in homes and community settings. These sources provide opportunities for machine learning models to identify disease patterns, predict clinical events, and support personalized interventions.

*Author For Correspondence

Bhavisha Vishalbhai Parvadiya
E-mail: bhavishaparvadia@gmail.com

¹Assistant Professor, Sardar Patel College of Administration and Management, SPEC Junagadh, India

²Student, Sardar Patel College of Administration and Management, SPEC Campus, Bakrol, India

Received Date: August 29, 2026

Accepted Date: September 26, 2026

Published Date: September 28, 2026

Citation: Bhavisha Vishalbhai Parvadiya, Sahidsha Abdulhamid Diwan. Evaluating Open-Source Large Language Models for Educational Applications. Journal of Open Source Developments. 2026; 13(2): 10–21p.

The difficulty is that healthcare data are not ordinary machine learning datasets. Medical information can reveal diagnoses, medications, physiological characteristics, genetic information, or behavioral patterns. In addition, healthcare

organizations operate under institutional policies, ethical requirements, and legal frameworks that restrict how patient information can be transferred and reused. Consequently, combining data from multiple organizations through a conventional centralized repository can be technically and administratively difficult [1].

Federated learning (FL) addresses this problem by changing where model training occurs. Instead of sending raw data to a central server, participating clients train a shared model locally and transmit selected model information to an aggregation mechanism. The central system then combines the contributions and distributes an updated model for another training round [2]. The original FedAvg work demonstrated that decentralized training could operate effectively despite unbalanced and non-IID data and highlighted communication as a major constraint in distributed learning [2].

The healthcare relevance of FL has grown considerably. A systematic review published in 2024 analyzed 612 healthcare FL publications and found that most studies remained proof-of-concept experiments, with only a small proportion reporting real-life applications [3]. This observation is important because strong experimental accuracy alone does not establish that an FL system is ready for clinical use.

Recent work has expanded the scope of FL beyond conventional hospital-to-hospital collaboration. Digital healthcare surveys now cover Internet of Medical Things (IoMT) devices, wearable systems, communication efficiency, privacy protection, energy consumption, and model heterogeneity [4, 5]. Medical imaging has become another major research area because images are difficult to centralize across institutions, while multicenter training can improve the diversity of training data [6, 7]. More recent research has also moved toward multimodal healthcare systems combining imaging, EHRs, wearable information, and clinical decision support applications [8].

Therefore, the objective of this study is not simply to describe FL as a privacy-preserving alternative to centralized machine learning. Instead, this study examines FL as a distributed healthcare infrastructure in which learning, privacy, communication, device capability, security, and clinical requirements must operate together. The review addresses five questions:

1. How can FL architecture be adapted to mobile healthcare?
2. What role do aggregation and optimization methods play in decentralized training?
3. Which privacy and security mechanisms are appropriate for healthcare FL?
4. Which healthcare applications benefit from decentralized learning?
5. What technical and clinical barriers must be resolved for practical deployment?

The remainder of this paper develops these questions through an architecture-oriented discussion rather than a simple application catalogue.

FEDERATED MOBILE HEALTHCARE ARCHITECTURE

Distributed Healthcare Learning Environment

A mobile healthcare FL system generally contains three functional layers: data-producing clients, an orchestration layer, and healthcare AI services (Figure 1).

The client layer contains hospitals, clinics, smartphones, wearable devices, imaging systems, and IoMT equipment. Data remains at these locations. The orchestration layer manages model distribution, client selection, aggregation, security controls, and training rounds. The service layer contains the resulting models and healthcare applications.

This separation is useful because the data owner and model coordinator do not necessarily need to be the same entity. A hospital can maintain control of its clinical database while participating in a multicenter learning process. Similarly, a patient's wearable device can locally process sensor information without transmitting the complete time series to a remote server.

