

A REVIEW OF DECENTRALIZED COLLABORATIVE MODEL TRAINING FOR MEDICAL DATA USING FEDERATED LEARNING WITH ENHANCED PRIVACY CONTROLS

Divyanshi Singh¹, Mr. Manish Kumar Soni²

¹Master of Technology, Computer Science and Engineering, Bansal Institute of Engineering & Technology, Lucknow, India

²Assistant Professor, Department of Computer Science and Engineering, Bansal Institute of Engineering & Technology, Lucknow, India

Abstract - The rapid digitization of healthcare systems has generated vast volumes of sensitive medical data, creating significant opportunities for advanced machine learning applications while simultaneously raising critical privacy concerns. Traditional centralized training approaches require data aggregation, which conflicts with regulatory frameworks and institutional data governance policies. Federated Learning (FL) has emerged as a promising paradigm that enables collaborative model training across distributed medical institutions without transferring raw patient data. This review systematically examines decentralized collaborative model training frameworks for healthcare applications, with a particular focus on enhanced privacy-preserving mechanisms integrated into FL systems. It analyzes architectural models, secure aggregation protocols, differential privacy techniques, homomorphic encryption schemes, and block chain-assisted decentralized infrastructures. The paper further evaluates performance trade-offs between model accuracy, communication efficiency, and privacy guarantees, highlighting challenges such as non-IID data distribution, adversarial attacks, scalability limitations, and regulatory compliance. By synthesizing current research trends and identifying persistent technical and ethical gaps, this review outlines future research directions aimed at achieving secure, scalable, and trustworthy decentralized medical AI systems. The study provides a structured foundation for researchers and practitioners developing privacy-aware federated learning frameworks in healthcare environments.

Key Words: Federated Learning; Decentralized Machine Learning; Medical Data Privacy; Differential Privacy; Secure Aggregation; Healthcare AI

1. INTRODUCTION

The integration of artificial intelligence (AI) into healthcare has transformed diagnostic systems, clinical decision support, medical imaging analytics, and predictive modeling. However, the effectiveness of machine learning (ML) models largely depends on access to large-scale, diverse, and high-quality datasets. In the medical domain, such data are inherently sensitive and governed by strict regulatory frameworks, making centralized data aggregation increasingly impractical. Federated Learning (FL) has

emerged as a distributed learning paradigm that enables collaborative model training without sharing raw data, thereby addressing privacy and compliance concerns while maintaining analytical utility (McMahan et al., 2017). This section contextualizes the motivation, challenges, and scope of decentralized collaborative training in medical environments.

1.1 Background and Motivation

Healthcare institutions generate heterogeneous data streams including electronic health records (EHRs), radiological images, genomic sequences, and biosignals. The application of deep learning techniques to such datasets has demonstrated superior performance in disease detection and outcome prediction compared to conventional statistical approaches (Esteva et al., 2017). Nevertheless, training robust models requires multi-institutional data collaboration to overcome local data bias and limited sample sizes.

Traditional centralized machine learning architectures require pooling data into a single repository, creating risks related to privacy breaches, data misuse, and regulatory non-compliance. Legislative frameworks such as GDPR and HIPAA impose strict limitations on cross-border and inter-institutional data sharing. Consequently, decentralized collaborative training has gained attention as a viable alternative that enables knowledge sharing without direct data exchange.

1.2 Challenges in Medical Data Sharing and Model Training

Medical data sharing is constrained by legal, ethical, and technical barriers. Privacy concerns arise due to the highly sensitive nature of patient information, where unauthorized access may lead to identity disclosure or discrimination. Even anonymized datasets remain vulnerable to re-identification attacks when combined with auxiliary information (Narayanan and Shmatikov, 2008).

From a technical standpoint, healthcare data are typically non-independent and identically distributed (non-IID), imbalanced, and institution-specific. Variations in imaging protocols, demographic distributions, and disease

prevalence reduce model generalizability. Furthermore, centralized infrastructures introduce single points of failure and increase susceptibility to cyber -attacks.

1.3 Role of Decentralized Learning and Federated Learning (FL)

1.3.1 Federated Learning Framework

Federated Learning is a distributed optimization framework in which participating clients train local models using private data and transmit only model updates to a coordinating server for aggregation. The foundational FedAvg algorithm demonstrated that decentralized stochastic gradient descent can achieve competitive performance compared to centralized training while reducing communication costs (McMahan et al., 2017).

In healthcare, FL enables hospitals to collaboratively develop diagnostic models for tasks such as tumor segmentation and disease prediction without exposing patient-level records (Rieke et al., 2020). This paradigm enhances data sovereignty while leveraging collective intelligence across institutions. Recent extensions incorporate peer-to-peer and block chain-based architectures to eliminate reliance on a central aggregator, thereby improving trust and fault tolerance.

1.4 Need for Enhanced Privacy Controls

Although FL prevents direct data sharing, it does not inherently guarantee complete privacy. Model updates can leak sensitive information through gradient inversion or membership inference attacks (Zhu et al., 2019). Therefore, enhanced privacy-preserving mechanisms are necessary to strengthen security guarantees.

