The Cost of Progress: Exploring Privacy Nightmares for AI in Precision Medicine

Ahmed Aziz[1,*], Noura Metawa[2]

[1] Tashkent state university of Economics, Tashkent, Uzbekistan

[2]American University in the Emirates, UAE

Emails: a.mohamed@tsue.uz; n.metawa@aue.ae

Abstract

Precision medicine is an innovative approach to healthcare that relies on the use of genomic data, electronic health records, and other types of medical data to develop personalized prevention, diagnosis, and treatment strategies for patients. The use of artificial intelligence (AI) in precision medicine has the potential to improve patient outcomes and reduce healthcare costs, but it also raises significant privacy concerns. This paper provides a comprehensive review of the privacy nightmares associated with the use of AI in precision medicine. We examine the potential risks and threats to patient privacy, including the use of personal data for unintended purposes, the risk of data breaches and hacking, and the potential for discrimination and bias. We also analyze the legal and ethical implications of using AI in precision medicine, including issues related to informed consent and data ownership. Our investigation highlights the need for strong data protection regulations and ethical frameworks to safeguard patient privacy in the age of AI in precision medicine. As the use of AI in precision medicine continues to expand, the paper presents a road for future directions for protecting patient privacy, including the use of privacy-preserving machine learning algorithms and the adoption of privacy-enhancing technologies.

Keywords: Artificial Intelligence; Privacy; Healthcare; Precision Medicine

1. Introduction

Precision medicine, also known as personalized medicine, is a rapidly advancing field in healthcare that aims to provide tailored prevention, diagnosis, and treatment strategies for individual patients based on their unique genetic makeup, environment, and lifestyle. This approach is a departure from traditional medicine, which typically relies on a one-size-fits-all approach that may not work as well for all patients. The rise of precision medicine has been driven by recent technological advancements, including the ability to analyze large amounts of genomic data and the development of targeted therapies. With precision medicine, doctors can identify specific genetic mutations or biomarkers that may be driving a patient's disease and develop targeted treatments to address those specific factors. Precision medicine has already shown promising results in the treatment of a variety of diseases, including cancer, cardiovascular disease, and rare genetic disorders. However, the field is still in its early stages, and there are many challenges to overcome, including the cost and complexity of genomic testing and the need for more robust clinical trials [1-3].

Artificial intelligence (AI) is playing an increasingly important role in precision medicine. AI can be used to analyze large amounts of data, including genomic data, electronic health records, and medical imaging, to identify patterns and insights that may be difficult or impossible for humans to detect. One of the main applications of AI in precision medicine is in the development of personalized treatment plans. AI algorithms can analyze a patient's genomic data,

DOI: https://doi.org/10.54216/JCIM.080205
Received May 12, 2021 Accepted: November 11, 2021

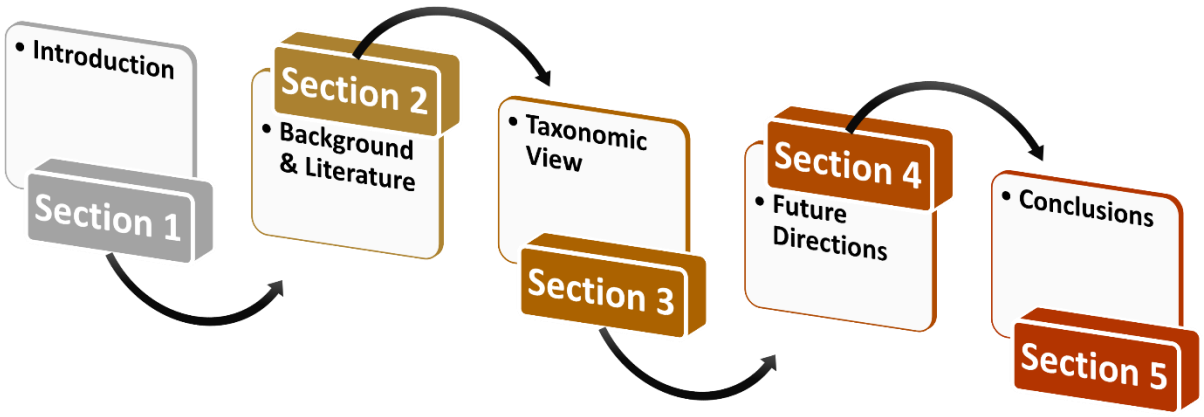

Figure 1: Organization of this study.

