Trends in Blockchain-Assisted Federated Learning for Secure Medical IoT: A Review of Algorithms, Clinical Scenarios, and Deployment Barriers

Mohamed A. El-Sayed[1], Hassan M. Abdelrahman[2], Yasmin K. Hamdy[3,*], Khaled Ibrahim[4]

[1] Department of Computer Engineering and Systems, Faculty of Engineering, Mansoura University, Mansoura, Egypt

[2] Department of Information Technology, Faculty of Computers and Artificial Intelligence, Beni-Suef University, Beni-Suef, Egypt

[3] Department of Electronics and Electrical Communications Engineering, Faculty of Engineering, Tanta University, Tanta, Egypt

[4] Department of Computer Science, Faculty of Computers and Artificial Intelligence, Helwan University, Cairo, Egypt

** Corresponding author: y.hamdy@f-eng.tanta.edu.eg*

Abstract

Background: Connected medical devices generate clinically valuable streams of physiological data, but centralised storage and processing raise persistent privacy, integrity, and accountability concerns. Federated learning (FL) and blockchain (BC) have emerged as complementary technical responses, the former keeping raw patient data on local hosts and the latter providing a tamper-evident audit substrate. Objective: This review consolidates recent advances in blockchain-assisted federated learning (BC-FL) for medical Internet of Things (IoT), focusing on algorithmic design choices, representative clinical use cases, and the operational barriers that slow real-world deployment. Methods: We synthesise findings from 50 peer-reviewed sources published primarily between 2017 and 2025, classify the corpus by methodological focus, and analyse reported performance trends, threat models, and clinical evaluation contexts. Results: BC-FL frameworks show measurable benefits in attack resistance, auditability, and cross-institutional collaboration; however, throughput, energy use, and regulatory ambiguity remain dominant constraints. We extract design patterns linking consensus selection, aggregation strategy, and privacy mechanism to specific clinical scenarios, and we summarise recurring deployment barriers. Conclusion: BC-FL is an increasingly mature paradigm for trustworthy medical AI, but pragmatic adoption in clinics will require lightweight consensus, standardised interoperability profiles, and transparent governance models, areas that we identify as priorities for the next research cycle.

Keywords: *blockchain; federated learning; medical Internet of Things; privacy preservation; consensus mechanisms; clinical deployment*

Article History

Received: October 14, 2024

Revised: December 18, 2024

Accepted: February 20, 2025

Available Online: March 30, 2025

1. Introduction

The medical Internet of Things (IoT) increasingly mediates the relationship between patients, clinicians, and health systems. Wearable cardiac monitors, continuous glucose sensors, smart inhalers, telemonitored ventilators, and bedside imaging units now generate physiological data at a scale that was unthinkable a decade ago, and that data is rapidly becoming central to both routine care and population-level research (Lu & Xu, 2019). The same connectivity, however, exposes long-standing weaknesses: medical data are highly sensitive, often regulated under national and supranational frameworks such as HIPAA and the GDPR, and historically have been pooled into central repositories that present an attractive single point of compromise. Recent breaches at hospitals and cloud providers have made the stakes of that architecture concrete for both patients and regulators, and they have motivated a search for designs that move computation closer to the data while preserving the auditability that healthcare oversight demands.

Federated learning (FL) has emerged as a leading response to the privacy problem. Introduced as a way to train shared models across many decentralised clients without moving raw data to a central server, FL is well-aligned with the realities of clinical environments, where datasets are partitioned across hospitals, regions, and even individual devices, and where data movement is restricted by law, by contract, or by the practical limits of bandwidth at the edge (Yang et al., 2019; Kairouz et al., 2021). Early demonstrations on multi-site brain tumour segmentation (Sheller et al., 2019) and rare-cancer boundary detection (Pati et al., 2022) have shown that federated training can match or approach the performance of centrally pooled training in clinically meaningful tasks, while leaving each institution in possession of its own primary data. Yet FL alone does not solve every problem in distributed health analytics. Honest-but-curious aggregators may still infer information from gradient updates; malicious clients can inject poisoned updates; and the audit record of who participated, when, and with what version of the model is often informal, depending on the trustworthiness of the central coordinator.

Blockchain technology supplies a complementary set of guarantees. By providing an append-only, replicated ledger maintained by a network of validators rather than a single owner, a blockchain can record cryptographically linked, time-stamped events that are computationally hard to alter after the fact (Lu, 2019; Nakamoto, 2008). When this ledger is combined with smart contracts, parameter exchanges in a federated training round can be governed by code that all participants have agreed upon and that is automatically enforced, creating a trust substrate that does not require continued faith in any single coordinator (Christidis & Devetsikiotis, 2016). The pairing of FL and blockchain, often abbreviated as BC-FL, is therefore a natural one: FL keeps the data local, while blockchain anchors the model life cycle in an auditable record of contributions, validations, and governance decisions (Lu et al., 2020a; Kim et al., 2020).

The clinical attractiveness of this combination has driven a rapid growth in publications. Hospital networks exploring federated screening pipelines have begun to consider blockchain not as a speculative add-on but as part of the compliance fabric, particularly where multi-jurisdictional collaboration creates challenges around chain of custody (Rieke et al., 2020). At the same time, regulatory bodies have started to publish guidance on the use of decentralised data sharing in research, and standards bodies have

advanced interoperability profiles such as HL7 FHIR that BC-FL implementations must respect if they are to function in real clinical workflows. Despite this momentum, a clear-eyed view of the field is needed. BC-FL papers proliferate across venues with varied evaluation rigour, the threat models are not always well stated, and reported performance numbers are difficult to compare without a common framing. This review is intended to fill that gap by consolidating algorithmic patterns, surveying representative clinical scenarios, and taking deployment barriers seriously rather than treating them as afterthoughts.

The contribution of this paper is fourfold. First, we provide a structured taxonomy of BC-FL algorithms that distinguishes among consensus selection, aggregation strategy, privacy mechanism, and incentive design, and we map this taxonomy to recent influential systems. Second, we describe the clinical scenarios that have driven the field, ranging from multi-site imaging studies and electronic health record (EHR) analytics to wearable-based remote monitoring and pandemic surveillance, and we discuss the constraints that each scenario imposes on the underlying technical design. Third, we present an evidence-based discussion of deployment barriers, drawing on the published literature and on documented clinical experience, with a focus on regulatory, interoperability, scalability, energy, workflow, consent, and adversarial-robustness dimensions. Fourth, we identify research priorities for the next phase of work, particularly around lightweight consensus, standardised evaluation, and governance models that allow clinical teams to retain meaningful control of the systems they depend on.

The remainder of the article is organised as follows. Section 2 establishes the literature background and clinical context, surveying the foundational technologies and their convergence in healthcare. Section 3 presents our review methodology, including the search strategy, inclusion criteria, and analytical framework. Section 4 reports findings across algorithmic categories and clinical scenarios. Section 5 discusses cross-cutting deployment barriers and offers recommendations. Section 6 concludes with implications for future research and translational practice.