Differential privacy introduces calibrated noise to model updates to limit individual data contribution, providing formal privacy bounds (Dwork, 2006). Secure multi-party computation and homomorphic encryption enable encrypted aggregation of model parameters without revealing intermediate updates. Secure aggregation protocols further ensure that the server cannot access individual client gradients. These techniques collectively mitigate adversarial risks while maintaining acceptable performance trade-offs.

2. FUNDAMENTALS

The deployment of advanced machine learning techniques in healthcare necessitates a clear understanding of computational paradigms, architectural models, and privacy-preserving mechanisms. This section outlines the foundational principles underpinning decentralized collaborative model training, particularly within federated learning (FL) environments designed for sensitive medical data.

2.1 Overview of Machine Learning in Healthcare

Machine learning (ML) has significantly transformed clinical diagnostics, disease prediction, medical imaging analysis, and personalized treatment planning. Supervised learning models, especially deep neural networks, have demonstrated high performance in radiology, dermatology, and pathology by leveraging large annotated datasets (Litjens et al., 2017). Similarly, predictive analytics applied to electronic health records (EHRs) enable early detection of adverse clinical events such as sepsis and hospital readmission.

Despite these advancements, medical ML systems face inherent constraints. Healthcare datasets are typically fragmented across institutions, heterogeneous in format, and subject to strict regulatory controls. The traditional assumption of centralized data availability is often unrealistic in clinical settings. Consequently, alternative learning paradigms that preserve data locality while enabling collaborative model improvement have gained increasing attention.

2.2 Centralized vs Decentralized Learning Paradigms

Centralized learning involves aggregating all training data into a single repository, where model optimization is performed on unified datasets. This approach simplifies coordination and often yields strong performance due to complete data visibility. However, it introduces significant privacy, security, and governance risks. Large centralized databases become attractive targets for cyber-attacks and may violate data protection regulations.

In contrast, decentralized learning distributes the training process across multiple nodes, each retaining its local data. Instead of transferring raw datasets, only intermediate model parameters or gradients are communicated. This paradigm enhances data sovereignty and reduces exposure to breaches. However, decentralized systems must address challenges such as communication overhead, synchronization complexity, and statistical heterogeneity (Kairouz et al., 2021). The trade-off between privacy preservation and system efficiency remains a central design consideration.

2.3 Federated Learning: Definition and Key Properties

2.3.1 Core Concept and Operational Mechanism

Federated Learning is a distributed optimization framework in which multiple clients collaboratively train a shared global model under the coordination of a central server or decentralized protocol. Each client performs local training using private data and periodically transmits model updates for aggregation. The canonical Federated Averaging (FedAvg) algorithm reduces communication rounds by

combining locally computed gradients through weighted averaging (McMahan et al., 2017).

Key properties of FL include data locality, communication efficiency, and iterative global aggregation. In healthcare contexts, this allows hospitals to jointly train predictive models while maintaining patient data within institutional boundaries. FL also supports cross-silo collaboration, where a limited number of reliable institutions participate, as opposed to cross-device settings common in mobile applications.

2.4 Privacy and Security Concepts in FL

Although FL mitigates direct data sharing risks, it does not inherently eliminate privacy vulnerabilities. Adversaries may reconstruct sensitive information from transmitted gradients or manipulate updates through poisoning attacks. To address these threats, several cryptographic and statistical techniques are integrated into federated systems.

2.4.1 Differential Privacy

Differential Privacy (DP) provides a mathematically rigorous framework for quantifying and limiting individual data contribution within a dataset. By injecting calibrated noise into gradients or model parameters, DP ensures that the inclusion or exclusion of a single record does not significantly influence the model output (Dwork, 2006). In FL, DP can be applied locally at client devices or globally during aggregation. While enhancing privacy guarantees, it introduces a trade-off between model accuracy and privacy budget (ϵ), requiring careful calibration in clinical applications.

2.4.2 Secure Multi-Party Computation

Secure Multi-Party Computation (SMPC) enables multiple participants to jointly compute a function over their inputs without revealing the inputs themselves. In federated settings, SMPC protocols facilitate secure aggregation of model updates, ensuring that the server cannot access individual gradients. Techniques such as secret sharing and cryptographic masking are commonly employed to achieve confidentiality during collaborative optimization (Bonawitz et al., 2017). SMPC strengthens trust in multi-institutional medical collaborations.

2.4.3 Homomorphic Encryption

Homomorphic Encryption (HE) allows computations to be performed directly on encrypted data without requiring decryption. In FL architectures, clients encrypt model updates before transmission, and aggregation is conducted in encrypted form. Only the final aggregated result is decrypted, preventing intermediate exposure of sensitive parameters. Although HE provides strong confidentiality guarantees, it is computationally intensive and may

introduce latency in large-scale healthcare systems (Gentry, 2009).

2.4.4 Trusted Execution Environments

Trusted Execution Environments (TEEs) are hardware-based secure enclaves that isolate sensitive computations from the main operating system. Within FL frameworks, TEEs can securely perform model aggregation while protecting against malicious server-side interference. Technologies such as Intel SGX create encrypted memory regions inaccessible to unauthorized processes. While TEEs reduce cryptographic overhead compared to fully homomorphic approaches, they rely on hardware trust assumptions and may be vulnerable to side-channel attacks (Costan and Devadas, 2016).

