Eur. Phys. J. Spec. Top.
https://doi.org/10.1140/epjs/s11734-026-02475-9

THE EUROPEAN PHYSICAL JOURNAL SPECIAL TOPICS

Review

Federated learning in healthcare: a review of the deployment gap between simulated and real-world implementations

Elena Pitsik[1,a], Ivan Shishkin[2,b], Alsu Ivanova[2,c], Nikita Muchanko[2,d], Ilya Khanchuk[3,e], Fu Xiren[2,f], Denis Andrikov[1,g], and Alexander Hramov[1,h]

[1] Research Institute of Applied Artificial Intelligence and Digital Solutions, Plekhanov Russian University of Economics, Stremyanny per., 36, Moscow, Russia115054
[2] Engineering Academy, Peoples' Friendship University of Russia named after Patrice Lumumba, Miklukho-Maklaya St., 6, Moscow, Russia117198
[3] Advanced Engineering School "Intelligent Theranostic Systems", Sechenov University, Trubetskaya St., 8s2, Moscow, Russia119991

Received 28 May 2026 / Accepted 19 June 2026

Abstract Federated learning (FL) has emerged as a promising paradigm for privacy-preserving collaborative model training in healthcare, yet a pronounced gap persists between its performance in simulated research settings and genuine clinical deployments. This scoping review, conducted following PRISMA-ScR guidelines, systematically surveyed FL research in healthcare indexed in PubMed from 2016 onward. Of 1,338 initially identified papers, 772 met inclusion criteria, of which only 25 (3.2%) represented genuine real-world FL deployments. Real-world deployments generally demonstrated near-equivalent performance to centralized baselines, though practical barriers, such as infrastructure complexity, data heterogeneity, partial labeling, and unvalidated privacy mechanisms, remain pervasive. These findings underscore that the federated learning deployment gap is primarily an infrastructural and organizational challenge rather than an algorithmic one, and identify key conditions required for broader clinical adoption.

1 Introduction

Despite the transformative potential of machine learning (ML) and artificial intelligence (AI) in healthcare—from disease diagnosis and prognostication to treatment personalization [1, 2]—a fundamental bottleneck remains: access to large, diverse, yet privacy-sensitive clinical data. Healthcare data are inevitably siloed across institutions, each with its own infrastructure, annotation standards, and regulatory constraints [3]. The rapid expansion of AI in medicine is not only technical but also reflected in publication activity; a network-based analysis by Karpov et al. [4] revealed distinct research clusters and emerging trends, confirming that AI for clinical applications has become a major scientific domain. Beyond traditional clinical domains, AI methods have also been extensively applied to mental health—from diagnosing conditions to supporting rehabilitation as recently reviewed by Khorev et al.

Elena Pitsik, Ivan Shishkin, Alsu Ivanova, Nikita Muchanko, Ilya Khanchuk, Fu Xiren, Denis Andrikov and Alexander Hramov have contributed equally to this work.

[a] e-mail: pitsikelena@gmail.com (corresponding author)
[b] e-mail: ivan081272@yandex.ru
[c] e-mail: aliva00alsu@gmail.com
[d] e-mail: much.nikita@gmail.com
[e] e-mail: second.originaccount.00@mail.ru
[f] e-mail: satoandrea47@gmail.com
[g] e-mail: andrikov@bmstu.ru
[h] e-mail: hramovae@gmail.com

Published online: 28 June 2026

[5]. These advances, however, are constrained by the difficulty of aggregating data from multiple sites without violating patient privacy.

Federated learning (FL), first formally introduced by McMahan et al. [6], offers a compelling resolution to this tension. FL enables a model to be trained collaboratively across decentralized nodes without requiring direct data exchange. This paradigm shift has rapidly attracted interest from both academia and industry [7, 8], and is widely regarded as one of the most promising frameworks for privacy-preserving machine learning at scale.