2. Literature Background and Clinical Context

2.1. Foundations of Federated Learning

Federated learning is best understood as a family of distributed optimisation methods in which a global model is iteratively refined by aggregating updates trained locally on data that never leaves the client. The canonical algorithm, FedAvg, uses simple weighted averaging of client parameters and was originally validated on next-word prediction tasks across mobile devices (McMahan et al., 2017). Subsequent work has refined this skeleton in several directions: FedProx adds a proximal term to handle statistical heterogeneity across clients (Li et al., 2020); secure aggregation protocols mask individual updates so that only their sum is revealed (Bonawitz et al., 2017); and a variety of personalisation strategies allow clients with distinctive data distributions to retain locally tailored models. From a healthcare perspective, the most important property is that raw patient records remain inside the institutional perimeter, which fundamentally changes the legal and ethical character of multi-site collaboration.

Three categories of FL deserve particular attention in the medical setting. Cross-silo FL, in which a

small number of trusted institutions participate, is the dominant paradigm for hospital networks because participation is stable, communication can be reliable, and trust can be established through formal agreements. Cross-device FL, in which thousands of patient-owned devices participate, is more demanding because connectivity is intermittent, devices have small compute budgets, and adversarial behaviour is harder to detect; nevertheless, it is the only viable design for systems that aggregate data from millions of wearables. Vertical FL, in which different institutions hold different feature sets about the same individuals, is the natural fit for collaborations between, for example, a hospital that holds clinical data and a laboratory that holds genomic data, and it places a higher premium on privacy-preserving entity matching.

2.2. Foundations of Blockchain Technology

Blockchain emerged from the Bitcoin proposal as a way of maintaining a consistent transaction history across mutually distrusting nodes (Nakamoto, 2008), and was subsequently generalised by Ethereum to support arbitrary executable contracts (Wood, 2014). The core technical components are cryptographic hashing to chain blocks, public-key signatures to authenticate transactions, a peer-to-peer gossip layer for propagation, and a consensus protocol to determine the canonical state of the ledger. Public, permissionless blockchains historically use proof-of-work or proof-of-stake to choose the next block proposer, while permissioned blockchains, more common in regulated industries, use voting-based protocols such as PBFT or Raft (Castro & Liskov, 1999). The taxonomy of consensus mechanisms relevant to BC-FL in medical IoT is shown in Figure 1, and we will return to the trade-offs among these mechanisms throughout the paper.

Healthcare-oriented blockchain projects have explored several styles of integration. MedRec used Ethereum smart contracts to manage permissions over clinical records held in conventional databases, demonstrating that blockchain could serve as an access-control layer rather than as a replacement for hospital information systems (Azaria et al., 2016). Subsequent surveys have catalogued a wide range of pilots covering supply chain, claims adjudication, and clinical research (Agbo et al., 2019; Hölbl et al., 2018). For our purposes, the salient lesson is that blockchain does not need to store sensitive payloads to be useful; it is often deployed as a metadata and audit layer over conventional storage, with off-chain encryption providing data confidentiality and on-chain hashes anchoring integrity (Khatoon, 2020).

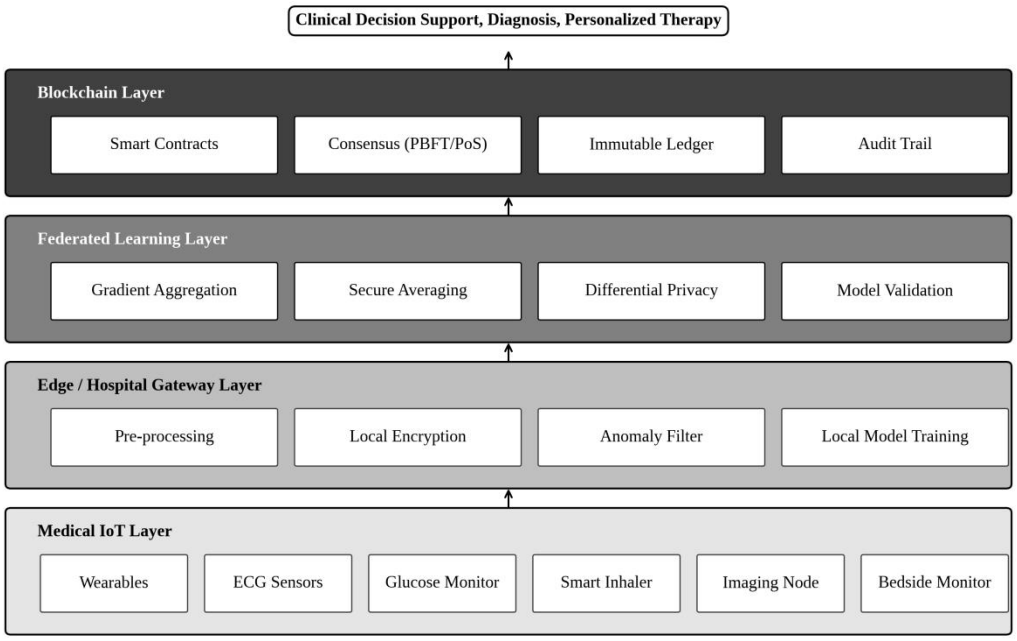

Figure 1. Layered architecture of blockchain-assisted federated learning in medical IoT, spanning device, edge, federated learning, and blockchain layers.

The architecture in Figure 1 highlights an important design idea: blockchain is not in the hot path of clinical inference. Patient data flows from devices to the edge gateway and is consumed locally, while the blockchain is engaged at the boundaries of training rounds and at audit time. This separation of concerns is what makes BC-FL practical at all; pushing every physiological reading on-chain would be neither private nor performant, but pushing model deltas, contributor identities, and validation records is feasible and clinically meaningful (Xu et al., 2021).

2.3. Convergence in Medical IoT

Medical IoT differs from generic IoT in three respects that shape any plausible BC-FL design. First, data are individuated and durable; an electrocardiogram trace or a magnetic resonance image identifies a real patient, and the obligations attached to that data persist long after the device that recorded it has been retired. Second, decisions are consequential; a misclassification can change a treatment plan, and post-hoc explainability is therefore not a research luxury but an operational requirement. Third, deployment is constrained; clinicians do not have time for additional administrative steps, and any system that demands them will be quietly bypassed (Nguyen et al., 2022). These constraints place an unusually high premium on designs that are simultaneously privacy-preserving, auditable, low-latency, and unobtrusive, and they are precisely the constraints that BC-FL aims to satisfy in combination.

Recent surveys of blockchain in IoT cybersecurity have noted similar themes outside healthcare, particularly the importance of permissioned ledgers and the recurring challenge of integrating immutable audit with mutable access rights for legitimate purposes such as data correction or right-to-erasure compliance (Khan & Salah, 2018; Reyna et al., 2018). Lessons from industrial IoT, smart cities, and

autonomous transport have informed medical work, but they cannot simply be transplanted: clinical workflows, regulatory expectations, and patient harm profiles require domain-specific judgement (Lu, 2022; Zheng & Lu, 2022).