along with other clinical and demographic information, to predict how that patient is likely to respond to different treatments. This information can help doctors develop more targeted and effective treatment plans, improving patient outcomes and reducing healthcare costs [4]. Another area where AI is being used in precision medicine is in drug discovery. By analyzing vast amounts of data, including genomic and proteomic data, AI algorithms can identify potential drug targets and help researchers develop new drugs more quickly and efficiently [5]. AI is also being used to improve diagnostic accuracy. For example, machine learning algorithms can analyze medical images, such as CT scans and MRIs, to identify patterns and abnormalities that may be indicative of disease. This can help doctors make more accurate diagnoses and develop more effective treatment plans [7].

The paper contributes to the growing body of research on the use of AI in precision medicine, providing a comprehensive review of the privacy nightmares associated with this technology. We also analyze the legal and ethical implications of using AI in precision medicine. Informed consent, data ownership, and privacy regulations are crucial considerations for the use of AI in precision medicine, and the paper provides insights and recommendations for addressing these issues. Finally, we present a comprehensive road map for future research on the private aspects of precision medicine. This road map can provide an insightful guide to the community, that aims to pave the way for further research in this field. The remainder of this research is organized into five sections presented in Figure 1.

2. Background & Literature

Precision medicine is a rapidly developing field that uses AI and machine learning to analyze large amounts of patient data to develop personalized treatment plans. While this approach promises to revolutionize healthcare, it also poses significant privacy challenges. The use of personal health information (PHI) raises concerns about the security and confidentiality of sensitive data. Patients have a right to expect that their PHI will be protected and used only for their benefit, and not shared or sold without their consent. In addition, the use of AI raises questions about the transparency and accountability of decision-making algorithms, which can have life-and-death consequences for patients. If AI-based systems are not developed and used responsibly, they could exacerbate existing health disparities and create new ones [8].

One of the primary challenges facing AI in precision medicine is the need to balance privacy and data access. On the one hand, researchers and healthcare providers need access to large amounts of data to develop accurate algorithms and treatment plans. On the other hand, patients have a right to control their PHI and to know who is accessing it and for what purpose. Another challenge is ensuring that AI-based systems are transparent and accountable. Patients have a right to understand how decisions about their health are being made and to have access to information that can help them make informed decisions about their treatment options. Finally, there is a need to address the potential for bias in AI algorithms, which could exacerbate health disparities by perpetuating existing biases and creating new ones. To overcome these challenges, stakeholders in the precision medicine field must work together to develop transparent and accountable AI-based systems that prioritize patient privacy and prioritize equity in healthcare delivery [9].

The research literature on the privacy challenges confronting AI in precision medicine is extensive and growing rapidly. Scholars have identified a range of privacy concerns related to the use of AI in healthcare, including the need to protect PHI, ensure algorithmic transparency, and address potential bias. Studies have also examined the legal and

DOI: https://doi.org/10.54216/JCIM.080205
Received May 12, 2021 Accepted: November 11, 2021

ethical implications of using AI in precision medicine, and have proposed various frameworks and guidelines to guide responsible AI development and deployment [10]. For example, some researchers have suggested that data governance frameworks be developed to ensure that data is collected and used in a way that is respectful of patient privacy and that data-sharing agreements be put in place to ensure that patient data is shared only with authorized parties. In addition, research has focused on the technical challenges of ensuring that AI-based systems are transparent and accountable. For example, scholars have explored ways to develop explainable AI models that allow patients and clinicians to understand how decisions are being made. Other research has focused on developing methods for detecting and mitigating bias in AI algorithms, such as using diverse training data and developing bias mitigation techniques. A summary of Literature studies on privacy preservation is given in Table 1.

Table 1: Summary of the literature studies on privacy preservation.