3. ARCHITECTURAL FRAMEWORKS FOR FEDERATED LEARNING IN MEDICAL SETTINGS

Federated learning (FL) architectures determine how model updates are exchanged, aggregated, and synchronized across participating medical institutions. In healthcare environments, architectural design must balance privacy preservation, computational efficiency, fault tolerance, and regulatory compliance. This section examines the principal architectural models adopted in federated medical AI systems.

3.1 Client-Server Federated Architecture

The client-server model represents the foundational architecture of federated learning. In this framework, a central coordinating server orchestrates training rounds by distributing a global model to participating hospitals or clinical centers (clients). Each client performs local optimization using its private dataset and returns model updates to the server for aggregation. The server computes a weighted average of updates and redistributes the improved model in iterative rounds (McMahan et al., 2017).

In medical applications, this architecture is particularly suited to cross-silo federated learning, where a limited number of trusted institutions collaborate. It offers structured coordination, simplified convergence control, and relatively low system complexity. However, the central server constitutes a single point of failure and may become a bottleneck in large-scale deployments. Additionally, despite secure aggregation mechanisms, trust assumptions remain concentrated around the coordinating entity, raising governance considerations in inter-hospital collaborations (Rieke et al., 2020).

3.2 Peer-to-Peer Decentralized Federated Models

Peer-to-peer (P2P) federated architectures eliminate the reliance on a central aggregator by enabling direct communication among participating nodes. In this decentralized topology, clients exchange model parameters

with neighboring nodes and iteratively update local models based on consensus algorithms. Distributed optimization techniques such as gossip protocols and decentralized stochastic gradient descent are typically employed.

Such architectures enhance fault tolerance and reduce centralized trust dependencies. In medical consortia where institutions seek equal authority in collaborative model governance, P2P systems promote fairness and transparency. However, convergence control becomes more complex, and communication overhead may increase due to multi-hop exchanges. Ensuring robustness against adversarial participants also requires stronger validation mechanisms compared to centralized setups (Lian et al., 2017).

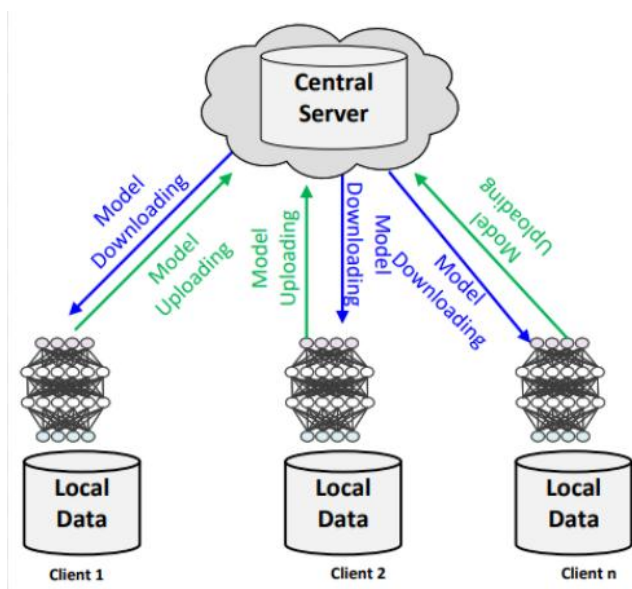


Figure-1: Decentralized Federated Models

3.3 Hybrid and Blockchain-Based Federated Systems

Hybrid federated architectures combine centralized coordination with decentralized trust mechanisms. One prominent approach integrates blockchain technology to maintain a tamper-resistant ledger of model updates and training transactions. Smart contracts can automate validation rules, enforce participation incentives, and ensure accountability across institutions.

Blockchain-assisted FL enhances transparency and mitigates risks associated with malicious parameter manipulation. Each update is cryptographically recorded, reducing the probability of undetected poisoning attacks. However, blockchain integration introduces computational latency, storage overhead, and scalability challenges, particularly when handling high-dimensional medical models (Li et al., 2020). Hybrid architectures therefore aim to balance coordination efficiency with decentralized trust reinforcement.

3.4 Communication Protocols and Synchronization

3.4.1 Federated Averaging (FedAvg)

FedAvg is the foundational aggregation protocol in federated learning. It reduces communication costs by allowing multiple local training epochs before transmitting updates to the server. The global model is computed as a weighted average of client parameters proportional to local dataset sizes (McMahan et al., 2017). In medical imaging and EHR-based prediction tasks, FedAvg demonstrates competitive accuracy while maintaining data locality. Nevertheless, its performance may degrade under non-independent and identically distributed (non-IID) data conditions, which are common in healthcare.

3.4.2 Federated Proximal (FedProx)

FedProx extends FedAvg by incorporating a proximal regularization term into the local objective function, constraining client updates to remain closer to the global model. This modification improves stability and convergence in heterogeneous environments with varying data distributions and computational capabilities (Li et al., 2020a). In multi-hospital collaborations where demographic and clinical variations are pronounced, FedProx enhances robustness against statistical divergence.

3.4.3 Synchronization Strategies

Federated systems may operate under synchronous or asynchronous update mechanisms. Synchronous protocols require all selected clients to complete local training before aggregation, ensuring consistency but potentially increasing latency. Asynchronous approaches allow incremental updates, improving scalability but introducing challenges related to stale gradients and convergence control. Selecting an appropriate synchronization strategy depends on institutional infrastructure, network bandwidth, and clinical workload constraints (Kairouz et al., 2021).