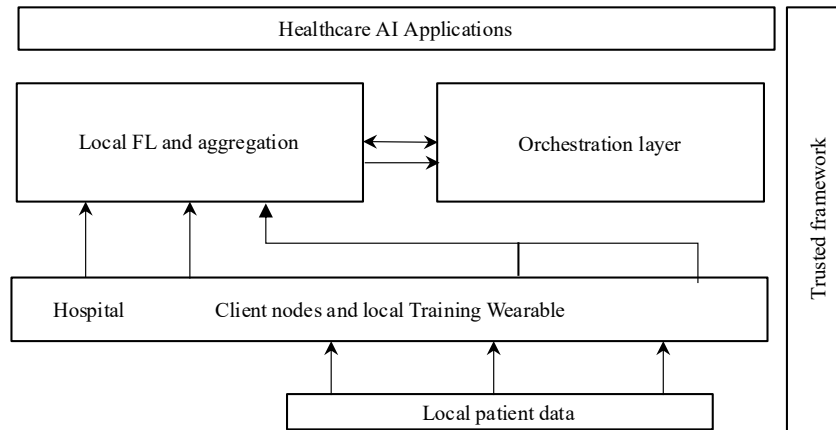


Figure 1. Layered architecture of privacy-preserving federated learning for mobile healthcare [9].

Architecture Illustrates an Important Principle

Raw healthcare data and model collaboration are separated. This separation reduces direct data movement but does not eliminate privacy risks, because model parameters and updates can contain information about local training data [5, 10].

Local Training and Global Aggregation

The basic FL cycle begins when the coordinating server distributes a global model to selected clients. Each client performs local optimization using its own data. The locally trained parameters are returned to the aggregation process, after which a new global model is generated.

The standard Federated Averaging formulation is:

$$W_{t+1} = \sum_{k=1}^K \frac{n_k}{N} W_{t+1}^k$$

Where,

$$N = \sum_{k=1}^K n_k$$

Here, W_{t+1} is the global model at round $t+1$, W_{t+1}^k is the locally trained model from client k , n_k is the number of local training samples at client k , and K represents the number of participating clients.

The coefficient

$$\frac{n_k}{N}$$

determines the relative contribution of client k . Thus, the standard FedAvg mechanism gives greater weight to clients containing more training samples [2].

However, in healthcare, sample count is not always an adequate indicator of contribution. A hospital with a smaller dataset may contain rare diseases or underrepresented patient populations. Similarly, a wearable device may provide fewer samples than a hospital but may contain high-frequency longitudinal information. Consequently, sample-weighted aggregation can produce undesirable effects when healthcare distributions are highly heterogeneous [3, 5].

This motivates adaptive aggregation methods that consider client reliability, distributional differences, model quality, fairness, and computational capability.

DATA AND SYSTEM HETEROGENEITY

Healthcare FL differs from many conventional distributed learning environments because heterogeneity exists simultaneously at the data, device, network, and institutional levels.

Statistical Heterogeneity

Patients do not generate identically distributed data. A cardiology hospital, pediatric clinic, oncology center, and general hospital may have very different patient populations. Even within a single institution, patients differ in age, disease status, medication, lifestyle, and treatment response.

This non-IID structure can cause local models to move in different directions during training. Therefore, a global model may converge more slowly or provide uneven performance across sites [2, 3].

Device Heterogeneity

Mobile healthcare introduces substantial hardware variations. A hospital server may have high-performance graphics processing units (GPUs), whereas a wearable sensor may have limited processing capability and battery capacity.

A practical FL system should therefore avoid assuming that every client can perform the same number of local training operations. Client selection and adaptive local computation can be used to accommodate resource differences.

Network Heterogeneity

Communication conditions also vary. Hospital networks may provide reliable high-bandwidth connections, whereas wearable devices may depend on intermittent wireless connectivity.

Because FL requires repeated communication rounds, network conditions can determine the total training time and energy consumption. The original FedAvg study identified communication as a central constraint in federated training [2].

Institutional Heterogeneity

Healthcare institutions may use different EHR schemas, terminology systems, imaging protocols, data preprocessing pipelines, and clinical workflows. Even if two institutions use the same machine learning model, their input data may not have identical meaning or structure.

Recent research on EHR harmonization identifies differences in healthcare data representation as a major obstacle to multi-partner FL [11].

PRIVACY AND SECURITY MECHANISMS

Keeping patient records locally is an important privacy measure; however, FL should not be described as inherently private. The transmitted model updates may reveal information about local datasets, and malicious participants can attempt to manipulate the training process.

Differential Privacy

Differential privacy (DP) provides a mathematical framework for limiting the influence of an individual record on the released result. In machine learning, noise can be added to gradients or model updates to reduce the possibility of reconstructing sensitive information [12].

The privacy parameter is commonly represented by ϵ . Lower values of ϵ generally correspond to stronger privacy guarantees, although the precise privacy-utility relationship depends on the mechanism and composition of multiple training steps.