Yet a pronounced deployment gap has emerged between federated learning as studied in controlled research settings and its actual deployment in clinical environments. The majority of published FL studies in healthcare rely on simulated federation—partitioning centralized public datasets such as MIMIC-III or ABIDE across artificial "clients" to mimic a multi-institutional scenario [9–11]. While methodologically convenient, this approach fails to capture the complexity of real-world deployment: heterogeneous institutional infrastructure, differences in data collection protocols, inconsistent annotation standards, variable network conditions, and unvalidated privacy guarantees [12]. As this review shows, only 3.2% of FL studies in healthcare represent genuine real-world deployments across independently operating institutional infrastructure. The remaining 96.8% rely on simulated settings. As a result, models that perform well in simulation frequently face significant degradation or practical infeasibility when confronted with genuine federated conditions [13, 14].

The systematic discrepancy between simulated and real-world FL performance and applicability—the deployment gap—has received growing attention [12]. However, its causes, manifestations, and mitigations are still being actively studied. As illustrated in Fig. 1, simulated studies operate under idealized conditions (centralized data partitions, controlled heterogeneity, perfect networks, identical preprocessing), whereas genuine deployments confront infrastructural complexity, natural domain shifts, partial labeling, IT constraints, and unvalidated privacy mechanisms. This figure highlights that the deployment gap is primarily infrastructural and organizational rather than algorithmic, and that only a tiny fraction (3.2%) of published research reflects real-world conditions.

This review focuses on key FL frameworks and their applications under both simulated and real-world conditions. We discuss various implementations of FL, their recurring limitations, and survey genuine real-world FL deployments in order to provide a clear view of the current state of the field, its unresolved problems, and the conditions required for FL to fulfill its potential as a widespread tool in healthcare.

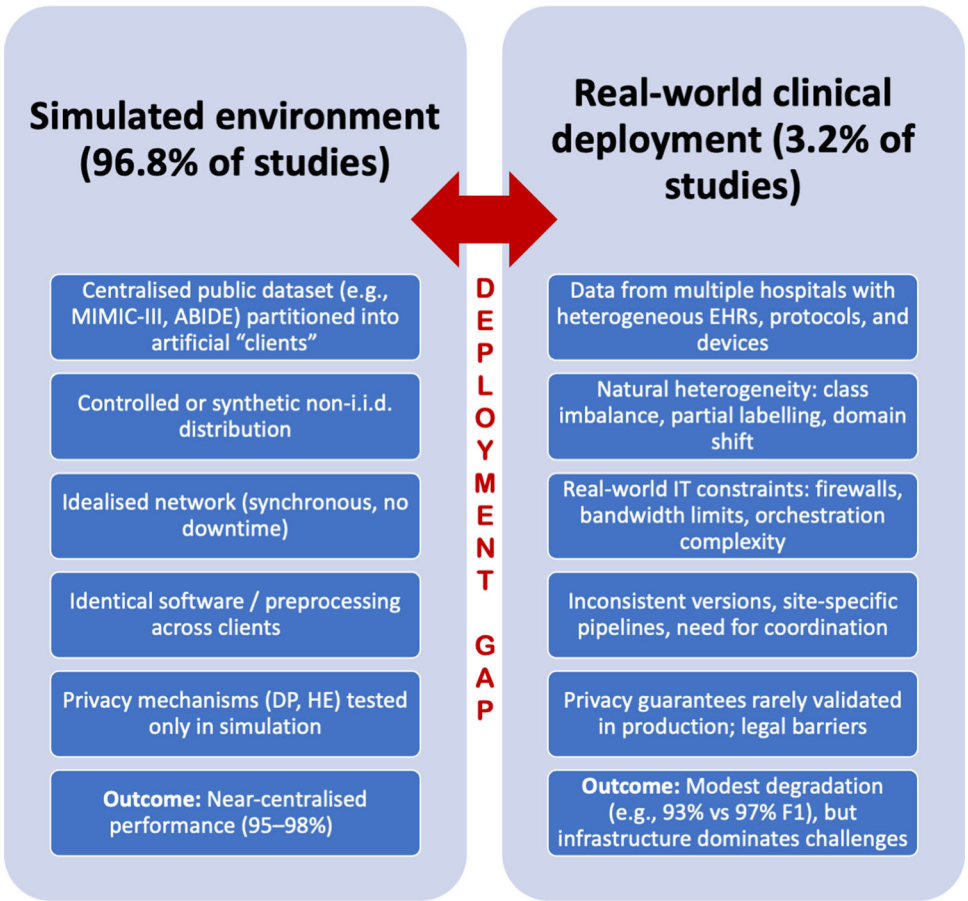

Fig. 1 The FL deployment gap: simulated versus real-world healthcare settings. The left column summarizes typical conditions in simulated studies (96.8% of the literature), while the right column shows challenges encountered in genuine clinical deployments (3.2%). The gap is characterized by infrastructural complexity, data heterogeneity, partial labeling, and unvalidated privacy mechanisms

