EDITORIAL

Navigating Ethical Boundaries in Federated Learning for Biomedical Research

Shahnila Rahim[1], **Xiao Kong**[2,*] **and Fatima Abdullah**[3]

[1] Applied Data Science, Noroff University College, Kristiansand, Norway
[2] School of Computer Science and Engineering, Kyungpook National University, Daegu, Republic of Korea
[3] School of Electrical Engineering and Computer Science, National Univsersity of Sciences and Technology, Islamabad, Pakistan

Abstract

Biomedical research is increasingly shaped by vast and diverse datasets, yet their integration is constrained by privacy concerns, regulatory barriers, and fragmented infrastructures. Federated learning (FL) has emerged as a promising paradigm that enables institutions to collaboratively train machine learning models while keeping sensitive data local. This approach has the potential to accelerate discovery in areas such as precision medicine, rare disease research, and population health by pooling knowledge without centralizing data. However, federated learning also introduces new ethical and governance challenges. Risks of information leakage, inequitable participation, algorithmic bias, unclear accountability, and regulatory complexity must be carefully addressed. This editorial highlights these boundaries and emphasizes that technical solutions alone are insufficient. We argue that responsible deployment of FL requires dedicated ethical frameworks, innovative governance structures, continuous auditing, and inclusive global participation. By embedding responsibility into its design and implementation, federated learning can not only advance biomedical science but also foster trust, equity, and sustainability in the future of data-driven health research.

Keywords: federated learning, biomedical ethics, data privacy, precision medicine, AI governance, fairness.

Biomedical research is undergoing a profound transformation. The past decade has witnessed an explosion of data generation from electronic health records, genomic sequencing, medical imaging, wearable devices, and digital health platforms [6, 11]. These data sources, in aggregate, hold immense promise for advancing precision medicine, accelerating drug discovery, and enabling predictive models for disease prevention [18, 19]. Yet the very abundance of biomedical data has become a double-edged sword. Information remains scattered across institutions, locked behind regulatory firewalls, and protected by ethical obligations that rightfully prioritize patient privacy.

Traditional approaches to collaborative research, such as centralized data repositories or pooled multicenter

Submitted: 22 August 2025
Accepted: 29 September 2025
Published: 09 December 2025

Vol. 1, **No.** 2, 2025.
10.62762/JAIB.2025.703433

***Corresponding author:**
✉ Xiao Kong
e-mail@e-mail.com

Citation
Rahim, S., Kong, X., & Abdullah, F. (2025). Navigating Ethical Boundaries in Federated Learning for Biomedical Research. *Journal of Artificial Intelligence in Bioinformatics*, 1(2), 72–78.

datasets, are becoming increasingly difficult to sustain. Regulations like the Health Insurance Portability and Accountability Act (HIPAA) in the United States and the General Data Protection Regulation (GDPR) in Europe impose strict safeguards on patient information [20]. Public trust in medical research has also grown more fragile in the wake of high-profile data breaches and misuse of health records [21, 22]. The central challenge is clear: how can we harness the collective power of biomedical data without compromising privacy, equity, or trust?

Federated learning has emerged as one of the most promising answers to this dilemma. In this paradigm, multiple institutions collaborate to train a shared machine learning model while keeping sensitive patient data within its original custodial environment. Instead of transferring raw datasets, only model updates, representing what the system has learned, are communicated to a coordinating server. This architectural change represents more than a technical modification. It offers a fundamental reimagining of data sharing in which knowledge can move freely while data remains local, thereby providing a new path for global-scale biomedical research [1, 3].

The enthusiasm surrounding federated learning is evident across the biomedical community. Hospitals that once hesitated to collaborate because of privacy concerns are now exploring federated frameworks. Pharmaceutical companies are piloting federated methods to accelerate drug discovery across geographically dispersed research centers. Academic consortia are experimenting with federated models for cancer detection, rare disease classification, and population health analytics [2, 23, 24]. At first glance, federated learning appears to remove barriers that have long constrained biomedical innovation.

Yet technological solutions are never purely technical. They embody values, assumptions, and power structures. As the adoption of federated learning expands, difficult questions arise. Does keeping data local truly guarantee privacy? Who gains the opportunity to participate in federated networks, and under what terms? Will models trained across unevenly distributed resources reinforce existing inequities in healthcare? How will accountability be assigned if a federated model contributes to a clinical error that harms a patient?