In healthcare, DP must therefore be evaluated together with predictive performance. Excessive noise can reduce diagnostic or monitoring accuracy, whereas weak privacy protection may not adequately address the threat model.

Secure Aggregation

Secure aggregation allows a server to obtain an aggregate of client updates without directly seeing each individual update. Bonawitz et al. developed a practical protocol designed for high-dimensional user-held data and demonstrated its relevance to FL [13].

This mechanism is particularly useful in healthcare because an institution should not necessarily have access to the detailed updates generated by another institution or patient device.

Homomorphic Encryption

Homomorphic encryption allows computations to be performed on encrypted information. In FL, it can protect model information during the selected stages of aggregation.

Its principal limitation is its computational cost. Healthcare FL involving large neural networks and resource-constrained mobile devices must therefore consider whether the additional encryption overhead is justified by the required security level.

Trusted Execution Environments

Trusted execution environments provide hardware-supported protected regions for sensitive computations. They can be used to protect aggregation or model-processing operations from unauthorized access.

TEE-based approaches can reduce dependence on computationally expensive cryptographic operations in some architectures; however, they introduce reliance on trusted hardware and implementation assumptions [5].

Hybrid Privacy Protection

Recent systematic reviews have shown increasing interest in combining DP, secure multiparty computation, homomorphic encryption, TEEs, zero-knowledge mechanisms, and blockchain-based controls [5, 10].

The appropriate combination depends on the application. A wearable monitoring system may prioritize energy efficiency and secure aggregation, whereas a multi-hospital imaging system may justify more computationally intensive protection.

HEALTHCARE APPLICATIONS

Electronic Health Records

EHRs contain structured and unstructured information, including diagnoses, laboratory results, medications, clinical observations, and demographic information. FL allows several institutions to train predictive models while retaining their original records locally (Figure 2).

Possible applications include:

- disease-risk prediction;
- readmission prediction;
- patient deterioration detection;
- treatment-outcome prediction; and
- clinical decision support.

The main difficulty is that EHRs are not standardized automatically across organizations. Data harmonization and common representations are necessary before meaningful collaborative learning can occur [11].

Medical Imaging

CT, X-ray, ultrasound, pathology, and other images contain sensitive patient information and may be difficult to transfer across institutions.

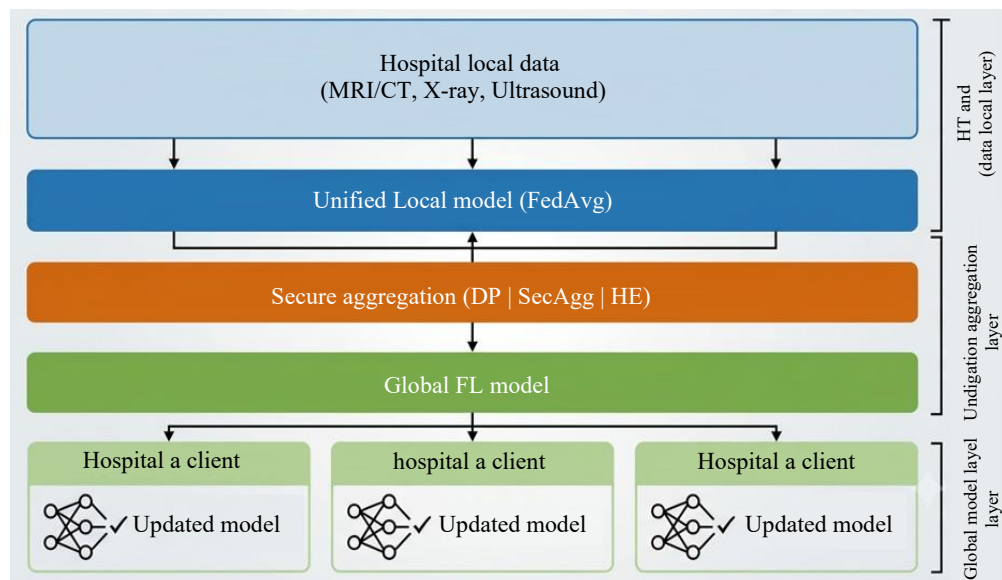


Figure 2. Federated workflow for medical imaging [7].

FL allows hospitals to train common segmentation, classification, or detection models while retaining the original images at each site [6, 7].