This review is structured as follows. Section 2 describes our PRISMA-ScR compliant methods. Section 3 presents the results: the overwhelming dominance of simulated studies (Sect. 3.1), the performance of genuine real-world deployments (Sect. 3.2), and a thematic analysis of recurring limitations — infrastructure, data heterogeneity, and privacy mechanisms (Sect. 3.3). Section 4 concludes with actionable recommendations for closing the deployment gap.

2 Methods

2.1 Inclusion criteria

The present scoping review was guided by the standards of the Preferred Reporting Items for Scoping Reviews (PRISMA) Statement [15]. Following inclusion criteria, in order to be included into this review, the paper should be:

1. Published in peer-review academic journal indexing in PubMed database;
2. Published not earlier than 2016;
3. Written in English;
4. Presenting an FL framework, system, or application in a healthcare context in simulated and/or real-world deployment.

We excluded conference proceedings, books, theses, reviews and pre-prints. A small number of highly relevant conference contributions identified during full-text review of cited/related work were retained as supporting evidence, but excluded from the final count of review results.

2.2 Search query

To make a distinction between real-world applications and simulated environment studies, we employed a block-based search query construction approach. The first block focuses the search on the federated learning:

(`federated learning` **OR** `federated machine learning` **OR** `federated` **OR** `federated deep learning` **OR** `federated training`)

The second block introduced the healthcare context:

(`healthcare` **OR** `clinical` **OR** `medical` **OR** `hospital` **OR** `patient data` **OR** `electronic health record` **OR** `medical imaging` **OR** `clinical trial`)

Using `AND` logical operator, we ensured that the search results contained research on FL in healthcare.

The aim of the present review, however, implies division of these results into two groups: real-world deployment of FL and simulated studies. This represents a challenge, as authors rarely report the simulated environment explicitly. In the attempt to review the struggles of real-world deployments of FL, the categorization of real-world deployment articles was performed manually by abstracts and full-text screening. The criteria for real-world deployment was determined as follows: FL system must involve independently operating institutional infrastructure, as opposed to a single centrally held dataset artificially partitioned into simulated clients.

The initial screening was performed using ASReview [16], and the results were revised manually, since not all articles clearly report a deployment type in abstracts.

3 Results and discussion

3.1 Query summary

The initial search yielded 1338 papers on federated learning in healthcare. 566 papers were excluded at initial screening due to compliance with one of the exclusion criteria. Among the rest 772 papers, 25 papers representing on-site deployment were manually selected.

Thematic grouping of included papers revealed three largest clusters of papers with the prevalence of medical imaging and segmentation-based research (118 papers). 53 papers are dedicated to development of new aggregation algorithms and 50 papers use EHR data modality for clinical prediction. The remaining included papers were distributed across smaller and more heterogeneous thematic groups that were not individually broken out in this

report. Therefore, only 3.2% of all FL research in healthcare found in this review represent genuine real-world deployment of FL architecture.

3.2 Real-world deployments

Evidence from genuinely real-world FL deployments suggests that federated models rarely match centralized, data-pooled baselines exactly, but the magnitude and direction of the gap vary considerably by task and setting. The most frequently observed outcome is near-equivalence. For instance, [17] reported that the FL model achieved non-inferior classification accuracy relative to the centralized comparator (center-level accuracies 78.3–98.5%), while [18] found that a federated model predicting coronary artery calcification scores reached sensitivity 67% and specificity 69%, compared with 66% and 70% for the centralized equivalent. Similarly, [19] demonstrated that FL-based re-training of an automated echocardiography measurement tool reduced measurement deviations from human reference values to non-significance in all but two of the evaluated cardiac parameters, which is a comparable result to centralized re-training. Where a genuine performance cost is detectable, it is modest: [20], deploying FL on real heterogeneous IoT edge devices for ECG arrhythmia detection, reported an F1-score of 93% under federation versus 97% for a centralized model. This four percentage point deficit the authors explicitly frame as an acceptable trade-off given the gains in privacy, scalability, and practical feasibility.