This editorial begins with optimism regarding what federated learning makes possible but emphasizes that optimism alone is not sufficient. The greater challenge lies in navigating the ethical boundaries that shape how federated learning is deployed, governed, and trusted in biomedical research. By raising these issues at the outset, the intention is not to temper enthusiasm but to ensure that enthusiasm is guided by responsibility.

1 The Promise of Federated Learning

The central promise of federated learning lies in its ability to reconcile two seemingly opposing demands in biomedical research: the need for large, diverse datasets and the obligation to protect patient privacy. Diseases such as cancer, Alzheimer's disease, and rare genetic disorders manifest differently across populations, environments, and healthcare systems. Robust and generalizable insights can only emerge when data from multiple institutions and geographies are brought together. Federated learning offers a path to realize this vision without requiring the centralization of sensitive health information.

1.1 Protecting Patient Privacy

In federated learning, patient-level data never leave the local institution. Instead, learning occurs on-site and only model updates are shared. This approach aligns with modern data protection regulations such as HIPAA and GDPR, which emphasize data minimization and accountability. Early studies have demonstrated that federated frameworks can produce models with performance comparable to those trained on centralized datasets while mitigating privacy risks [1, 2].

1.2 Enabling Global Collaboration

Federated learning lowers barriers to participation in cross-institutional research. Hospitals and research centers in different countries can contribute to global models without exposing raw data to external parties. This inclusivity is particularly important for building models that are representative of diverse populations, an essential step toward equitable precision medicine. By supporting cross-border collaborations, federated approaches also foster a culture of trust and cooperation in biomedical research.

1.3 Advancing Rare Disease Research

Rare conditions, by definition, affect too few patients at individual sites to permit robust machine learning analyses. Federated learning overcomes this limitation by pooling knowledge across distributed cohorts, thereby strengthening statistical power

without centralizing data. Recent experiments in multi-institutional tumor segmentation and neuroimaging consortia illustrate the feasibility of this approach [1, 2].

1.4 Accelerating Biomedical Discovery

By facilitating access to large-scale, heterogeneous data, federated learning can accelerate discovery in areas such as drug response, disease subtyping, and clinical outcome prediction. Pharmaceutical companies are exploring federated frameworks to support collaborative drug development, while academic groups have demonstrated success in federated analysis of medical imaging and genomics. The resulting models are more generalizable, more robust, and potentially more impactful than those trained in isolated institutional silos.

1.5 A New Model of Scientific Collaboration

Beyond technical advantages, federated learning represents a cultural shift in the way biomedical research is conducted. It encourages openness, inclusivity, and shared responsibility, resonating with the broader movement toward open science. By balancing innovation with privacy and accountability, federated learning sets the stage for a new era of cooperative biomedical discovery.

2 Navigating Ethical Boundaries

While the advantages of federated learning are compelling, the technology is not ethically neutral. It raises profound questions about how biomedical data are used, who benefits from the resulting models, and how risks are distributed. To ensure that federated approaches strengthen rather than undermine trust in medical research, it is essential to critically examine the ethical boundaries that shape their deployment. Several interrelated domains require particular attention.

2.1 Privacy and Consent

The promise of federated learning rests on the assumption that keeping data local inherently protects privacy. Yet, recent studies have demonstrated that model updates can, under certain conditions, be reverse-engineered to reveal sensitive patient information [4, 5]. Techniques such as gradient inversion attacks show that even without direct access to raw data, adversaries may reconstruct identifiable features from shared parameters.

From the perspective of informed consent, this introduces a subtle but important dilemma. Patients may agree to participate in "privacy-preserving" research without fully appreciating that absolute anonymity cannot be guaranteed. Consent frameworks therefore require revision. Instead of broad, static agreements, institutions should consider dynamic consent models in which participants are informed about evolving risks and given greater agency in managing how their data contribute to federated projects. Emerging approaches such as differential privacy and secure multiparty computation may reduce leakage risks, but they must be transparently communicated to participants.

2.2 Equity and Inclusion

Federated learning networks are often resource-intensive, requiring robust computational infrastructure, high-speed connectivity, and skilled technical staff. As a result, well-funded hospitals and academic centers may dominate these collaborations, while institutions in low- and middle-income countries risk marginalization. The outcome is a federated model that reflects the realities of high-resource environments while neglecting underrepresented populations.