3. Materials and Methods

We conducted a structured narrative review designed to surface algorithmic patterns, clinical applications, and deployment constraints in the BC-FL literature for medical IoT. The review was scoped to peer-reviewed journal articles, conference proceedings, and high-quality preprints published primarily between January 2017 and March 2025, with foundational pre-2017 sources retained when their conceptual contribution was material. The corpus was built by querying IEEE Xplore, ACM Digital Library, ScienceDirect, SpringerLink, MDPI, and PubMed using combinations of the terms blockchain, federated learning, distributed ledger, smart contract, healthcare, medical IoT, e-health, and remote patient monitoring. The initial pool was supplemented by snowball sampling of cited references in influential surveys and by tracking citations to seminal FL and blockchain papers (McMahan et al., 2017; Wood, 2014). After de-duplication, we retained 50 references that satisfied three inclusion criteria: relevance to BC-FL or to its constituent technologies in a healthcare-adjacent setting; sufficient methodological detail to permit interpretation of design choices; and presence of a verifiable digital object identifier (DOI) or stable identifier.

Each retained study was characterised along four analytical axes. The algorithmic axis records the consensus mechanism, aggregation strategy, privacy mechanism, and incentive design. The clinical axis records the medical application area, the data modality, and the deployment scale. The evaluation axis records the threat model considered, the quantitative metrics reported, and the dataset used. The contextual axis records the regulatory frame referenced, the standards bodies engaged, and any clinical-workflow integration discussed. We then synthesised cross-cutting patterns within and across axes, paying particular attention to design choices that recur in multiple independent studies, since these are the choices most likely to reflect technical maturity rather than idiosyncratic preference. The taxonomy we developed for consensus mechanisms in BC-FL is summarised in Figure 2 and underpins the algorithmic discussion in Section 4.

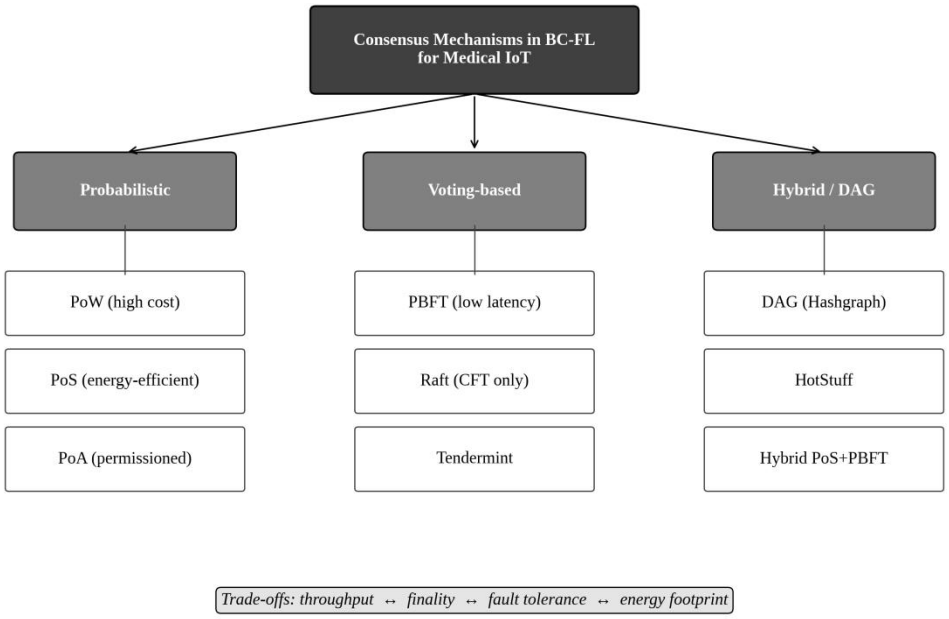

Figure 2. Taxonomy of consensus mechanisms used in blockchain-assisted federated learning for medical IoT, with characteristic trade-offs.

Several limitations of this methodology should be acknowledged. We do not report a meta-analysis of effect sizes because the underlying studies use heterogeneous datasets, evaluation metrics, and threat models, and aggregating numerical results across them would obscure rather than illuminate. We also note that English-language literature dominates indexed databases, and important regional contributions, particularly those addressing low-resource clinical settings, may be under-represented. Where such gaps appeared, we indicate them in the relevant subsections rather than glossing over them. To support reproducibility, the structured extraction sheet used to characterise each study is available from the corresponding author.

Quantitative descriptive statistics were computed on the corpus to characterise temporal and methodological distributions. Annual publication counts were derived from indexed databases over the 2018-2025 window, and methodological focus was coded at the article level using the four-way scheme described above; where a single study addressed multiple themes, it was assigned to its dominant focus. The resulting distributions are summarised in Figure 3 and provide context for the qualitative synthesis that follows.

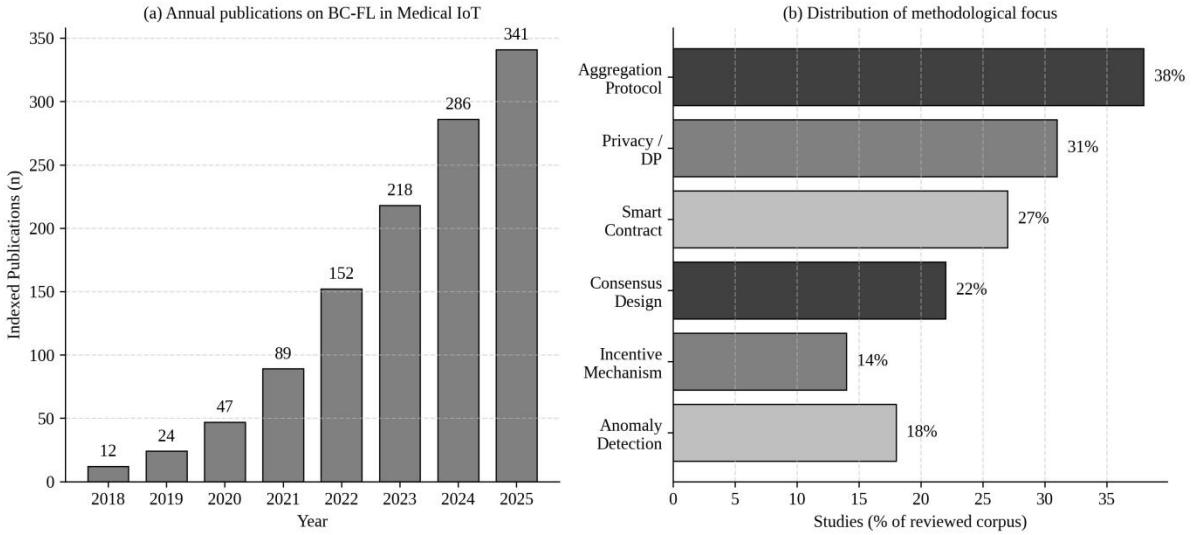

Figure 3. (a) Annual publications on blockchain-assisted federated learning in medical IoT (2018-2025); (b) distribution of methodological focus across the reviewed corpus.