Importantly, this is one of the very few real-world studies to report this gap honestly rather than treating it as negligible. [21], conducting real-world lung pathology segmentation experiments across six German university hospitals within the RACOON network, reported that the FL model outperformed all less complex alternatives, including pseudo-centralisation approaches, suggesting that in sufficiently heterogeneous real-world settings, FL can actually exceed what a naive centralized approach achieves when data pooling is imperfect or when domain shift between sites is pronounced.

3.3 Limitations and challenges

Reported limitations could be clustered into four recurring categories across the reviewed studies.

3.3.1 Infrastructure and deployment complexity

Unlike centralized architectures, FL requires setting up the central server, managing communication protocols across several local clients, and ensuring consistent preprocessing pipelines across clients, which required substantial coordination effort [17]. In [21], authors report a detailed taxonomy of practical hurdles encountered during real-world FL deployment at six German university hospitals, including orchestration complexity, logging and experiment management, and the need for IT expertise at each site. Authors of [22] report that hospital and corporate network restrictions required an open port to be configured on the FL server, which was resolved by deploying the server on AWS in a semi-public network. Long experiment durations were also a practical issue, limiting the frequency of experimental iterations. Various attempts to reduce deployment barriers were made, including leveraging pre-configured Raspberry Pi microcomputers [23].

3.3.2 Data heterogeneity and partial labeling

Partial labeling and data heterogeinity represent a challenge for FL models for several reasons. Authors of [24] directly address partial labeling as a structural challenge: different hospitals had only annotated data relevant to their own research focus, leaving large volumes of unlabelled data unused. In FL setting, local models do not 'see' a complete picture of the input during training, placing pressure on the aggregation function [19, 25]. Also, real-world heterogeneity causes substantial variation in local model performance [26] – an effect rarely present in simulated studies.

3.3.3 Data privacy

A considerable body of research focuses on enhancing privacy in FL systems through privacy-preserving technologies integrated directly into the learning pipeline. Among all included papers, 228 articles (29.5%) mentioned privacy as a design objective, of which 85 (11.0%) proposed or evaluated a specific privacy-enhancing mechanism such as differential privacy, homomorphic encryption, or secure aggregation.

It should be noted that 'privacy' in this context refers specifically to privacy-preserving mechanisms operating during the federated training process itself rather than to data anonymisation or de-identification procedures applied prior to data sharing, which fall outside the scope of the FL-specific mechanisms discussed below.

Differential privacy (DP) is the most frequently employed formal privacy mechanism in the reviewed literature. The dominant implementation is differentially private stochastic gradient descent (DPSGD), in which calibrated Gaussian or Laplace noise is injected into model gradients before aggregation, bounding the influence of any individual training example and providing a quantifiable privacy guarantee expressed through the privacy budget. Applications span a wide range of clinical tasks, most prominently COVID-19 detection from chest X-ray and CT imaging, breast cancer classification, polyp and surgical instrument segmentation, PET artifact correction, arrhythmia diagnosis from ECG signals, multi-omics survival analysis, and postoperative mortality prediction in colorectal surgery. A recurring finding across these studies is the privacy-utility trade-off. In particular, tighter privacy budgets reliably degrade model performance, sometimes substantially [27, 28].

Several papers propose technical mitigations for this trade-off. Knowledge distillation combined with ensemble weighting is used to recover utility under DP noise constraints [29–31]. A minority of papers employ local differential privacy (LDP), in which noise is applied at the client before transmission rather than centrally [32]. The multicenter colorectal surgery mortality study [33] and the multi-institutional PET imaging paper [34] describe genuinely distributed deployments, underscoring that formal privacy guarantees remain largely unvalidated in production healthcare environments.