Such exclusion has significant ethical implications. If models trained in wealthy settings are later deployed in disadvantaged communities, the resulting inequities may widen. To address this, consortia should implement mechanisms of capacity building, such as shared infrastructure funding, standardized toolkits, and cloud-based federated platforms that lower technical barriers to entry. Ensuring that all stakeholders, regardless of geography, can contribute meaningfully is not simply a matter of fairness but a prerequisite for building generalizable models that serve global populations.

2.3 Bias and Fairness

Federated learning does not eliminate bias; it can, in fact, amplify it. Local datasets are shaped by the demographics, diagnostic practices, and clinical environments of their institutions. When combined without explicit bias mitigation, these local disparities may produce global models that perform unevenly across subgroups. For example, a federated cancer detection model trained predominantly on imaging from European populations may systematically underperform in African or Asian populations [6].

Addressing fairness requires both technical and

governance innovations. Technically, strategies such as re-weighting updates, stratified sampling of gradients, or fairness-aware optimization can help balance underrepresented cohorts. At the governance level, federated consortia should adopt mandatory bias auditing and reporting, similar to how clinical trials are required to disclose demographic representation. By embedding fairness into both the algorithmic and organizational layers, federated learning can avoid reinforcing structural inequities in healthcare.

2.4 Transparency and Accountability

Biomedical research requires not only accurate but also interpretable models. Clinicians must be able to understand why a model has made a particular recommendation if they are to integrate it into decision-making. Yet many federated learning applications rely on opaque deep learning architectures that resist straightforward interpretation. Without transparency, trust in federated outcomes may be undermined.

Accountability is equally pressing. In a multi-institutional federated consortium, responsibility for errors is diffuse. If a patient is harmed by a flawed model, is accountability borne by the local institution that trained part of the model, the coordinating server, or the consortium as a whole? Current liability frameworks are ill-suited to distributed machine learning. To address this, federated networks should establish clear accountability agreements before deployment, outlining who holds responsibility for model performance, monitoring, and remediation. Regulatory agencies may also need to develop new guidelines tailored specifically for distributed AI systems in healthcare.

2.5 Regulatory Complexity

Federated learning often spans multiple jurisdictions, each with distinct legal frameworks for health data. HIPAA in the United States, GDPR in the European Union, and diverse national laws elsewhere create a regulatory patchwork that complicates international collaboration. Questions of model ownership, intellectual property rights, and data sovereignty remain unsettled. For instance, if a global federated model trained across European and Asian hospitals leads to a profitable therapeutic application, how should benefits be distributed among participants?

Novel governance mechanisms are urgently needed. One possibility is the establishment of federated data trusts independent entities that oversee access, auditing, and benefit-sharing for distributed models. Another is the use of smart contracts on blockchain systems to automatically enforce contribution tracking and equitable redistribution of value. While such proposals are in early stages, they illustrate how ethical reflection can inspire innovative solutions to regulatory challenges.

3 Toward Responsible Deployment

If federated learning is to fulfill its promise in biomedical research, its deployment must be guided by deliberate strategies that ensure ethical integrity, technical robustness, and social legitimacy. The central task is to transform federated learning from a promising proof-of-concept into a trustworthy socio-technical infrastructure. Several interrelated directions illustrate how this transformation can be achieved as shown in Figure 1.

3.1 Building Ethical Frameworks for Federated Research

A first priority is the development of ethical frameworks tailored specifically for federated learning in biomedicine. Existing guidelines for data sharing and biomedical ethics do not fully anticipate the risks introduced by distributed machine learning. For instance, the possibility of gradient inversion attacks requires explicit consent procedures that go beyond conventional data use agreements. Frameworks such as the World Health Organization's guidance on ethics and governance of AI for health [7] and the European Commission's Ethics Guidelines for Trustworthy AI [8] provide useful starting points but must be adapted to federated contexts. Concrete measures include dynamic consent models, periodic ethics audits, and patient advisory boards embedded within federated consortia.

3.2 Innovative Governance Structures

Traditional governance approaches are inadequate for distributed networks where responsibility is shared across institutions. Novel structures such as *data trusts* and *federated governance boards* have been proposed to oversee model training, auditing, and benefit-sharing [9, 10]. Smart contracts implemented via blockchain systems may provide automated mechanisms for recording contributions, enforcing data use restrictions, and distributing benefits fairly [11]. Such governance innovations can reduce ambiguity in accountability while also incentivizing