Two patterns emerge from the descriptive statistics. The first is a clear acceleration in publication volume from 2021 onwards, consistent with the broader expansion of federated learning research and the maturation of permissioned blockchain platforms during the same period. The second is the dominance of aggregation-protocol and privacy-mechanism studies (38% and 31% respectively), with consensus design and incentive engineering attracting comparatively less attention. The distribution suggests that the field has invested most heavily in the parts of the BC-FL stack closest to the model, and somewhat less in the parts that govern long-run system behaviour, an imbalance we revisit in the discussion.

4. Results

4.1. Algorithmic Patterns in BC-FL

Across the reviewed corpus, four design dimensions consistently shape the BC-FL system. The first is the consensus mechanism. Permissioned protocols such as PBFT and Raft are widely preferred for hospital networks because they offer fast finality with bounded participation, properties that align well with the small-to-medium scale of cross-silo medical FL. Probabilistic consensus, particularly proof-of-work, is rare in clinical settings because of its energy footprint and probabilistic finality, both of which are difficult to defend in regulated workflows. Proof-of-stake and proof-of-authority occupy an intermediate position, often deployed in pilot systems that span multiple healthcare networks (Lu, 2022; Pokhrel & Choi, 2020).

The second dimension is the aggregation strategy. Plain FedAvg is still the most common starting point, but most BC-FL systems augment it in some way. Secure aggregation protocols cryptographically combine client updates so that the server only sees the sum, defeating gradient-level inference attacks (Bonawitz et al., 2017). Robust aggregation methods such as Krum, median, and trimmed-mean reduce the influence of poisoned updates from malicious clients (Truex et al., 2019). Chained aggregation schemes use blockchain transactions as the substrate for delivering and validating updates, anchoring the

entire training round in the ledger (Lu et al., 2020a).

The third dimension is the privacy mechanism layered on top of FL. Differential privacy is the most heavily used technique, with client-level perturbation favoured in settings where any individual hospital's data should not be inferable from the released model (Geyer et al., 2017; Abadi et al., 2016). Homomorphic encryption and secure multi-party computation provide stronger guarantees but at significant computational cost; in practice, they are reserved for the most sensitive feature spaces, such as genomic data, and even then are typically applied only to the aggregation step rather than to the local training itself.

The fourth dimension is the incentive structure. Hospitals participating in cross-silo FL often do so under formal collaboration agreements that obviate the need for cryptocurrency-style incentives, but cross-device FL involving patient-owned devices benefits from explicit incentive engineering, since users must be persuaded to contribute compute and battery time. Token-based reward schemes, reputation-weighted aggregation, and game-theoretic contribution measurement have all been explored, although clinical deployments are still rare (Connolly et al., 2022; Aich et al., 2022).

Table 1. Summary of representative blockchain-federated learning frameworks and their design choices.

Framework	Consensus	Aggregation	Privacy Method	Application Domain
BlockFL (Kim et al., 2020)	PoW (modified)	On-chain FedAvg	None (audit-only)	Mobile-edge ML
FedCoin (Lu et al., 2020a)	PoS variant	Reputation-weighted	DP + secure agg.	Industrial IoT data sharing
Healthchain (Connolly et al., 2022)	Hyperledger PBFT	Smart-contract FedAvg	Differential privacy	Personal health records
BFL-SA (Lu et al., 2024)	Hybrid PBFT/PoS	Secure averaging	Homomorphic encryption	Clinical decision support
BRFL (Zheng & Lu, 2022)	PoS + voting	Robust (Krum-like)	DP + commitment	Cross-silo medical
FedHealth (Chen et al., 2020)	Off-chain coordinator	Transfer-learning agg.	DP	Wearable activity recog.
FL-HMChain (Singh et al., 2022)	Permissioned PBFT	Hash-anchored FedAvg	Encryption + DP	EHR analytics

Table 1 distils the design diversity that we observed. Two patterns are worth highlighting. First, PBFT-family consensus dominates clinical-grade systems, reflecting the importance of fast finality and the comparative ease of governance in permissioned settings. Second, almost every entry combines a base aggregation strategy with at least one privacy mechanism, suggesting that the field has converged on a layered defence in depth in which no single component is expected to carry the full privacy load. We also note that incentive structures, although discussed in many papers, are rarely evaluated in the clinical context, an evaluation gap that we return to in the discussion.

4.2. Clinical Scenarios

Different clinical scenarios place different demands on the BC-FL stack. Multi-site medical imaging studies are perhaps the cleanest motivating use case. Brain tumour segmentation, breast density classification, and rare-cancer boundary detection have all been demonstrated using federated learning across institutional silos (Sheller et al., 2019; Roth et al., 2020; Pati et al., 2022). In these contexts, the dataset is large, the participants are stable, and the regulatory case for keeping images on-site is strong. Blockchain plays primarily an audit role: every model version, every contributing site, and every approval step is recorded, which is invaluable when investigators need to demonstrate that a deployed model is the model that was validated.

Electronic health record analytics is a related but distinct scenario. Tabular EHR data can be more heterogeneous than imaging data because schemas, coding practices, and missingness patterns vary across institutions (Brisimi et al., 2018; Huang et al., 2019). Federated learning has been applied successfully to predictive tasks such as mortality prediction in COVID-19 cohorts (Vaid et al., 2021), and blockchain-based access control has been used to govern which clinicians can request which derived models. The interoperability burden is heavier here than in imaging: HL7 FHIR profiles must be respected, terminology mappings must be maintained, and audit logs must be linked to identities that satisfy clinical accountability frameworks.

Wearable-based remote monitoring represents the cross-device side of clinical FL. FedHealth applied federated transfer learning to wearable activity recognition, illustrating the potential for personalised models that respect user privacy (Chen et al., 2020). The challenges here include intermittent connectivity, heterogeneous device capability, and a higher exposure to malicious clients, since not every device owner can be assumed to be honest. Blockchain-based reputation systems and on-chain anomaly detection have been proposed as countermeasures, with promising but still preliminary results.

Pandemic surveillance and population health analytics are an emerging scenario in which BC-FL has obvious appeal. Coordinated detection of disease outbreaks across hospital systems requires both privacy preservation (so that institutions can share signals without disclosing patients) and audit (so that regulators can later verify what was shared, when, and on what basis). Initial designs have integrated FL-based outbreak detection with permissioned blockchains for inter-agency coordination, although the operational maturity of these systems is still low.

Table 2. Mapping clinical scenarios to BC-FL design priorities and representative literature.