Homomorphic encryption (HE) represents a cryptographically stronger privacy guarantee than DP. Rather than adding noise to gradients, HE allows computations to be performed directly on encrypted model updates, so that the aggregating server never observes plaintext weights. The most commonly used scheme is Paillier encryption, a partial HE scheme supporting addition over ciphertexts, which is sufficient for FedAvg-style weighted averaging and is employed by papers including the Paillier-based FL-SSL framework for COVID-19 CT classification [35] and the PriCell framework for disease-associated cell classification [36].

More recent contributions extend this to fully homomorphic or dynamic HE. FedGraphHE [37] proposes dynamic HE tailored for federated graph neural networks in smart healthcare networks, and Health-FedNet [38] combines HE with differential privacy for chronic disease prediction on MIMIC-III, representing one of the few papers to layer both mechanisms.

A consistent and openly acknowledged limitation across HE-based FL papers is computational overhead. Encrypting, transmitting, and aggregating ciphertext model updates introduces substantial latency and memory cost relative to plaintext FL. The paper [39] directly benchmarks HE against DP and standard FL on a cloud training scenario, concluding that HE suffers from prohibitive overhead when model updates are independently encrypted by numerous clients, while DP introduces accuracy degradation. Addressing the overhead problem is a central motivation for several papers. In particular, a lightweight HE-FL framework for skin cancer diagnosis [40] and an efficient HE approach for diabetic retinopathy classification [41] both explicitly prioritize computational efficiency as a design objective, proposing optimized encryption pipelines that reduce latency while preserving privacy guarantees. Split federated learning combined with HE [42] further attempts to redistribute the computational burden between clients and the server. As with DP, all HE-FL studies in this corpus operate in simulated or laboratory environments; none describes a production deployment across independently operating hospital infrastructure.

Another privacy-preserving technology, largest privacy-adjacent category in the corpus, and also the most heterogeneous in terms of its functional role within the FL pipeline, is blockchain. Unlike DP and HE, which protect the privacy of gradient updates during aggregation, blockchain primarily addresses trust, auditability, and model integrity. The clinical applications in this group are notably concentrated around COVID-19 detection and IoT-based healthcare [43], with numerous papers combining FL, blockchain, and edge computing for pandemic surveillance, CT-based COVID diagnosis, lung cancer detection, and wearable monitoring [44–46]. A notable subset positions blockchain as a mechanism for ensuring equitable model contribution and transparent governance across institutions, extending its role beyond privacy into accountability [47, 48].

4 Conclusion

This scoping review confirms that federated learning in healthcare remains predominantly a simulation-based research endeavor: only 3.2% of reviewed studies represent genuine deployments across independently operating institutional infrastructure. Where real-world deployments do exist, federated models generally achieve performance comparable to centralized baselines, suggesting that algorithmic maturity is not the primary barrier to adoption. Instead, the evidence consistently points to infrastructural complexity, cross-site data heterogeneity, and inconsistent annotation standards as the dominant challenges. Privacy-preserving mechanisms, while extensively studied, remain largely unvalidated in production settings. For FL to fulfill its potential as a widespread tool in healthcare, future work must prioritize reproducible deployment frameworks, standardized communication protocols, and prospective evaluation under genuine federated conditions, moving the field from proof-of-concept simulation toward accountable clinical practice.

Data availability Data sharing is not applicable to this article as no new datasets were generated or analyzed during the current study.