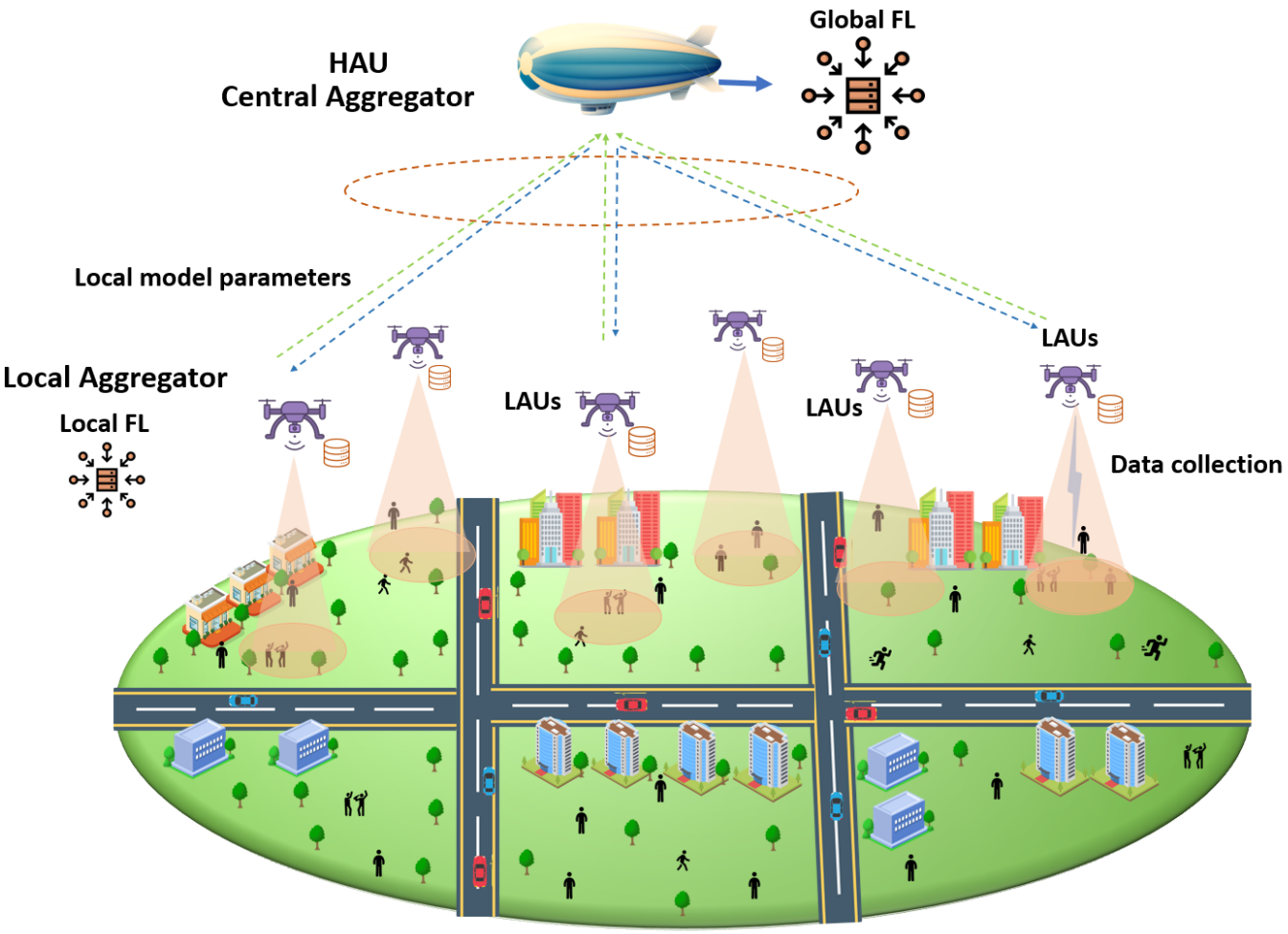

Figure 1. Conceptual diagram of federated learning in biomedical research with surrounding ethical boundaries.

institutions of varying sizes to participate in federated collaborations.

3.3 Technical Safeguards: Privacy and Security

Responsible deployment requires embedding privacy-preserving techniques into the core architecture of federated learning. Differential privacy, secure multiparty computation, and homomorphic encryption are being integrated into federated frameworks to mitigate risks of data leakage [12, 13]. Emerging work on hybrid federated systems combines multiple techniques to provide stronger guarantees. For example, the Federated Tumor Segmentation (FeTS) challenge has demonstrated the feasibility of secure aggregation protocols in multi-institutional medical imaging [24]. Continued innovation in these safeguards will be essential to ensuring public trust.