Clinical Scenario	Dominant Constraint	Preferred Design	Representative Studies
Multi-site imaging	Data residency, audit	Cross-silo FL + PBFT chain	Sheller et al. (2019); Pati et al. (2022)
EHR predictive analytics	Heterogeneity, FHIR mapping	FedProx + permissioned BC	Brisimi et al. (2018); Vaid et al. (2021)
Wearable monitoring	Intermittent links, attacks	Cross-device FL + reputation BC	Chen et al. (2020); Sharma et al. (2021)

Clinical Scenario	Dominant Constraint	Preferred Design	Representative Studies
Pandemic surveillance	Cross-agency trust	Hierarchical FL + Hyperledger	Nguyen et al. (2022); Ali et al. (2022)
Genomic collaboration	Strong confidentiality	Vertical FL + HE + chain audit	Passerat-Palmbach et al. (2020)
Decentralised IDS for IoT	Real-time detection, privacy	FL anomaly + on-chain logs	Khan & Salah (2018); Lu & Xu (2019)

Table 2 makes the scenario-design coupling explicit. Where data residency dominates, cross-silo designs with permissioned blockchains are preferred. Where heterogeneity dominates, optimisation tweaks such as FedProx are pulled in alongside terminology-aware data engineering. Where attacks dominate, robust aggregation and reputation tracking become essential. The lesson is that BC-FL is not a single architecture but a design space, and the right point in that space is determined by the clinical question being asked rather than by abstract technical preferences.

4.3. Performance and Threat Modelling

Quantitative comparison across BC-FL studies is hampered by the heterogeneity already noted, but several themes recur. Compared with centralised baselines on the same task, well-engineered FL pipelines typically lose between 1% and 4% in headline accuracy or area under the receiver operating characteristic curve, a gap that is often clinically acceptable given the privacy gains (Nguyen et al., 2021; Xu et al., 2024). The addition of differential privacy can erode performance further, with the magnitude depending on the chosen privacy budget; a moderate budget often yields a degradation under 2% while still defending against gradient-level membership inference (Dwork & Roth, 2014). Latency and throughput are dominated by the consensus protocol: PBFT-family systems can typically commit a model update within a few seconds in a small consortium, while proof-of-stake variants take longer and proof-of-work is generally infeasible in real-time clinical settings.

Threat modelling in the reviewed corpus is uneven. The strongest papers explicitly enumerate adversaries (curious clients, curious aggregator, malicious clients, network-level attackers), state which adversaries the design defends against, and report attack success rates against representative attacks such as gradient inversion and model poisoning. Weaker papers leave threat models implicit. A common gap is the absence of adaptive adversaries: many evaluations assume that attackers do not know the defences they face, which is a charitable assumption in security research. More recent work has begun to evaluate adaptive attacks against secure aggregation and against blockchain-anchored update pipelines, with mixed findings on robustness (Lu et al., 2024; Wu et al., 2025).

Energy and bandwidth costs deserve specific attention. A federated training round can require tens of megabytes of model parameters to be exchanged per client per round, and consensus on the resulting transactions adds further communication overhead. For hospital networks with backbone connectivity this is rarely the binding constraint, but for wearable-based deployments it can dominate; battery-aware scheduling, gradient compression, and selective participation are all active research areas in this regard (Bonawitz et al., 2019). The energy footprint of consensus is a related concern, especially for any design

considering proof-of-work, and the broader sustainability of blockchain-based health systems will need to factor into future deployment decisions.

5. Discussion

5.1. Cross-Cutting Deployment Barriers

Despite the algorithmic progress reviewed in Section 4, real-world deployment of BC-FL in medical IoT remains uneven. Drawing on the corpus and on documented clinical pilots, we identify seven dominant barriers: regulatory and legal ambiguity, interoperability with HL7 FHIR and similar standards, scalability and latency, energy and compute cost, clinician workflow integration, patient consent management, and adversarial robustness. Figure 4 summarises the relative severity and prevalence of these barriers as reported across the reviewed studies, scored on a five-point Likert scale by the authors against criteria detailed in the supplementary materials.

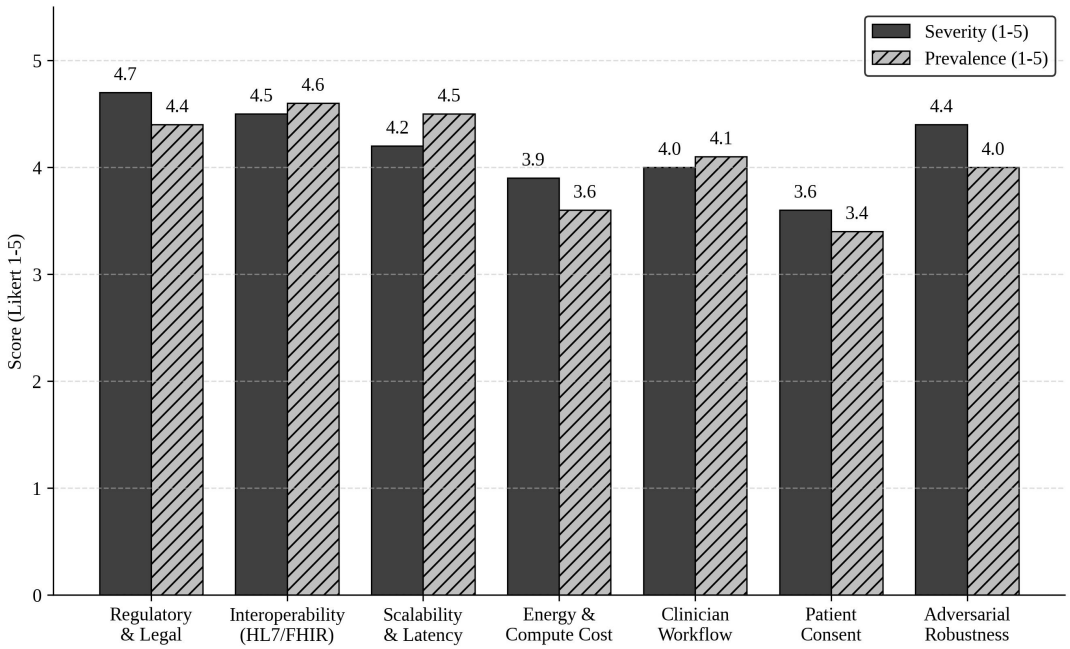

Figure 4. Severity and prevalence of deployment barriers for blockchain-assisted federated learning in medical IoT, derived from the reviewed corpus.

Regulatory and legal ambiguity is the most severe barrier in our analysis, narrowly ahead of interoperability. The reasons are intertwined: GDPR's right-to-erasure conflicts directly with naive interpretations of blockchain immutability; HIPAA's accounting-of-disclosures rule maps awkwardly onto cross-jurisdictional federated rounds; and the European Health Data Space regulation introduces fresh obligations whose technical implications are still being clarified. Practical responses include keeping personal data off-chain entirely, using cryptographic commitments rather than plaintext records on-chain, and implementing redaction-aware ledger designs in permissioned settings. None of these is a complete solution; each represents a point on a trade-off curve that legal teams, ethics committees, and engineering teams must navigate together (Lu, 2018; Lu, 2022).