References	Technologies	Contributions	Downsides	Year
[15]	IoT, IIoT, Security, CPS	A detailed analysis of IIoT security concerns and attacks on layered architectures.	Propose only a study of recent states-of-art	2018
[16]	BSCS ABE	Solve supply chain resource interaction security issues	Lack of comparative results	2019
[17]	Blockchain, Smart Contract,	A comprehensive survey on Blockchain scalability and data capacity of IIoT	Propose only a study of existing solutions	2019
[18]	Blockchain, DRL	Optimize IIoT scalability/throughput.	Lack of comparative results	2019
[19]	Data mining	Defines typical supply chain risk and how the company controls network risk.	No case study was proposed	2019
[20]	Blockchain, BLMS	Study of a blockchain-based logistics monitoring system (BLMS).	Lack of comparative results	2019
[21]	Big data	Discuss supply chain management success aspects that earlier studies missed.	No case study was proposed	2019
[14]	Blockchain, Smart Contract, Classification	Summarize security issues of blockchain-based IoT and IIoT	No case study was proposed	2020
[5]	e-Healthcare, IoT	data anonymization/access control mechanisms.	No case study was proposed	2020
[13]	e-Healthcare, IoT	FL	Narrow scope	2020
[8]	wireless sensor networks	classification trees for the multifaceted challenge of privacy protection in healthcare	No case study was proposed	2020

3. Privacy Preservation Solutions

In this section, we will discuss the different types of privacy preservation methods used in precision medicine, including data Anonymization, pseudonymization, differential privacy, homomorphic encryption, and federated

DOI: https://doi.org/10.54216/JCIM.080205
Received May 12, 2021 Accepted: November 11, 2021

learning (FL) (See Figure 2). We will explore the benefits and drawbacks of each approach and discuss how they can be used to protect patient privacy while still enabling the development of effective AI-based healthcare systems.

3.1. Anonymization and pseudonymization techniques

Anonymization and pseudonymization are two common techniques used to preserve privacy in precision medicine. Anonymization involves removing all identifiable private data from a record, rendering it untraceable to an individual patient. This technique is widely used in medical datasets, and anonymization software is often built into clinical data archiving systems, making it a straightforward method in practice. However, errors in the anonymization process can render the protection ineffective, and the definition of "sufficient" de-identification varies across different jurisdictions, complicating the establishment of international standards. Furthermore, de-identification techniques are typically employed as a preparation for data transfer or sharing, which presents issues if a patient withdraws their consent or if legislation changes.