3.4 Auditing, Monitoring, and Explainability

Trustworthy federated models must be continuously monitored for fairness, bias, and interpretability. Auditing tools are beginning to emerge that allow for decentralized performance evaluation without exposing raw data [14]. Techniques for federated explainability, such as distributed SHAP or gradient-based saliency methods, offer clinicians insights into model decisions even when the underlying model is trained across multiple sites [15]. Embedding auditing and explainability into the federated pipeline can prevent harm, enhance transparency, and satisfy regulatory demands for model interpretability.

3.5 Capacity Building and Equitable Participation

Without deliberate intervention, federated learning risks reinforcing global inequities in research. Institutions in low- and middle-income countries may lack the computational infrastructure needed to join federated networks. Capacity building must therefore be a central pillar of responsible deployment. Cloud-based federated platforms, standardized toolkits such as TensorFlow Federated and Flower, and shared funding mechanisms can lower barriers to entry [3]. Equally important is ensuring that the

resulting global models are validated on data from underrepresented populations, thereby enhancing their generalizability and fairness.

3.6 Interdisciplinary Collaboration and Policy Alignment

Finally, responsible federated learning cannot be achieved by technical innovation alone. Interdisciplinary collaboration is essential. Ethicists, legal scholars, clinicians, and patient advocates must be involved in federated projects from inception through deployment. Moreover, alignment with policy frameworks is critical. Initiatives such as the Global Alliance for Genomics and Health (GA4GH) and the OECD's recommendations on AI governance provide platforms for harmonizing federated practices across borders [16, 17]. Integrating federated learning into these broader policy ecosystems will help avoid fragmentation and ensure sustainability.

3.7 From Promise to Practice

Early initiatives demonstrate that responsible deployment is feasible. The Federated Tumor Segmentation (FeTS) challenge, federated analyses of diabetic retinopathy screening, and pharmaceutical collaborations in drug discovery illustrate that privacy-preserving multi-institutional learning is not only possible but can achieve performance comparable to centralized models [23, 24]. What remains is to scale these demonstrations into durable infrastructures, supported by ethical frameworks, governance structures, and continuous monitoring. By embedding responsibility into every layer—technical, organizational, and regulatory—federated learning can evolve from a promising innovation to a cornerstone of biomedical research.

4 Conclusion

Federated learning has the potential to reshape biomedical research by enabling collaboration without compromising patient privacy. It allows diverse institutions to contribute to shared models while keeping data secure at the source. This approach can unite fragmented knowledge, strengthen rare disease research, and build models that reflect the diversity of global populations. The challenge ahead is to ensure that this promise is realized responsibly. Privacy safeguards, fair participation, transparent governance, and continuous monitoring must become integral to every federated initiative. These measures are not optional; they are essential to building trust and ensuring equitable outcomes. The future of biomedical discovery will depend not only on technological innovation but also on the values that guide its use. Federated learning offers an opportunity to advance science while protecting human dignity. How we choose to navigate this opportunity will define its lasting impact.

Data Availability Statement

Not applicable.

Funding

This work was supported without any funding.

Conflicts of Interest

The authors declare no conflicts of interest.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Sheller, M. J., Edwards, B., Reina, G. A., Martin, J., Pati, S., Kotrotsou, A., ... & Bakas, S. (2020). Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. *Scientific reports, 10*(1), 12598. [CrossRef]

[2] Li, X., Gu, Y., Dvornek, N., Staib, L. H., Ventola, P., & Duncan, J. S. (2020). Multi-site fMRI analysis using privacy-preserving federated learning and domain adaptation: ABIDE results. *Medical image analysis, 65*, 101765. [CrossRef]

[3] 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. [CrossRef]

[4] Hitaj, B., Ateniese, G., & Perez-Cruz, F. (2017, October). Deep models under the GAN: Information leakage from collaborative deep learning. In *Proceedings of the 2017 ACM SIGSAC conference on computer and communications security* (pp. 603-618). [CrossRef]

[5] Melis, L., Song, C., De Cristofaro, E., & Shmatikov, V. (2019, May). Exploiting unintended feature leakage in collaborative learning. In *2019 IEEE symposium on security and privacy (SP)* (pp. 691-706). IEEE. [CrossRef]

[6] Kaissis, G. A., Makowski, M. R., Rückert, D., & Braren, R. F. (2020). Secure, privacy-preserving and federated machine learning in medical imaging. *Nature Machine Intelligence, 2*(6), 305-311. [CrossRef]

[7] World Health Organization. (2024). *Ethics and governance of artificial intelligence for health: large multi-modal models*. WHO guidance. World Health Organization.