4. LITERATURE REVIEW

This section synthesizes prior research on federated learning (FL) in medical contexts, structured thematically to capture methodological evolution, privacy enhancements, architectural diversification, and domain-specific applications. The review highlights datasets, experimental settings, performance outcomes, and unresolved research challenges.

4.1 Early Works on FL in Healthcare

4.1.1 Initial Implementations in Medical Imaging and EHRs

The earliest applications of federated learning in healthcare primarily focused on medical imaging tasks, where deep convolutional neural networks were collaboratively trained across institutions without centralizing radiological datasets. A landmark study demonstrated multi-institutional brain tumor segmentation using federated training on MRI data, achieving performance comparable to centralized baselines while preserving institutional data boundaries (Sheller et al., 2018). Similar efforts extended to histopathology and chest X-ray classification tasks, validating the feasibility of cross-silo collaboration.

Parallel research explored FL for electronic health record (EHR) analytics, particularly for predictive modeling of clinical outcomes such as mortality and readmission risk. These studies confirmed that decentralized gradient aggregation could maintain predictive accuracy despite heterogeneous data distributions across hospitals.

4.1.2 Initial Privacy Challenges Identified

Although early implementations validated functional viability, they also exposed privacy vulnerabilities inherent in gradient sharing. Research demonstrated that adversaries could reconstruct input data from model updates using gradient inversion techniques (Zhu et al., 2019). Additionally, membership inference attacks revealed risks of identifying whether specific patient records contributed to model training. These findings underscored that FL alone does not guarantee privacy and necessitated additional protective mechanisms.

4.2 Privacy Enhancements Integrated with FL

4.2.1 Differential Privacy in Federated Settings

Subsequent research incorporated differential privacy (DP) into federated optimization workflows. By adding calibrated noise to gradients before aggregation, DP-FL frameworks provided formal privacy guarantees quantified by a privacy budget parameter. Empirical studies showed that moderate privacy budgets preserved acceptable accuracy in medical image classification while limiting information leakage (Geyer et al., 2017). However, excessive noise injection degraded convergence stability, highlighting the need for adaptive privacy calibration strategies.

4.2.2 Cryptographic Secure Aggregation and SMPC

To mitigate risks from untrusted servers, secure aggregation protocols based on secure multi-party computation (SMPC) were introduced. These methods ensured that the central aggregator could only access encrypted or masked parameter sums rather than individual updates (Bonawitz et

al., 2017). In healthcare collaborations, secure aggregation enhanced trust among participating hospitals by preventing exposure of institution-specific gradient information.

4.2.3 Homomorphic Encryption and Trusted Execution

Homomorphic encryption (HE) and Trusted Execution Environments (TEEs) further strengthened confidentiality guarantees. HE-enabled federated frameworks allowed encrypted model updates to be aggregated without decryption, albeit at higher computational cost (Gentry, 2009). TEEs, such as Intel SGX, provided hardware-isolated aggregation environments that protected model parameters during computation. Comparative evaluations indicated that cryptographic methods offer stronger theoretical guarantees, while TEEs provide improved efficiency under controlled hardware trust assumptions (Costan and Devadas, 2016).

4.3 Decentralized / Distributed Model Training Approaches

4.3.1 Peer-to-Peer Decentralized Medical FL

Moving beyond centralized orchestration, decentralized peer-to-peer FL models were proposed to eliminate reliance on a coordinating server. These systems employed consensus-based optimization and gossip protocols to propagate updates among hospitals. Studies reported improved fault tolerance and resilience against single-point failure, although convergence rates were sensitive to network topology and communication delays (Lian et al., 2017).

4.3.2 Block chain and Ledger-Assisted FL Systems

Block chain integration emerged as a trust-enhancing mechanism in federated healthcare systems. Distributed ledgers recorded model updates immutably, while smart contracts enforced aggregation policies and participation incentives. Empirical evaluations demonstrated improved transparency and tamper resistance, particularly in multi-stakeholder environments involving research hospitals and diagnostic centers (Li et al., 2020). Nevertheless, block chain latency and scalability remain limiting factors in large-scale medical deployments.

4.3.3 Centralized vs Decentralized Paradigm Comparison

Comparative analyses reveal that centralized FL architectures generally achieve faster convergence due to structured coordination, whereas decentralized frameworks enhance trust distribution and robustness. Centralized models are computationally efficient but vulnerable to server compromise, while fully decentralized systems require sophisticated synchronization protocols. The choice of paradigm depends on institutional governance structures and risk tolerance levels.

4.4 Applications and Case Studies

4.4.1 Medical Imaging

Federated learning has been widely applied to MRI-based tumor segmentation, CT-based COVID-19 diagnosis, and mammographic cancer detection. Multi-center imaging collaborations demonstrated that federated models can generalize better across demographic variations compared to single-institution training (Rieke et al., 2020). These studies typically employed convolutional neural networks and evaluated performance using Dice similarity coefficients and AUC metrics.