References

1. E.J. Topol, High-performance medicine: the convergence of human and artificial intelligence. Nat. Med. **25**(1), 44–56 (2019)
2. P. Rajpurkar, E. Chen, O. Banerjee, E.J. Topol, Ai in health and medicine. Nat. Med. **28**(1), 31–38 (2022)
3. W.N. Price, I.G. Cohen, Privacy in the age of medical big data. Nat. Med. **25**(1), 37–43 (2019)
4. O.E. Karpov, E.N. Pitsik, S.A. Kurkin, V.A. Maksimenko, A.V. Gusev, N.N. Shusharina, A.E. Hramov, Analysis of publication activity and research trends in the field of ai medical applications: network approach. Int. J. Environ. Res. Public Health **20**(7), 5335 (2023)
5. V. Khorev, A. Kiselev, A. Badarin, V. Antipov, O. Drapkina, S. Kurkin, A. Hramov, Review on the use of ai-based methods and tools for treating mental conditions and mental rehabilitation. The Eur. Phys. J. Spec. Top. **234**(15), 4139–4158 (2025)
6. B. McMahan, E. Moore, D. Ramage, S. Hampson, B.A. Arcas, Communication-efficient learning of deep networks from decentralized data. In: Artificial Intelligence and Statistics, pp. 1273–1282 (2017). Pmlr
7. Q. Yang, Y. Liu, T. Chen, Y. Tong, Federated machine learning: concept and applications. ACM Trans. Intell. Syst. Technol. (TIST) **10**(2), 1–19 (2019)
8. P. Kairouz, H.B. McMahan, Advances and open problems in federated learning. Found. Trends Mach. Learn. **14**(1–2), 1–210 (2021)
9. G.H. Lee, S.-Y. Shin, Federated learning on clinical benchmark data: performance assessment. J. Med. Internet Res. **22**(10), 20891 (2020)
10. S. Li, D. Miao, Q. Wu, C. Hong, D. D'Agostino, X. Li, Y. Ning, Y. Shang, Z. Wang, M. Liu, Federated learning in healthcare: a benchmark comparison of engineering and statistical approaches for structured data analysis. Health Data Sci. **4**, 0196 (2024)
11. X. Li, Y. Gu, N. Dvornek, L.H. Staib, P. Ventola, J.S. Duncan, Multi-site fmri analysis using privacy-preserving federated learning and domain adaptation: abide results. Med. Image Anal. **65**, 101765 (2020)
12. Z.L. Teo, L. Jin, N. Liu, S. Li, D. Miao, X. Zhang, W.Y. Ng, T.F. Tan, D.M. Lee, K.J. Chua, Federated machine learning in healthcare: a systematic review on clinical applications and technical architecture. Cell Rep. Med. **5**(2), 101419 (2024)
13. M. Li, P. Xu, J. Hu, Z. Tang, G. Yang, From challenges and pitfalls to recommendations and opportunities: implementing federated learning in healthcare. Med. Image Anal. **101**, 103497 (2025)
14. J. Terrail, S.-S. Ayed, E. Cyffers, F. Grimberg, C. He, R. Loeb, P. Mangold, T. Marchand, O. Marfoq, E. Mushtaq, Flamby: datasets and benchmarks for cross-silo federated learning in realistic healthcare settings. Adv. Neural. Inf. Process. Syst. **35**, 5315–5334 (2022)
15. D. Moher, A. Liberati, J. Tetzlaff, D.G. Altman, Preferred reporting items for systematic reviews and meta-analyses: the prisma statement. BMJ **339**, b2535 (2009)
16. R. Van De Schoot, J. De Bruin, R. Schram, P. Zahedi, J. De Boer, F. Weijdema, B. Kramer, M. Huijts, M. Hoogerwerf, G. Ferdinands, An open source machine learning framework for efficient and transparent systematic reviews. Nat. Mach. Intell. **3**(2), 125–133 (2021)
17. A.R. Ran, X. Wang, P.P. Chan, M.O. Wong, H. Yuen, N.M. Lam, N.C. Chan, W.W. Yip, A.L. Young, H.-W. Yung, Developing a privacy-preserving deep learning model for glaucoma detection: a multicentre study with federated learning. Br. J. Ophthalmol. **108**(8), 1114–1123 (2024)
18. J. Wolff, J. Matschinske, D. Baumgart, A. Pytlik, A. Keck, A. Natarajan, C.E. Schacky, J.K. Pauling, J. Baumbach, Federated machine learning for a facilitated implementation of artificial intelligence in healthcare-a proof of concept study for the prediction of coronary artery calcification scores. J. Integr. Bioinform. **19**(4), 20220032 (2022)
19. C. Morbach, G. Gelbrich, M. Schreckenberg, M. Hedemann, D. Pelin, N. Scholz, O. Miljukov, A. Wagner, F. Theisen, N. Hitschrich, Population data-based federated machine learning improves automated echocardiographic quantification of cardiac structure and function: the automatisierte vermessung der echokardiographie project. Eur. Heart J-Dig. Health **5**(1), 77–88 (2024)
20. A.T. Shumba, D. Cantoro, T. Montanaro, G. Semeraro, I. Sergi, M. De Vittorio, L. Patrono, A laboratory-based federated learning deployment on real devices for ecg-based clinical decision support systems, in: 2025 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). (IEEE), pp.1–6(2025)
21. M.R. Bujotzek, Ü. Akünal, S. Denner, P. Neher, M. Zenk, E. Frodl, A. Jaiswal, M. Kim, N.R. Krekiehn, M. Nickel, Real-world federated learning in radiology: hurdles to overcome and benefits to gain. J. Am. Med. Inform. Assoc. **32**(1), 193–205 (2025)
22. L.A. Schoenpflug, R.B. Benavides, M. Nowak, F. Sheikhzadeh, A. Moayyedi, K. Wasag, J. Reimers, M. Zhou, R. Venugopal, B. Sobottka, Navigating real-world challenges: a case study on federated learning in computational pathology. J. Pathol. Inform. **18**, 100464 (2025)