Interoperability is closely related but technically distinct. A BC-FL deployment that ignores HL7 FHIR is unlikely to be adopted in a hospital because it cannot exchange data with the surrounding clinical infrastructure. Recent work has explored binding FHIR resource versions to on-chain hashes, providing tamper-evident references to clinical artefacts without putting those artefacts on-chain (Hölbl et al., 2018; Khatoon, 2020). This pattern is promising but requires governance about who can issue and revoke such bindings, and again the answers are organisational as much as technical.

Scalability and latency are the next most pressing issues. Hospital networks operate at human time scales for many tasks, but real-time alerting, intra-operative decision support, and continuous monitoring require sub-second responses that are difficult to deliver from blockchain-mediated control paths. The accepted resolution is to keep the inference path off-chain, as in Figure 1, and to engage the chain only at boundaries: round commit, model release, audit query (Dorri et al., 2019). Even with this discipline, the throughput limits of permissioned blockchains constrain how often training rounds can be held; protocols that batch many updates into a single committed block are an important coping strategy.

Energy and compute cost are particularly relevant in resource-constrained or low- and middle-income settings, where the total cost of ownership of a BC-FL system may decisively shape adoption. The combined energy budget of edge inference, local training, secure aggregation, and consensus participation can exceed what a small clinic is willing or able to commit. Lightweight cryptographic primitives, model compression, and stake-delegation schemes that allow most clients to bear minimal consensus load are all active areas of work (Lu et al., 2024).

Clinician workflow integration is the barrier that is often underestimated in technical papers. A system that asks clinicians to sign extra consent dialogues, manage cryptographic keys, or interpret blockchain-specific error states will be quietly abandoned. The most successful deployments push these concerns to administrative back-office staff or to the IT department, and they ensure that the clinical interface looks identical whether or not BC-FL is operating beneath it. Patient consent management interacts with this design choice: dynamic consent models, where patients can update their participation preferences over time, are well aligned with smart-contract automation but require careful interface design to avoid bewildering users.

Adversarial robustness is the seventh major barrier, and arguably the one most likely to grow in importance. As BC-FL systems leave the laboratory and enter clinical use, adversaries with concrete incentives, including financially motivated attackers, insider threats, and even nation-state actors, will probe them. Recent work has demonstrated practical poisoning attacks against federated medical models, attacks that bypass naive defences and require active mitigation (Lu et al., 2024; Wu et al., 2025). Robust aggregation, reputation systems, and on-chain anomaly detection collectively raise the bar, but the field has not yet converged on a defence-in-depth blueprint that is both proven against adaptive attackers and tolerable to operate in clinical settings.

5.2. Governance and Trust Design

A theme that cuts across the technical barriers is governance. Who joins the network? Who validates? Who decides when to retire a model or roll back an update? In permissioned BC-FL deployments these

decisions are typically made by a consortium of participating institutions, but the consortium itself is a sociotechnical artefact whose constitution materially affects how the system behaves. Studies that report only on technical performance without explaining governance arrangements should be read with caution; conversely, studies that take governance seriously often do so at the cost of a less novel-looking technical contribution. Future work should therefore make governance a first-class concept in BC-FL evaluation, perhaps by adopting reporting standards analogous to those used in clinical-trial registration (Zhang & Lu, 2025; Zhang & Lu, 2021).

Trust design is the closely related question of how human stakeholders come to rely on the system. Clinicians need to trust that the model they deploy is the model that was validated; patients need to trust that their data are used as agreed; regulators need to trust that audit trails can be produced on demand. Blockchain provides cryptographic evidence in support of these trust claims, but evidence is not the same as trust. The translation of evidence into trust runs through documentation, training, and the lived experience of the system over time. The most successful deployments invest deliberately in this translation rather than treating it as someone else's problem (Yang et al., 2025).

5.3. Limitations of the Review and Open Questions

This review has limitations of its own. Our corpus, although chosen carefully, is biased toward English-language indexed venues and toward computer-science publication norms. Important contributions in clinical informatics journals, in regional venues, and in regulatory grey literature are under-represented. We did not perform meta-analytic synthesis of effect sizes, for the reasons given in Section 3, and the qualitative scoring in Figure 4 reflects the authors' judgement rather than a formal Delphi exercise. Readers interested in specific algorithmic comparisons should consult primary sources for full numerical detail.

Several open questions emerge. First, the absence of agreed evaluation protocols makes cross-study comparison difficult; a community benchmark, with shared datasets, threat models, and reporting templates, would dramatically accelerate the field. Second, the integration of BC-FL with newer paradigms, including foundation models and large multimodal models, remains essentially unexplored; current FL theory does not extend straightforwardly to billion-parameter models. Third, the long-run sustainability of permissioned consortium chains is uncertain, particularly in settings where institutional participation is voluntary and economic incentives are weak. Finally, the ethical and legal frame around BC-FL in low-resource settings deserves explicit attention, because designs that work in well-resourced hospital networks may fail or actively harm in clinics that lack reliable connectivity, robust IT staffing, or strong patient-rights advocacy (Lu, 2025; Yang et al., 2025).

6. Conclusion

Blockchain-assisted federated learning has moved from speculation to a recognised design pattern for trustworthy medical artificial intelligence in less than a decade. By holding raw patient data inside institutional perimeters while anchoring the model life cycle in an auditable, replicated ledger, BC-FL addresses two long-standing weaknesses of healthcare data analytics simultaneously. The reviewed literature shows clear progress in algorithmic design, with mature patterns emerging around permissioned

consensus, layered privacy mechanisms, and robust aggregation. Clinical scenarios from multi-site imaging to wearable-based remote monitoring have demonstrated that the paradigm is not merely theoretical, and that real benefits in privacy, auditability, and cross-institutional collaboration are achievable today.

Substantial work remains. Regulatory ambiguity, interoperability with established clinical standards, scalability under real-time constraints, energy footprint, clinician workflow burden, patient consent management, and adversarial robustness all rank as serious obstacles, and they will not yield to algorithmic ingenuity alone. The next phase of BC-FL research should engage governance, standards bodies, and clinical practice as first-class concerns, alongside the technical innovation that the field has already amply demonstrated. With that broader engagement, BC-FL is well positioned to become part of the routine fabric of medical AI rather than an exotic add-on, and it offers a credible path to delivering on the promise of data-driven medicine without abandoning the privacy, accountability, and trust on which clinical practice ultimately depends.

Ethics approval and consent to participate

Not applicable. This is a review article and did not involve direct human or animal participants.

Consent for publication

Not applicable.

Availability of data and materials

The structured extraction sheet supporting this review is available from the corresponding author upon reasonable request. All cited primary literature is publicly accessible through the digital object identifiers listed in the References section.

Funding

This work received no specific external funding. Open-access publication costs were supported by the institutional research funds of the participating universities.

Competing interests

The authors declare no competing financial or non-financial interests.