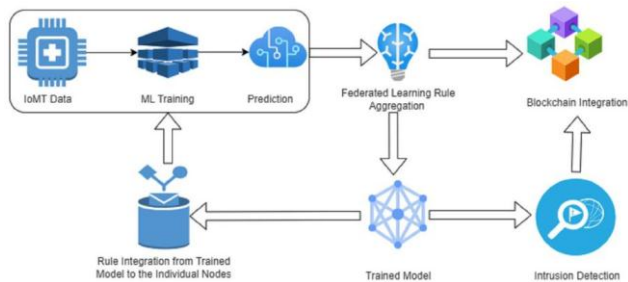


Figure-2: Federated learning in the Medical

4.4.2 Electronic Health Records Analytics

FL-based predictive models for EHR data have addressed tasks such as ICU mortality prediction and sepsis detection. Recurrent neural networks and transformer-based architectures were trained across geographically distributed hospitals. Results indicated improved robustness against local data bias while maintaining compliance with regulatory constraints.

4.4.3 Genomic and Biosignal Frameworks

Emerging research extends FL to genomic sequence analysis and wearable biosignal monitoring. Distributed training across genomic repositories has enabled collaborative variant classification without exposing sensitive genetic markers. Similarly, federated frameworks applied to ECG and EEG data demonstrate privacy-preserving cardiac and neurological monitoring systems.

5. PRIVACY AND SECURITY ENHANCEMENTS

Federated learning (FL) mitigates direct data sharing risks by retaining patient data within institutional boundaries; however, the exchange of model parameters introduces new security and privacy vulnerabilities. Adversarial participants or compromised aggregation servers may exploit gradient information to infer sensitive data or manipulate model behaviour. Consequently, robust privacy-preserving algorithms and threat-aware system designs are essential for secure decentralized medical AI deployments.

5.1 Differential Privacy in FL: Algorithms and Trade-offs

5.1.1 Differentially Private Gradient Mechanisms

Differential Privacy (DP) provides formal privacy guarantees by ensuring that the inclusion or exclusion of a single data record does not significantly influence model outputs. In federated settings, DP is commonly implemented through gradient clipping followed by calibrated noise injection before transmission to the aggregator. This approach bounds sensitivity and limits potential information leakage from individual client updates (Dwork, 2006). Client-level DP further strengthens protection by masking the contribution of entire institutional datasets, which is particularly relevant in cross-silo healthcare collaborations.

5.1.2 Privacy-Utility Trade-offs

While DP enhances confidentiality, it introduces a measurable trade-off between model accuracy and privacy budget (ϵ). Excessive noise reduces convergence speed and predictive performance, especially in high-dimensional medical imaging tasks. Adaptive privacy accounting and dynamic noise scaling strategies have been proposed to balance diagnostic accuracy with regulatory compliance requirements (Geyer, Klein and Nabi, 2017). Selecting optimal privacy parameters remains a context-dependent challenge in safety-critical healthcare systems.

5.2 Secure Aggregation and Encryption Techniques

5.2.1 Secure Aggregation Protocols

Secure aggregation ensures that individual client updates remain confidential during the aggregation process. Protocols based on cryptographic masking and secret sharing allow the server to compute only the aggregated sum of gradients without accessing intermediate values (Bonawitz et al., 2017). This technique reduces the risk of server-side inference attacks while maintaining computational efficiency suitable for multi-hospital federated deployments.

5.2.2 Homomorphic Encryption and Multi-Party Computation

Homomorphic encryption (HE) enables mathematical operations to be performed directly on encrypted data, allowing aggregation without exposing plaintext parameters. Although HE provides strong confidentiality guarantees, it imposes significant computational overhead, which may limit real-time medical applications (Gentry, 2009). Secure Multi-Party Computation (SMPC) offers an alternative cryptographic approach, distributing computation among participants such that no single entity learns the complete input. SMPC-based federated systems enhance trust in collaborative clinical networks where mutual distrust may exist.

5.3 Block chain and Smart Contracts for Trust Management

Block chain technology introduces decentralized trust management into federated learning ecosystems. By recording model updates and aggregation events in an immutable distributed ledger, block chain enhances transparency and auditability. Smart contracts automate validation rules, enforce access control policies, and manage incentive mechanisms among participating institutions.

In healthcare FL frameworks, block chain mitigates risks of model tampering and unauthorized modification by ensuring traceability of updates. However, consensus mechanisms introduce latency and scalability constraints, particularly when handling large parameter exchanges. Hybrid designs combining block chain logging with off-chain aggregation have been proposed to address efficiency limitations (Li et al., 2020).

5.4 Threat Models and Attack Vectors

Understanding adversarial strategies is critical for designing resilient federated medical systems. Threat models typically consider malicious clients, compromised servers, or external eavesdroppers.

5.4.1 Poisoning Attacks

Poisoning attacks occur when adversarial clients intentionally submit manipulated gradients to corrupt the global model. In medical diagnosis systems, such attacks may bias predictions or degrade accuracy. Data poisoning modifies local training data, whereas model poisoning directly alters gradient updates. Robust aggregation rules and anomaly detection mechanisms are necessary to mitigate such threats (Kairouz et al., 2021).

5.4.2 Model Inversion and Reconstruction Attacks

Model inversion attacks aim to reconstruct sensitive input data from shared gradients or model parameters. Empirical demonstrations show that gradient leakage can reveal patient images or training samples under certain conditions (Zhu, Liu and Han, 2019). These vulnerabilities highlight the necessity of combining FL with differential privacy and secure aggregation to prevent unauthorized reconstruction of medical records.