23. A.A. Soltan, A. Thakur, J. Yang, A. Chauhan, L.G. D'Cruz, P. Dickson, M.A. Soltan, D.R. Thickett, D.W. Eyre, T. Zhu, A scalable federated learning solution for secondary care using low-cost microcomputing: privacy-preserving development and evaluation of a covid-19 screening test in uk hospitals. The Lancet Dig. Health **6**(2), 93–104 (2024)
24. M. Tölle, P. Garthe, C. Scherer, J.M. Seliger, A. Leha, N. Krüger, S. Simm, S. Martin, S. Eble, H. Kelm, Real world federated learning with a knowledge distilled transformer for cardiac ct imaging. npj Dig. Med. **8**(1), 88 (2025)
25. K.V. Sarma, S. Harmon, T. Sanford, H.R. Roth, Z. Xu, J. Tetreault, D. Xu, M.G. Flores, A.G. Raman, R. Kulkarni, Federated learning improves site performance in multicenter deep learning without data sharing. J. Am. Med. Inform. Assoc. **28**(6), 1259–1264 (2021)
26. S. Fu, H. Jia, M. Vassilaki, V.K. Keloth, Y. Dang, Y. Zhou, M. Garg, R.C. Petersen, J. St Sauver, S. Moon, Fedfsa: hybrid and federated framework for functional status ascertainment across institutions. J. Biomed. Inform. **152**, 104623 (2024)
27. M. Adnan, S. Kalra, J.C. Cresswell, G.W. Taylor, H.R. Tizhoosh, Federated learning and differential privacy for medical image analysis. Sci. Rep. **12**(1), 1953 (2022)
28. A. Qayyum, K. Ahmad, M.A. Ahsan, A. Al-Fuqaha, J. Qadir, Collaborative federated learning for healthcare: multi-modal covid-19 diagnosis at the edge. IEEE Open J. Comput. Soc. **3**, 172–184 (2022)
29. R. Annan, H. Qin, R. Newman, M. Siddula, L. Qingge, Privacy-preserving federated learning with optimized ensemble weighting and knowledge distillation for covid-19 detection from non-iid medical imaging data. Scientific Reports (2026)
30. S.M. Bokhari, S. Sohaib, M. Shafi, Fusion of personalized federated learning (pfl) with differential privacy (dp) learning for diagnosis of arrhythmia disease. PLoS ONE **20**(7), 0327108 (2025)
31. S.C. Messinis, N.E. Protonotarios, N. Doulamis, Differentially private client selection and resource allocation in federated learning for medical applications using graph neural networks. Sensors **24**(16), 5142 (2024)
32. T.-T. Ho, K.-D. Tran, Y. Huang, Fedsgdcovid: federated sgd covid-19 detection under local differential privacy using chest x-ray images and symptom information. Sensors **22**(10), 3728 (2022)
33. I. Shiri, Y. Salimi, M. Maghsudi, E. Jenabi, S. Harsini, B. Razeghi, S. Mostafaei, G. Hajianfar, A. Sanaat, E. Jafari, Differential privacy preserved federated transfer learning for multi-institutional 68ga-pet image artefact detection and disentanglement. Eur. J. Nucl. Med. Mol. Imag. **51**(1), 40–53 (2023)
34. M.M. Maurer, B. Pfitzner, R.P. Water, L. Faraj, C. Riepe, D. Zuluaga, F. Krenzien, N. Raschzok, R. Siegel, C. Schineis, Privacy preserving federated learning for 90-day mortality prediction in colorectal surgery: a multicenter retrospective development and comparison study. Int. J. Surg. **111**(12), 9065–9074 (2025)