[8] Cannarsa, M. (2021). Ethics guidelines for trustworthy AI. *The Cambridge handbook of lawyering in the digital age*, 283-297. [CrossRef]

[9] Delacroix, S., & Lawrence, N. D. (2019). Bottom-up data trusts: Disturbing the 'one size fits all'approach to data governance. *International data privacy law, 9*(4), 236-252. [CrossRef]

[10] Morley, J., Murphy, L., Mishra, A., Joshi, I., & Karpathakis, K. (2022). Governing data and artificial intelligence for health care: developing an international understanding. *JMIR formative research, 6*(1), e31623. [CrossRef]

[11] Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H. R., Albarqouni, S., ... & Cardoso, M. J. (2020). The future of digital health with federated learning. *NPJ digital medicine, 3*(1), 119. [CrossRef]

[12] Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., ... & Roselander, J. (2019). Towards federated learning at scale: System design. *Proceedings of machine learning and systems, 1*, 374-388.

[13] Zhang, C., Xie, Y., Bai, H., Yu, B., Li, J., & Gao, Y. (2021). A survey on federated learning. *Knowledge-Based Systems, 216*, 106775. [CrossRef]

[14] Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., ... & Zhao, S. (2021). Advances and open problems in federated learning. *Foundations and trends® in machine learning, 14*(1–2), 1-210. [CrossRef]

[15] López-Blanco, R., Alonso, R. S., González-Arrieta, A., Chamoso, P., & Prieto, J. (2023, July). Federated learning of explainable artificial intelligence (FED-XAI): A review. In *International Symposium on Distributed Computing and Artificial Intelligence* (pp. 318-326). Cham: Springer Nature Switzerland. [CrossRef]

[16] Terry, S. F. (2014). The global alliance for genomics & health. *Genetic testing and molecular biomarkers, 18*(6), 375-376. [CrossRef]

[17] Yeung, K. (2020). Recommendation of the council on artificial intelligence (OECD). *International legal materials, 59*(1), 27-34. [CrossRef]

[18] Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature medicine, 25*(1), 44-56. [CrossRef]

[19] Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., ... & Dean, J. (2019). A guide to deep learning in healthcare. *Nature medicine, 25*(1), 24-29. [CrossRef]

[20] Shabani, M., Dyke, S. O., Joly, Y., & Borry, P. (2015). Controlled access under review: improving the governance of genomic data access. *PLoS Biology, 13*(12), e1002339. [CrossRef]

[21] Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society, 3*(2), 2053951716679679. [CrossRef]

[22] Pfitzner, B., Steckhan, N., & Arnrich, B. (2021). Federated learning in a medical context: a systematic literature review. *ACM Transactions on Internet Technology (TOIT), 21*(2), 1-31. [CrossRef]

[23] Dayan, I., Roth, H. R., Zhong, A., Harouni, A., Gentili, A., Abidin, A. Z., ... & Li, Q. (2021). Federated learning for predicting clinical outcomes in patients with COVID-19. *Nature medicine, 27*(10), 1735-1743. [CrossRef]

[24] Pati, S., Baid, U., Edwards, B., Sheller, M., Wang, S. H., Reina, G. A., ... & Poisson, L. (2022). Federated learning enables big data for rare cancer boundary detection. *Nature communications, 13*(1), 7346. [CrossRef]

Shahnila Rahim received the Ph.D. degree in computer science and engineering from the School of Computer Science and Engineering, Kyungpook National University (KNU), Daegu, Republic of Korea. She is currently an Assistant Professor with the Department of Applied Data Science, Noroff University College, Kristiansand, Norway. Her research interests include machine learning, deep learning, federated learning, aerial computing, reinforcement learning, B5G communication networks, and the Internet of Things (IoT). (Email: Shahnila.Rahim@noroff.no)

Xiao Kong received the B.S degree in Computer Science and Engineering from Qilu University of Technology, China, in 2021. Received master's degree in computer science and engineering from Kyungpook National University, South Korea, in 2025. (Email: kongxiao@knu.ac.kr)

Fatima Abdullah received the Ph.D. degree in Computer Science and Engineering from Kyungpook National University, Daegu, South Korea, in 2024. Her research interests include the Internet of Things (IoT), fog computing, edge computing, streaming data analytics, and latency optimization within fog networks. (Email: fatima.abdullah@seecs.edu.pk)