5.4.3 Eavesdropping and Relay Attacks

Eavesdropping attacks exploit insecure communication channels to intercept model updates during transmission. Relay or man-in-the-middle attacks may alter updates before they reach the aggregator. Secure communication protocols employing end-to-end encryption and authenticated channels are essential to prevent interception and tampering. Network-level security must complement

algorithmic privacy measures to ensure comprehensive protection in distributed healthcare infrastructures.

6. EVALUATION METRICS AND BENCHMARKING

Robust evaluation of federated learning (FL) systems in healthcare requires multidimensional assessment, encompassing predictive performance, privacy guarantees, computational efficiency, and dataset representativeness. Unlike conventional centralized models, federated frameworks introduce additional variables such as communication cost, client heterogeneity, and cryptographic overhead. This section outlines the principal evaluation metrics and benchmarking strategies adopted in medical FL research.

6.1 Performance Metrics for Federated Models

6.1.1 Accuracy

Accuracy represents the proportion of correctly classified instances over the total number of predictions. In federated medical classification tasks—such as tumor detection or disease diagnosis—accuracy provides a general performance indicator. However, it may be misleading in imbalanced clinical datasets where negative cases significantly outnumber positive ones (Saito and Rehmsmeier, 2015). Consequently, additional metrics are required to ensure clinically meaningful evaluation.

6.1.2 Area Under the Curve (AUC)

The Area Under the Receiver Operating Characteristic Curve (AUC-ROC) evaluates a model's ability to discriminate between classes across different threshold settings. In multi-institutional FL studies, AUC is widely used for assessing diagnostic robustness across heterogeneous data sources. It is particularly valuable in medical risk prediction models, where threshold-independent evaluation is necessary to balance sensitivity and false-positive rates (Bradley, 1997).

6.1.3 Sensitivity and Specificity

Sensitivity (true positive rate) measures the ability to correctly identify patients with a condition, while specificity (true negative rate) evaluates correct identification of non-affected individuals. These metrics are critical in safety-sensitive healthcare applications, such as cancer screening or infectious disease detection. High sensitivity minimizes missed diagnoses, whereas high specificity reduces unnecessary interventions. Federated frameworks often report these metrics to demonstrate clinical reliability (Powers, 2011).

6.2 Privacy and Security Metrics

6.2.1 Privacy Loss (ϵ)

In differentially private federated learning, privacy guarantees are quantified using the privacy budget parameter (ϵ), which measures the maximum information leakage attributable to an individual record. Lower ϵ values correspond to stronger privacy but typically introduce greater noise and reduced model utility. Privacy accounting techniques, including moment's accountant methods, are used to track cumulative privacy loss across training rounds (Abadi et al., 2016). Selecting an appropriate ϵ in healthcare systems requires balancing regulatory compliance with diagnostic performance.

6.2.2 Cryptographic Overhead

Cryptographic mechanisms such as secure aggregation, homomorphic encryption, and secure multi-party computation introduce additional computational and communication overhead. Evaluation commonly includes encryption time, decryption latency, and increased message size. These metrics are crucial for determining system feasibility in resource-constrained hospital networks. Empirical analyses indicate that fully homomorphic encryption significantly increases processing time compared to secure aggregation schemes, highlighting trade-offs between theoretical security strength and practical deployability (Gentry, 2009).

6.3 Communication and Computation Overheads

Federated learning inherently requires repeated communication between clients and aggregators. Communication overhead is typically measured in terms of transmitted bytes per round and total communication rounds to convergence. Optimization strategies such as model compression, quantization, and sparse updates are frequently evaluated to reduce bandwidth consumption (Kairouz et al., 2021).

Computation overhead includes local training time, memory usage, and server-side aggregation latency. In medical environments with heterogeneous computational infrastructure, performance benchmarking must account for variability in hardware capabilities. Efficient federated protocols aim to minimize synchronization delays while maintaining convergence stability under non-independent and identically distributed (non-IID) data conditions.

7. CONCLUSION

Decentralized collaborative model training using federated learning (FL) represents a transformative paradigm for privacy-preserving artificial intelligence in healthcare. This review has examined architectural frameworks, privacy-enhancing mechanisms, decentralized coordination strategies, and domain-specific applications across medical

imaging, electronic health records, and genomic analytics. The synthesis of existing literature indicates that FL effectively mitigates direct data-sharing risks while enabling multi-institutional knowledge integration. However, federated systems are not inherently secure; vulnerabilities such as gradient leakage, poisoning attacks, and inference risks necessitate the integration of differential privacy, secure aggregation, cryptographic protocols, and trust-enhancing mechanisms such as block chain. Evaluation metrics must extend beyond predictive accuracy to incorporate privacy budgets, communication costs, and computational feasibility, particularly in heterogeneous clinical environments. Although substantial progress has been achieved, unresolved challenges remain in handling non-IID data distributions, scalability constraints, regulatory harmonization, and real-world deployment readiness. Future research should prioritize adaptive privacy-utility optimization, robust aggregation under adversarial conditions, and standardized benchmarking frameworks. Overall, privacy-aware federated learning offers a viable pathway toward secure, scalable, and ethically responsible medical AI systems capable of supporting collaborative healthcare innovation.