35. S.S. Chowa, M.R.I. Bhuiyan, M.S. Tahosin, A. Karim, S. Montaha, M.M. Hassan, M.A. Shah, S. Azam, An automated privacy-preserving self-supervised classification of covid-19 from lung ct scan images minimizing the requirements of large data annotation. Sci. Rep. **15**(1), 226 (2025)
36. S. Sav, J.-P. Bossuat, J.R. Troncoso-Pastoriza, M. Claassen, J.-P. Hubaux, Privacy-preserving federated neural network learning for disease-associated cell classification. Patterns **3**(5), 100487 (2022)
37. A. Zuo, Z. Feng, Y. Ping, S. Tao, H. Sun, Y. Chen, Fedgraphhe: a privacy-preserving federated graph neural network framework with dynamic homomorphic encryption and robust aggregation. PLoS ONE **21**(1), 0339881 (2026)
38. A. Ali, V. Snášel, J. Platoš, Health-fednet: a privacy-preserving federated learning framework for scalable and secure healthcare analytics. Results Eng. **27**, 106484 (2025)
39. G. Kaissis, A. Ziller, J. Passerat-Palmbach, T. Ryffel, D. Usynin, A. Trask, I. Lima Jr., J. Mancuso, F. Jungmann, M.-M. Steinborn, End-to-end privacy preserving deep learning on multi-institutional medical imaging. Nat. Mach. Intell. **3**(6), 473–484 (2021)
40. C. Nielsen, M. Wilms, N.D. Forkert, Highly efficient homomorphic encryption-based federated learning for diabetic retinopathy classification. J. Med. Imag. **12**(3), 034504–034504 (2025)
41. J. Lin, J. Chen, J. Xiong, D. Jiao, W. Zhao, Y. Xiang, A lightweight privacy-preserving federated learning framework for heterogeneity-resilient skin cancer diagnosis. IEEE J. Biomed. Health Inform. , (2025)
42. Z. Yang, Y. Chen, H. Huangfu, M. Ran, H. Wang, X. Li, Y. Zhang, Dynamic corrected split federated learning with homomorphic encryption for u-shaped medical image networks. IEEE J. Biomed. Health Inform. **27**(12), 5946–5957 (2023)
43. B. Bhasker, P.M. Rao, P. Saraswathi, S.G.K. Patro, J.K. Bhutto, S. Islam, M. Kareemullah, A.F. Emma, Blockchain framework with iot device using federated learning for sustainable healthcare systems. Sci. Rep. **15**(1), 26736 (2025)
44. E. Ashraf, N.F. Areed, H. Salem, E.H. Abdelhay, A. Farouk, Fidchain: Federated intrusion detection system for blockchain-enabled iot healthcare applications. In: Healthcare, (10), pp.1110 (2022). MDPI
45. C. Lin, D. He, X. Huang, X. Xie, K.-K.R. Choo, Ppchain: a privacy-preserving permissioned blockchain architecture for cryptocurrency and other regulated applications. IEEE Syst. J. **15**(3), 4367–4378 (2020)
46. D.K. Murala, L. Vemulapalli, Y. Balagoni, E. Patnala, B. Romeo, Medledgerfl: a hybrid blockchain-federated learning framework for secure remote healthcare services. Sci. Rep. **16**, 8218 (2026)
47. X. Liang, J. Zhao, Y. Chen, E. Bandara, S. Shetty, Architectural design of a blockchain-enabled, federated learning platform for algorithmic fairness in predictive health care: design science study. J. Med. Internet Res. **25**, 46547 (2023)
48. M.R. Hasan, Q. Li, U. Saha, J. Li, Decentralized and secure collaborative framework for personalized diabetes prediction. Biomedicines **12**(8), 1916 (2024)