Recent medical imaging reviews have identified several persistent problems, including scanner variation, acquisition differences, non-IID data, privacy leakage, and limited external clinical validation [6, 7].

Wearable and Smartphone Monitoring

Wearables provide continuous physiological and behavioral measurements. Smartphones can collect activity, mobility, and application-derived health information.

An FL system can train models for activity recognition, sleep monitoring, cardiovascular monitoring, or abnormal-event detection without transferring the complete sensor stream.

Recent work has specifically addressed wearable health devices by combining FL with encrypted aggregation to reduce the exposure of sensitive sensor data while addressing communication overhead [14].

However, wearable systems require careful energy management. Frequent model updates can consume battery power and wireless bandwidth. Therefore, client participation should be selective rather than continuous.

Remote Patient Monitoring

Remote patient monitoring extends healthcare beyond hospitals. Patients can transmit measurements from connected devices while receiving continuous or periodic model-based assessments.

FL can allow models to learn across many patients without requiring their complete monitoring histories to be centralized. Recent IoMT research has explored federated frameworks for decentralized patient monitoring and data fusion [15, 9].

The major practical issue is reliability. A remote device may become disconnected, produce noisy measurements, or operate with limited computational capacity. Therefore, FL systems require mechanisms for incomplete participation and unreliable clients.

Multimodal Healthcare

Healthcare decisions increasingly depend on multiple information sources. A single patient may have EHR information, medical images, laboratory results, wearable measurements, and clinical notes. Multimodal FL attempts to learn from these heterogeneous modalities without centralizing all underlying data. Recent reviews have identified multimodal healthcare FL as an emerging area covering biomedical imaging, EHRs, wearables, and clinical decision support [8, 16].

However, multimodal FL is more complicated than single-modality learning. Different institutions may possess different combinations of modalities. One hospital may have imaging and EHR data, while another may have EHR and wearable data but no imaging. Future systems must therefore handle missing modalities and asynchronous information.

COMPARATIVE ANALYSIS OF APPLICATION AREAS

The comparison demonstrates that the privacy problem is shared across applications, but the engineering constraints are different. Consequently, healthcare FL should be designed around the characteristics of the target application rather than treated as a single standardized technology (Table 1).

SECURITY THREATS AND MITIGATION

Model Inference

Attackers may attempt to infer information from transmitted updates. Even when raw data remain local, model parameters can contain statistical information about the underlying records.

Secure aggregation and DP can reduce this risk, although neither mechanism should be considered universally sufficient.

Model Poisoning

A malicious client can send manipulated updates to influence the global model. In healthcare, this creates a particularly important safety concern because an attacker can attempt to alter a diagnostic model.

Robust aggregation, client authentication, anomaly detection, and update validation are possible mitigation strategies [5, 10, 17].

Membership Inference

Membership inference attempts to determine whether a particular patient's information was included in the model training. Such attacks are especially relevant when rare medical conditions are involved. Privacy mechanisms should therefore be selected according to the sensitivity of the application and the anticipated adversary.

Communication Attacks

Healthcare FL also requires protection during communication. Authentication, encryption, access control, secure channels, and appropriate key management are required to prevent unauthorized modification or interception of updates.

Table 1. Federated learning applications in mobile healthcare.

Application	Primary data	Typical objective	Main technical issue
EHR analytics	Clinical records	Risk and outcome prediction	Data harmonization
Medical imaging	CT, MRI, X-ray, pathology	Detection and segmentation	Site variability
Wearables	Physiological signals	Continuous monitoring	Energy consumption
Smartphones	Activity and behavioral data	Personalization	Device heterogeneity
Remote monitoring	Time-series measurements	Early warning	Connectivity
Multimodal systems	EHR, images, sensors	Integrated prediction	Missing modalities

Source: Synthesized from [3–8, 14, 15].

Table 2. Key performance dimensions for mobile healthcare FL.

Dimension	Measurement examples	Importance
Predictive performance	Accuracy, F1-score, AUROC, sensitivity, specificity	Clinical usefulness
Communication	Bytes transferred, rounds, bandwidth	Mobile scalability
Computation	Training time, FLOPs, memory	Device feasibility
Energy	Battery consumption, energy per round	Wearable deployment
Privacy	Privacy budget, attack success rate	Data protection
Robustness	Poisoning tolerance, failure tolerance	System security
Fairness	Subgroup/site performance	Equitable performance
Generalization	External-site performance	Clinical reliability

Source: Consolidated from [3–5, 7, 10].