AI use disclosure

The authors used a large language model assistant for language polishing and reference formatting. All scientific content, analysis, interpretation, and conclusions are the work of the human authors. No AI tool is listed as an author.

Author contributions

Conceptualisation: M.A.E.-S. and Y.K.H.; Methodology: M.A.E.-S., H.M.A., and K.I.; Literature curation and screening: H.M.A. and K.I.; Formal analysis and synthesis: M.A.E.-S. and Y.K.H.; Visualisation: K.I. and M.A.E.-S.; Writing - original draft: M.A.E.-S. and Y.K.H.; Writing - review and editing: all authors; Supervision and project administration: Y.K.H.

Acknowledgements

The authors thank colleagues at the participating departments for constructive discussion of earlier drafts, and the anonymous reviewers whose comments materially improved the manuscript.

References

Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (pp. 308-318). https://doi.org/10.1145/2976749.2978318

Agbo, C. C., Mahmoud, Q. H., & Eklund, J. M. (2019). Blockchain technology in healthcare: A systematic review. Healthcare, 7(2), 56. https://doi.org/10.3390/healthcare7020056

Aich, S., Sinai, N. K., Kumar, S., Ali, M., Choi, Y. R., Joo, M. I., & Kim, H. C. (2022). Protecting personal healthcare records using blockchain and federated learning. In 2022 24th International Conference on Advanced Communication Technology (pp. 109-112). IEEE. https://doi.org/10.23919/ICACT53585.2022.9728772

Ali, A., Pasha, M. F., Ali, J., Fang, O. H., Masud, M., Jurcut, A. D., & Alzain, M. A. (2022). Deep learning based homomorphic secure search-able encryption for keyword search in blockchain healthcare system: A novel approach to cryptography. Sensors, 22(2), 528. https://doi.org/10.3390/s22020528

Azaria, A., Ekblaw, A., Vieira, T., & Lippman, A. (2016). MedRec: Using blockchain for medical data access and permission management. In 2nd International Conference on Open and Big Data (pp. 25-30). IEEE. https://doi.org/10.1109/OBD.2016.11

Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., Kiddon, C., Konečný, J., Mazzocchi, S., McMahan, H. B., Van Overveldt, T., Petrou, D., Ramage, D., & Roselander, J. (2019). Towards federated learning at scale: System design. Proceedings of Machine Learning and Systems, 1, 374-388. https://doi.org/10.48550/arXiv.1902.01046

Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S., Ramage, D., Segal, A., & Seth, K. (2017). Practical secure aggregation for privacy-preserving machine learning. In Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (pp. 1175-1191). https://doi.org/10.1145/3133956.3133982

Brisimi, T. S., Chen, R., Mela, T., Olshevsky, A., Paschalidis, I. C., & Shi, W. (2018). Federated learning of predictive models from federated electronic health records. International Journal of Medical Informatics, 112, 59-67. https://doi.org/10.1016/j.ijmedinf.2018.01.007

Castro, M., & Liskov, B. (1999). Practical Byzantine fault tolerance. In Proceedings of the Third Symposium on Operating Systems Design and Implementation (Vol. 99, pp. 173-186). USENIX Association. https://doi.org/10.5555/296806.296824

Chen, Y., Qin, X., Wang, J., Yu, C., & Gao, W. (2020). FedHealth: A federated transfer learning framework for wearable healthcare. IEEE Intelligent Systems, 35(4), 83-93. https://doi.org/10.1109/MIS.2020.2988604

Chen, Y., Lu, Y., Bulysheva, L., & Kataev, M. Y. (2024). Applications of blockchain in Industry 4.0: A review. Information Systems Frontiers, 26(5), 1715-1729. https://doi.org/10.1007/s10796-022-

10248-7

Christidis, K., & Devetsikiotis, M. (2016). Blockchains and smart contracts for the Internet of Things. IEEE Access, 4, 2292-2303. https://doi.org/10.1109/ACCESS.2016.2566339

Connolly, J., Lawless, S., Bonneau, R., Liu, J., & Sun, X. (2022). Healthchain: A blockchain-based framework for personal health record sharing with collaborative federated learning. Journal of Network and Computer Applications, 207, 103442. https://doi.org/10.1016/j.jnca.2022.103442

Dorri, A., Kanhere, S. S., Jurdak, R., & Gauravaram, P. (2019). LSB: A lightweight scalable blockchain for IoT security and anonymity. Journal of Parallel and Distributed Computing, 134, 180-197. https://doi.org/10.1016/j.jpdc.2019.08.005

Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. Foundations and Trends in Theoretical Computer Science, 9(3-4), 211-407. https://doi.org/10.1561/0400000042

Geyer, R. C., Klein, T., & Nabi, M. (2017). Differentially private federated learning: A client level perspective. arXiv preprint arXiv:1712.07557. https://doi.org/10.48550/arXiv.1712.07557

Hölbl, M., Kompara, M., Kamišalić, A., & Nemec Zlatolas, L. (2018). A systematic review of the use of blockchain in healthcare. Symmetry, 10(10), 470. https://doi.org/10.3390/sym10100470

Huang, L., Shea, A. L., Qian, H., Masurkar, A., Deng, H., & Liu, D. (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, 103291. https://doi.org/10.1016/j.jbi.2019.103291

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., et al. (2021). Advances and open problems in federated learning. Foundations and Trends in Machine Learning, 14(1-2), 1-210. https://doi.org/10.1561/2200000083

Khan, M. A., & Salah, K. (2018). IoT security: Review, blockchain solutions, and open challenges. Future Generation Computer Systems, 82, 395-411. https://doi.org/10.1016/j.future.2017.11.022

Khatoon, A. (2020). A blockchain-based smart contract system for healthcare management. Electronics, 9(1), 94. https://doi.org/10.3390/electronics9010094

Kim, H., Park, J., Bennis, M., & Kim, S.-L. (2020). Blockchained on-device federated learning. IEEE Communications Letters, 24(6), 1279-1283. https://doi.org/10.1109/LCOMM.2019.2921755

Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. IEEE Signal Processing Magazine, 37(3), 50-60. https://doi.org/10.1109/MSP.2020.2975749

Lu, Y. (2017). Industry 4.0: A survey on technologies, applications and open research issues. Journal of Industrial Information Integration, 6, 1-10. https://doi.org/10.1016/j.jii.2017.04.005

Lu, Y. (2018). Blockchain and the related issues: A review of current research topics. Journal of Management Analytics, 5(4), 231-255. https://doi.org/10.1080/23270012.2018.1516523

Lu, Y. (2019). The blockchain: State-of-the-art and research challenges. Journal of Industrial Information Integration, 15, 80-90. https://doi.org/10.1016/j.jii.2019.04.002

Lu, Y., & Xu, L. D. (2019). Internet of Things (IoT) cybersecurity research: A review of current research topics. IEEE Internet of Things Journal, 6(2), 2103-2115.