8. LIMITATIONS OF REVIEW

This review is subject to several limitations. First, the rapidly evolving nature of federated learning research means that emerging techniques and preprint contributions may not be comprehensively covered. Second, the analysis primarily focuses on cross-silo healthcare settings, with comparatively limited discussion of cross-device medical IoT scenarios. Third, many referenced studies rely on simulated federated environments rather than fully distributed real-world deployments, which may restrict external validity. Additionally, quantitative meta-analysis was not performed due to heterogeneity in evaluation protocols, datasets, and reporting standards across studies. Finally, regulatory, ethical, and socio-technical dimensions—such as patient consent frameworks and governance interoperability—were discussed conceptually but not examined through empirical policy analysis.

REFERENCES

1. Abadi, M., Chu, A., Goodfellow, I., McMahan, H.B., Mironov, I., Talwar, K. and Zhang, L. (2016) 'Deep learning with differential privacy', Proceedings of the ACM Conference on Computer and Communications Security (CCS), pp. 308–318.
2. Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H.B., Patel, S., Ramage, D., Segal, A. and Seth, K. (2017) 'Practical secure aggregation for privacy-preserving machine learning', Proceedings of the ACM Conference on Computer and Communications Security (CCS), pp. 1175–1191.

3. Bradley, A.P. (1997) 'The use of the area under the ROC curve in the evaluation of machine learning algorithms', *Pattern Recognition*, 30(7), pp. 1145–1159.
4. Costan, V. and Devadas, S. (2016) 'Intel SGX explained', *IACR Cryptology ePrint Archive*, 2016(086), pp. 1–118.
5. Dwork, C. (2006) 'Differential privacy', *Proceedings of the 33rd International Colloquium on Automata, Languages and Programming (ICALP)*, pp. 1–12.
6. Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M. and Thrun, S. (2017) 'Dermatologist-level classification of skin cancer with deep neural networks', *Nature*, 542(7639), pp. 115–118.
7. Gentry, C. (2009) 'Fully homomorphic encryption using ideal lattices', *Proceedings of the 41st ACM Symposium on Theory of Computing (STOC)*, pp. 169–178.
8. Geyer, R.C., Klein, T. and Nabi, M. (2017) 'Differentially private federated learning: A client level perspective', *NIPS Workshop on Private Multi-Party Machine Learning*.
9. Kairouz, P., McMahan, H.B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A.N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R., D'Oliveira, R.G.L., Eichner, H., El Rouayheb, S., Evans, D., Garcia-Saavedra, A., et al. (2021) 'Advances and open problems in federated learning', *Foundations and Trends in Machine Learning*, 14(1–2), pp. 1–210.
10. Li, T., Sahu, A.K., Talwalkar, A. and Smith, V. (2020) 'Federated optimization in heterogeneous networks', *Proceedings of Machine Learning and Systems (MLSys)*, pp. 429–450.
11. Li, X., Jiang, M., Zhang, X., Kamp, M. and Dou, Q. (2020) 'A blockchain-based federated learning framework for data privacy preservation', *IEEE Transactions on Industrial Informatics*, 16(6), pp. 4287–4296.
12. Lian, X., Zhang, C., Zhang, H., Hsieh, C.-J., Zhang, W. and Liu, J. (2017) 'Can decentralized algorithms outperform centralized algorithms? A case study for decentralized parallel stochastic gradient descent', *Advances in Neural Information Processing Systems*, 30, pp. 5330–5340.
13. Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., Van der Laak, J.A.W.M., Van Ginneken, B. and Sánchez, C.I. (2017) 'A survey on deep learning in medical image analysis', *Medical Image Analysis*, 42, pp. 60–88.
14. McMahan, H.B., Moore, E., Ramage, D., Hampson, S. and Arcas, B.A.Y. (2017) 'Communication-efficient learning of deep networks from decentralized data', *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, pp. 1273–1282.
15. Narayanan, A. and Shmatikov, V. (2008) 'Robust de-anonymization of large sparse datasets', *IEEE Symposium on Security and Privacy*, pp. 111–125.
16. Powers, D.M.W. (2011) 'Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation', *Journal of Machine Learning Technologies*, 2(1), pp. 37–63.
17. Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H.R., Albarqouni, S., Bakas, S., Galtier, M.N., Landman, B.A., Maier-Hein, K., Ourselin, S., Sheller, M. and Cardoso, M.J. (2020) 'The future of digital health with federated learning', *npj Digital Medicine*, 3(119), pp. 1–7.
18. Saito, T. and Rehmsmeier, M. (2015) 'The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets', *PLoS ONE*, 10(3), e0118432.
19. Sheller, M.J., Reina, G.A., Edwards, B., Martin, J. and Bakas, S. (2018) 'Multi-institutional deep learning modeling without sharing patient data: A feasibility study on brain tumor segmentation', *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pp. 92–104.
20. Sheller, M.J., Edwards, B., Reina, G.A., Martin, J., Pati, S., Kotrotsou, A., Milchenko, M., Xu, W., Marcus, D., Colen, R.R. and Bakas, S. (2020) 'Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data', *Scientific Reports*, 10, 12598.
21. Zhu, L., Liu, Z. and Han, S. (2019) 'Deep leakage from gradients', *Advances in Neural Information Processing Systems*, 32, pp. 14774–14784.
22. Akhmetov, A., Latif, Z., Tyler, B. and Yazici, A. (2025) 'Enhancing healthcare data privacy and interoperability with federated learning', *PeerJ Computer Science*, 11:e2870.
23. Adnan, M., Kalra, S., Cresswell, J.C., Taylor, G.W. and Tizhoosh, H.R. (2022) 'Federated learning and differential privacy for medical image analysis', *Scientific Reports*, 12, 1953.
24. Ali, M.S., Ahsan, M.M., Tasnim, L. et al. (2024) 'Federated learning in healthcare: Model misconducts, security, challenges, applications, and future research directions', *arXiv*, 2405.13832.
25. Daram, S. (2025) 'Federated learning in medical AI: Advancing privacy-preserving data sharing for collaborative healthcare research', *International*