PERFORMANCE AND RESOURCE CONSIDERATIONS

Privacy alone cannot determine whether a healthcare FL system is useful. A practical system must balance several dimensions.

A study reporting only predictive accuracy therefore provides an incomplete assessment of an FL system. For mobile healthcare, communication, energy, privacy, and cross-site generalization should be reported alongside conventional machine learning metrics (Table 2).

INTEROPERABILITY AND CLINICAL DEPLOYMENT

One of the most important differences between an experimental FL system and a deployable healthcare system is interoperability.

Two hospitals may store the same clinical concept using different terminology or coding systems. Imaging systems may use different acquisition protocols. Wearables may sample physiological signals at different frequencies. These differences can prevent direct federation, even when the institutions agree to collaborate.

Data harmonization is therefore an important prerequisite [11]. The solution does not necessarily require complete centralization. Instead, common schemas, standardized metadata, compatible preprocessing procedures, and agreed-upon outcome definitions can be implemented locally.

Clinical deployment introduces additional requirements. A model should be evaluated using data from institutions that were not involved in its development. This provides evidence of generalization. A 2024 systematic review of healthcare FL found that real-world deployment remained limited compared with proof-of-concept research [3].

This gap remains important in 2026. Recent reviews continue to identify reproducibility, clinical validation, interoperability, and scalability as unresolved issues [5, 6, 8].

COMPARISON OF PRIVACY MECHANISMS

A hybrid design is often more realistic than relying on a single technique. For example, secure aggregation can hide individual updates from the server, whereas DP can reduce information leakage from the aggregate model (Table 3).

RESEARCH DIRECTIONS

Adaptive Client Selection

Future systems can consider data diversity, battery level, network conditions, previous contribution, and reliability instead of selecting clients solely according to availability. This can reduce unnecessary communication while preserving representative participation [17, 16].

Table 3. Privacy-preserving methods for healthcare federated learning.

Method	Protection mechanism	Strength	Limitation
Differential Privacy	Adds controlled noise	Formal privacy framework	Accuracy/privacy trade-off
Secure Aggregation	Hides individual updates	Protects client contributions	Protocol overhead
Homomorphic Encryption	Computation on encrypted values	Strong confidentiality	High computation cost
TEE	Hardware-isolated processing	Efficient protected execution	Hardware dependency
SMPC	Joint secure computation	No single party receives complete input	Communication complexity
Blockchain	Distributed audit trail	Traceability	Scalability overhead

Source: Synthesized from [5, 10, 13].

Personalized Federated Learning

A single global model may not be equally appropriate for all patients. Personalized FL can maintain a shared representation while adapting selected model components locally.

This direction is particularly relevant to mobile healthcare because individual health conditions and behavioral patterns can vary substantially.

Communication-Efficient Learning

Communication remains a fundamental limitation of FL [2]. Future mobile systems can combine model compression, quantization, update sparsification, asynchronous participation, and adaptive communication intervals.

Robust and Trustworthy Aggregation

Aggregation algorithms should distinguish malicious updates from legitimate updates generated by naturally different healthcare populations. This is challenging because an unusual update is not necessarily an attack.

Future research should therefore combine robust statistics with contextual information about clients.

Multimodal Federation

Multimodal FL is likely to become increasingly important as healthcare systems integrate images, EHRs, sensors, and clinical information [8]. Research is needed for missing-modality scenarios, modality-specific privacy, and efficient cross-modal representation learning.

Explainability

Clinical users require understandable explanations for model predictions. Explainable FL should therefore provide useful information about model behavior without exposing patient information.

Standardized Benchmarks

The field needs standardized evaluation protocols. Studies should report not only accuracy but also privacy, communication, computation, energy, robustness, fairness, and external-site performance.

Clinical Validation

The transition from laboratory experiments to real clinical environments should become a central research objective. Multi-institutional prospective studies can provide stronger evidence than single-dataset simulations [9].

PROPOSED REFERENCE WORKFLOW

Based on the reviewed literature, a practical mobile healthcare FL workflow can be organized into seven stages:

1. *Local data preparation:* Each institution or device preprocesses its healthcare data locally.