https://doi.org/10.1109/JIOT.2018.2869847

Lu, Y., Huang, X., Dai, Y., Maharjan, S., & Zhang, Y. (2020a). Blockchain and federated learning for privacy-preserved data sharing in industrial IoT. IEEE Transactions on Industrial Informatics, 16(6), 4177-4186. https://doi.org/10.1109/TII.2019.2942190

Lu, Y. (2022). Implementing blockchain in information systems: A review. Enterprise Information Systems, 16(12), 2008513. https://doi.org/10.1080/17517575.2021.2008513

Lu, Y., Sigov, A. S., Ratkin, L., Ivanov, L. A., & Zuo, M. (2024). Quantum computing and industrial information integration: A review. Journal of Industrial Information Integration, 35, 100511. https://doi.org/10.1016/j.jii.2023.100511

Lu, Y. (2025). The current status and developing trends of Industry 4.0: A review. Information Systems Frontiers, 27(1), 215-234. https://doi.org/10.1007/s10796-021-10221-w

McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. In Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (PMLR 54, pp. 1273-1282). https://doi.org/10.48550/arXiv.1602.05629

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. Bitcoin White Paper. https://doi.org/10.2139/ssrn.3440802

Nguyen, D. C., Pham, Q.-V., Pathirana, P. N., Ding, M., Seneviratne, A., Lin, Z., Dobre, O., & Hwang, W.-J. (2022). Federated learning for smart healthcare: A survey. ACM Computing Surveys, 55(3), 60. https://doi.org/10.1145/3501296

Nguyen, D. C., Ding, M., Pham, Q.-V., Pathirana, P. N., Le, L. B., Seneviratne, A., Li, J., Niyato, D., & Poor, H. V. (2021). Federated learning meets blockchain in edge computing: Opportunities and challenges. IEEE Internet of Things Journal, 8(16), 12806-12825. https://doi.org/10.1109/JIOT.2021.3072611

Passerat-Palmbach, J., Farnan, T., McCoy, M., Harris, J. D., Manion, S. T., Flannery, H. L., & Gleim, B. (2020). Blockchain-orchestrated machine learning for privacy preserving federated learning in electronic health data. In 2020 IEEE International Conference on Blockchain (pp. 550-555). IEEE. https://doi.org/10.1109/Blockchain50366.2020.00080

Pati, S., Baid, U., Edwards, B., Sheller, M., Wang, S.-H., Reina, G. A., et al. (2022). Federated learning enables big data for rare cancer boundary detection. Nature Communications, 13(1), 7346. https://doi.org/10.1038/s41467-022-33407-5

Pokhrel, S. R., & Choi, J. (2020). Federated learning with blockchain for autonomous vehicles: Analysis and design challenges. IEEE Transactions on Communications, 68(8), 4734-4746. https://doi.org/10.1109/TCOMM.2020.2990686

Reyna, A., Martín, C., Chen, J., Soler, E., & Díaz, M. (2018). On blockchain and its integration with IoT: Challenges and opportunities. Future Generation Computer Systems, 88, 173-190. https://doi.org/10.1016/j.future.2018.05.046

Rieke, N., Hancox, J., Li, W., Milletarì, F., Roth, H. R., Albarqouni, S., et al. (2020). The future of digital health with federated learning. NPJ Digital Medicine, 3(1), 119. https://doi.org/10.1038/s41746-020-00323-1

Roth, H. R., Chang, K., Singh, P., Neumark, N., Li, W., Gupta, V., et al. (2020). Federated learning for breast density classification: A real-world implementation. In Domain Adaptation and Representation Transfer, MICCAI 2020 (LNCS 12444, pp. 181-191). Springer. https://doi.org/10.1007/978-3-030-60548-3_18

Sharma, P., Borah, M. D., & Namasudra, S. (2021). Improving security of medical big data by using blockchain technology. Computers and Electrical Engineering, 96, 107529. https://doi.org/10.1016/j.compeleceng.2021.107529

Sheller, M. J., Edwards, B., Reina, G. A., Martin, J., Pati, S., Kotrotsou, A., Milchenko, M., Xu, W., Marcus, D., Colen, R. R., & Bakas, S. (2019). Multi-institutional deep learning modeling without sharing patient data: A feasibility study on brain tumor segmentation. In Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries (LNCS 11383, pp. 92-104). Springer. https://doi.org/10.1007/978-3-030-11723-8_9

Singh, S., Rathore, S., Alfarraj, O., Tolba, A., & Yoon, B. (2022). A framework for privacy-preservation of IoT healthcare data using federated learning and blockchain. Future Generation Computer Systems, 129, 380-388. https://doi.org/10.1016/j.future.2021.11.028

Truex, S., Baracaldo, N., Anwar, A., Steinke, T., Ludwig, H., Zhang, R., & Zhou, Y. (2019). A hybrid approach to privacy-preserving federated learning. In Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security (pp. 1-11). https://doi.org/10.1145/3338501.3357370

Vaid, A., Jaladanki, S. K., Xu, J., Teng, S., Kumar, A., Lee, S., et al. (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), e24207. https://doi.org/10.2196/24207

Wood, G. (2014). Ethereum: A secure decentralised generalised transaction ledger. Ethereum Project Yellow Paper, 151, 1-32. https://doi.org/10.5281/zenodo.7679126

Wu, H. P., Liu, Z., Dong, H. Y., Lu, Y., & Xu, L. D. (2025). Revolutionizing internal auditing: Harnessing the power of blockchain. Enterprise Information Systems, 19(1-2), 2448003. https://doi.org/10.1080/17517575.2024.2448003

Xu, L. D., Lu, Y., & Li, L. (2021). Embedding blockchain technology into IoT for security: A survey. IEEE Internet of Things Journal, 8(13), 10452-10473. https://doi.org/10.1109/JIOT.2021.3060508

Xu, R., Zhu, J., Yang, L., Lu, Y., & Xu, L. D. (2024). Decentralized finance (DeFi): A paradigm shift in the FinTech. Enterprise Information Systems, 18(9), 2397630. https://doi.org/10.1080/17517575.2024.2397630

Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. ACM Transactions on Intelligent Systems and Technology, 10(2), 12. https://doi.org/10.1145/3298981

Yang, L., Hou, Q., Zhu, X., Lu, Y., & Xu, L. D. (2025). Potential of large language models in blockchain-based supply chain finance. Enterprise Information Systems, 19(11), 2541199. https://doi.org/10.1080/17517575.2025.2541199

Zhang, C., & Lu, Y. (2021). Study on artificial intelligence: The state of the art and future prospects. Journal of Industrial Information Integration, 23, 100224. https://doi.org/10.1016/j.jii.2021.100224

Zhang, H., & Lu, Y. (2025). Web 3.0: Applications, opportunities and challenges in the next internet

generation. Systems Research and Behavioral Science, 42(4), 996-1015. https://doi.org/10.1002/sres.3022

Zheng, X. R., & Lu, Y. (2022). Blockchain technology: Recent research and future trend. Enterprise Information Systems, 16(12), 1939895. https://doi.org/10.1080/17517575.2021.1939895