- Journal of Artificial Intelligence, Data Science, and Machine Learning, 6(2).
26. Dendukuri, S.V. (2025) 'Federated learning in healthcare: Protecting patient privacy while advancing analytics', *Journal of Computer Science and Technology Studies*, 7(7), pp.840–845.
 27. Gu, X., Sabrina, F., Fan, Z. and Sohail, S. (2023) 'A review of privacy enhancement methods for federated learning in healthcare systems', *International Journal of Environmental Research and Public Health*, 20(15), 6539.
 28. HariPriya, R., Khare, N. and Pandey, M. (2025) 'Privacy-preserving federated learning for collaborative medical data mining in multi-institutional settings', *Scientific Reports*, 15, 12482.
 29. HariPriya, R., Khare, N., Pandey, M. and Biswas, S. (2025) 'A privacy-enhanced framework for collaborative Big Data analysis in healthcare using adaptive federated learning aggregation', *Journal of Big Data*, 12, 113.
 30. Manjula, N.J., Randhi, K. and Bandarapu, S.R. (2023) 'Federated learning for healthcare: Balancing data privacy and model accuracy', *American Journal of Computing and Engineering*, 6(1), pp.69–80.
 31. Pati, S., Kumar, S. et al. (2024) 'Privacy preservation for federated learning in health care', *Patterns (N Y)*, 5(7), 100974.
 32. Sandhu, S.S., Taheri Gorji, H. et al. (2023) 'Medical imaging applications of federated learning', *Diagnostics*, 13(19), 3140.
 33. Shah, S.T., Ali, Z., Waqar, M. and Kim, A. (2025) 'Federated learning in public health: Decentralized, equitable, and secure disease prevention approaches', *Healthcare*, 13(21), 2760.
 34. Verma, A. and Gonzalez, M. (2023) 'Privacy-preserving federated learning for healthcare data sharing', *International Journal of Recent Advances in Engineering and Technology*, 12(2), pp.14–20.
 35. Zafar, A., Saad, M. and Haque, S.B.U. (2025) 'Efficient and privacy-enhanced federated learning for medical imaging in resource-limited environments', *Journal of Electrical Systems*, 21(01).
 36. Koutsoubis, N., Waqas, A., Yilmaz, Y., Ramachandran, R.P., Schabath, M. and Rasool, G. (2024) 'Future-proofing medical imaging with privacy-preserving federated learning and uncertainty quantification: A review', arXiv, 2409.16340.
 37. Dendukuri, S.V. (2025) 'Federated Learning in Healthcare: Protecting Patient Privacy ...', *Journal of Computer Science and Technology Studies*, 7(7), pp.840–845.
 38. Leveraging federated learning for rare disease EHR analysis (2024) *Journal of Electronic Clinical Data Science*, DOI: 10.1016/j.ject.2024.11.001.
 39. Federated Learning in Smart Healthcare (2024) 'Federated learning in smart healthcare: Privacy, security and IoT predictive analytics', *Healthcare*, 12(24), 2587.
 40. Anonymizing Data for Privacy-Preserving FL (2020) 'Anonymizing data for privacy-preserving federated learning', arXiv, 2002.09096.
 41. Comprehensive surveys of FL methods in medical imaging (2023) 'Federated learning for medical image analysis: A survey', PubMed.
 42. Darzidehkalani, E., Ghasemi-Rad, M. et al. (2022) 'Federated learning in medical imaging: Methods, challenges, and considerations', *Journal of the American College of Radiology*, 19(8):975–982.
 43. Federated Learning Approaches for Healthcare AI (2025) 'Federated learning approaches for privacy-preserving AI in healthcare data science', *Journal of Informatics Education and Research*, 5(2).
 44. Systematic review of FL in healthcare ethics (2026) 'Federated learning in healthcare ethics: Privacy-preserving and equitable medical AI', *Healthcare*, 14(3), 306.
 45. Systematic Pattern analysis of FL privacy (2024) 'Privacy preservation for federated learning in health care', *Patterns*, 5(7), 100974.
 46. Comprehensive review of FL challenges (2024) 'Federated learning in healthcare: Model misconducts, security, challenges', arXiv:2405.13832.
 47. IEEE-scale review of privacy enhancement FL (2025) 'Efficient and privacy-enhanced federated learning', *Journal of Electrical Systems*.
 48. Medical AI review on FL and privacy (2025) Daram, S. IJAIDSML.
 49. FL in healthcare analytics framework (2025) Richardson, A. 'Federated Learning Framework for Privacy-Preserving Healthcare Analytics', *Eureka Journal of Computing Science & Digital Innovation*, 1(1), pp.8–14.