2. *Model initialization*: A baseline model is distributed to selected participants.
3. *Local training*: Participants perform a predefined number of training epochs.
4. *Privacy protection*: Updates are clipped, encrypted, noised, or otherwise protected according to the threat model.
5. *Secure aggregation*: Protected updates are combined without unnecessarily exposing individual contributions.
6. *Global evaluation*: The updated model was evaluated using appropriate local and external validation procedures.
7. *Controlled redistribution*: The new model is distributed to the next set of participants.

This workflow emphasizes that FL is not simply a replacement for centralized training. It is a complete operational process that involves data preparation, distributed computation, privacy protection, aggregation, validation, and governance.

DISCUSSION

The literature demonstrates that FL provides a practical mechanism for collaborative healthcare AI; however, the phrase “privacy-preserving” should be interpreted carefully. FL primarily changes the location of raw data processing. Additional mechanisms are required to protect model updates, aggregation processes, and communication channels [5, 10].

The most important technical issue is heterogeneity. Healthcare participants are inherently different, and forcing them into an identical training configuration can reduce model effectiveness. This suggests that future systems should be more adaptive rather than simply increasing the complexity of the global model [18, 19].

Another important observation is the movement toward multimodal and edge-based healthcare. Recent work includes IoMT monitoring, wearable systems, medical imaging, EHRs, and multimodal clinical decision support [8, 14, 15]. These developments broaden the potential usefulness of FL but also make system design more complicated.

The evidence also suggests a gap between algorithmic research and clinical deployment. A 2024 systematic review reported that only a small proportion of healthcare FL research had reached real-world implementation [3]. More recent reviews continue to emphasize interoperability, reproducibility, secure aggregation, scalability, and clinical validation [5, 6].

Therefore, future research should evaluate FL systems as healthcare infrastructures rather than only as machine learning algorithms. A clinically useful system must be accurate, secure, efficient, interoperable, explainable, and maintainable.

CONCLUSION

FL provides a distributed approach for developing healthcare AI while reducing the need to centralize sensitive patient information. Its architecture is particularly relevant to mobile healthcare because information is naturally distributed across hospitals, clinics, smartphones, wearable devices, and Internet of Things (IoMT) platforms.

The review examined the underlying FL process, FedAvg aggregation, healthcare applications, privacy mechanisms, security threats, system heterogeneity, and deployment requirements. EHR analytics, medical imaging, wearable monitoring, remote patient monitoring, and multimodal healthcare represent important application areas.

However, FL should not be regarded as a complete privacy solution by itself. DP, secure aggregation, encryption, trusted execution, and other mechanisms may be required depending on the threat model. Similarly, technical performance cannot be judged solely through predictive accuracy. Communication overhead, energy consumption, robustness, fairness, interoperability, and external clinical validation must also be considered.

Recent 2025–2026 research indicates that the field is moving toward more comprehensive healthcare FL ecosystems involving IoMT, multimodal data, secure aggregation, privacy-aware architectures, and clinical translation. The next stage of research should therefore emphasize adaptive and personalized learning, efficient edge computation, robust aggregation, standardized evaluation, explainability, and multi-institutional clinical validation.

FL has considerable potential for privacy-aware mobile healthcare; however, its long-term value will depend on the ability to integrate machine learning performance with privacy engineering, healthcare interoperability, security, and clinical requirements.

REFERENCES

1. Rieke N, Hancox J, Li W, Milletari F, Roth HR, Albarqouni S, et al. The future of digital health with federated learning. *NPJ Digit Med*. 2020;3(1):119. doi:10.1038/s41746-020-00323-1. PMID:33015372.
2. McMahan B, Moore E, Ramage D, Hampson S, Aguera y Arcas BA. Communication-efficient learning of deep networks from decentralized data. 20th International Conference on Artificial Intelligence and Statistics (AISTATS), Fort Lauderdale, FL, USA. 2017. p.1273–82.
3. Teo ZL, Jin L, Li S, Miao D, Zhang X, Ng WY, et al. Federated machine learning in healthcare: a systematic review on clinical applications and technical architecture. *Cell Rep Med*. 2024;5(2):101419. doi:10.1016/j.xcrm.2024.101419. PMID:38340728.
4. Abbas SR, Abbas Z, Zahir A, Lee SW. Federated learning in smart healthcare: a comprehensive review on privacy, security, and predictive analytics with IoT integration. *Healthcare (Basel)*. 2024;12(24):2587. doi:10.3390/healthcare12242587. PMID:39766014.
5. Kanauzia R, Singh M, Gulati S, Singh BM, Arora N, Gola KK. A comprehensive survey on federated learning for privacy preservation in digital healthcare applications. *Knowl Inf Syst*. 2026;68:55. doi:10.1007/s10115-025-02673-2.
6. Ghosh D, Mehjabin M, Rayed ME, Mridha MF, Kabir MM. Advancements and challenges of federated learning in medical imaging: a systematic literature review. *Artif Intell Rev*. 2026;59(2):87. doi:10.1007/s10462-025-11489-z.
7. Guan H, Yap PT, Bozoki A, Liu M. Federated learning for medical image analysis: a survey. *Pattern Recognit*. 2024;151:110424. doi:10.1016/j.patcog.2024.110424. PMID:38559674.
8. Nemane S, Sarad VM, Tulaskar P, Marode TP, Bhangdiya V, Agrawal L, et al. Federated learning in multimodal healthcare diagnostics: privacy-preserving AI for biomedical imaging, electronic health records, wearables, and clinical decision support. *Arch Comput Methods Eng*. 2026:1–34. doi:10.1007/s11831-026-10629-0.
9. Nguyen DC, Pham QV, Pathirana PN, Ding M, Seneviratne A, Lin Z, et al. Federated learning for smart healthcare: a survey. *ACM Comput Surv*. 2022;55(3):1–37. doi:10.1145/3501296.
10. Kumar KAS, Nelson L, Jibinsingh BR. Systematic review of privacy-preserving federated learning in decentralized healthcare systems. *Franklin Open*. 2025;13:100440. doi:10.1016/j.fraope.2025.100440.
11. Lorenzer L, Schrempf M, Jauk S, Kalabakov S, Pierri G, Suzuki KMF, et al. Harmonization of electronic medical records for federated learning: addressing challenges in international healthcare collaborations. *Stud Health Technol Inform*. 2025;324:141–147. doi:10.3233/SHTI250175. PMID:40270399.
12. Abadi M, Chu A, Goodfellow IJ, McMahan HB, Mironov I, Talwar K, et al. Deep learning with differential privacy. 23rd ACM SIGSAC Conference on Computer and Communications Security (CCS), Vienna, Austria. 2016. p.308–18. doi:10.1145/2976749.2978318.
13. Bonawitz K, Ivanov V, Kreuter B, Marcedone A, McMahan HB, Patel S, et al. Practical secure aggregation for privacy-preserving machine learning. 24th ACM SIGSAC Conference on Computer and Communications Security (CCS), Dallas, TX, USA. 2017. p.1175–91. doi:10.1145/3133956.3133982.
14. Khan H, Kavati R, Pulkaram SS, Jalooli A. End-to-end privacy-aware federated learning for wearable health devices via encrypted aggregation in programmable networks. *Sensors (Basel)*. 2025;25(22):7023. doi:10.3390/s25227023. PMID:41305229.

-
15. Almogadwy B, Alqarafi A. Fused federated learning framework for secure and decentralized patient monitoring in healthcare 5.0 using IoMT. *Sci Rep*. 2025;15(1):24263. doi:10.1038/s41598-025-06574-w. PMID:40624105.
 16. Kairouz P, McMahan HB. Advances and open problems in federated learning. *Found Trends Mach Learn*. 2021;14(1-2):1–210. doi:10.1561/22000000083.
 17. Thrasher J, Devkota A, Siwakoti P, Chivukula R, Poudel P, Hu C, et al. Multimodal federated learning in healthcare: a review. *J Healthc Inform Res*. 2025;10:419–448. doi:10.1007/s41666-025-00226-4.
 18. Raza A, Guzzo A, Ianni M, Lappano R, Zanolini A, Maggiolini M, et al. Federated learning in radiomics: a comprehensive meta-survey on medical image analysis. *Comput Methods Programs Biomed*. 2025;267:108768. doi:10.1016/j.cmpb.2025.108768. PMID:40279838.
 19. Tanveer F, Iradat F, Iqbal W, Alsagri HS, Alhakbani HAA, Ahmad A, et al. Balancing privacy and performance in healthcare: a federated learning framework for sensitive data. *Digit Health*. 2025;11:20552076251381769. doi:10.1177/20552076251381769. PMID:41018509.