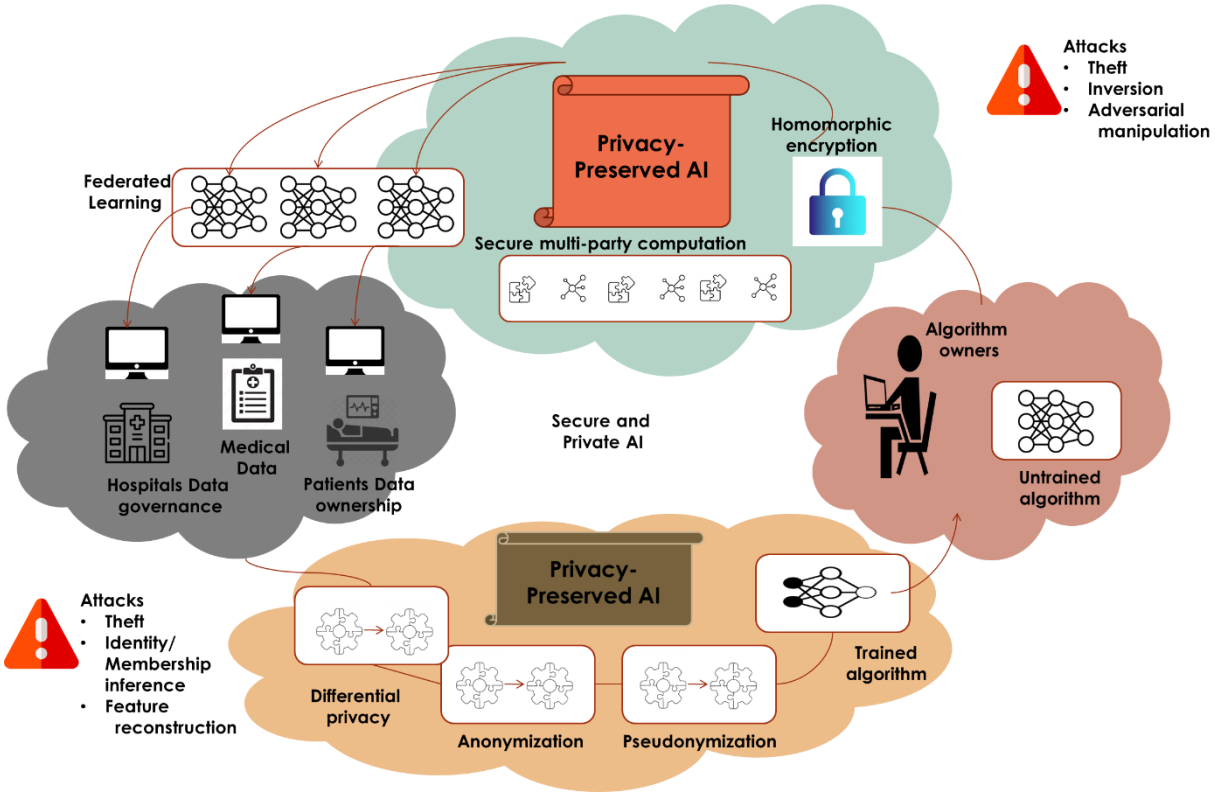

Figure 2: visualization of process of privacy-preservation in precision medicine.

Pseudonymization, on the other hand, involves replacing sensitive entries with artificially generated ones while still allowing re-attribution using a look-up table. This technique requires data manipulation rather than just data deletion, and the look-up table must be safely kept separately. While pseudonymization provides an additional layer of protection by making it more difficult to trace a record back to an individual, it poses additional difficulties compared to anonymization. For example, the look-up table can be problematic in the setting of insecure storage, risking data theft. Technical errors can also render the protection ineffective, and retaining institution names can make an entire dataset identifiable [11-14]. The distinction between different aspects of the above methods is given in Table 2.

Table 2: Comparison between Anonymization and pseudonymization

Criteria	Anonymization	Pseudonymization
Definition	Removal of private data from a record	Replacement of sensitive entries with artificially generated ones while still allowing re-attribution using a look-up table

DOI: https://doi.org/10.54216/JCIM.080205
Received May 12, 2021 Accepted: November 11, 2021

Methodology	Data deletion	Data manipulation
Look-up table	Not required	Required
Safekeeping	Not required	Required
Data theft risk	Low	High
Protection	Straightforward	Additional layer
Ineffectiveness	Errors in the process	Technical errors, retaining institution names
Data Ownership	Uncouples data governance from data ownership	Uncouples data governance from data ownership
Legislation	Issues of legislation changes	Issues of legislation changes
Imaging dataset	Varies in difficulty to link back to an individual	Varies in difficulty to link back to an individual
Re-identification risk	Low, but not zero	Low with appropriate security measures in place
Utility loss	High if data is removed, some if masked	Low if generated data is realistic
Data analysis	Limited by removal of sensitive data	More robust, but requires a look-up table
Regulatory compliance	May meet HIPAA standards	May meet HIPAA standards
Ethical considerations	May not respect patient autonomy	May not respect patient autonomy
Transparency	No guarantee of transparency	More transparent with a look-up table
Data Sharing	Possible with appropriate security measures in place	Possible with appropriate security measures in place

3.2. federated machine learning

FL is an approach to AI in precision medicine that has the potential to preserve privacy by allowing different institutions or individuals to collaborate on a shared model without sharing their private data. This is achieved by training a model using data that remains in its original location, while only the model parameters are shared with a central server for aggregation. This approach ensures that sensitive patient data is not shared across different institutions, reducing the risk of data breaches and privacy violations. Additionally, federated machine learning can improve the accuracy of models by allowing larger and more diverse datasets to be used for training. One key advantage of federated machine learning is that it allows institutions to maintain control over their own data while still benefiting from a collaborative model [15]. This is particularly useful in precision medicine, where institutions may have limited access to large, diverse datasets due to data privacy concerns. By combining their data in a secure and privacy-preserving way, institutions can improve the accuracy and effectiveness of their models, leading to better patient outcomes. Additionally, federated machine learning can help address issues of data bias by allowing models to be trained on more diverse datasets from different institutions and populations. Overall, federated machine learning has the potential to revolutionize AI in precision medicine by enabling collaboration while maintaining data privacy and security [16]. A summary of FL algorithms is given in Table 3.

Table 3: An overview of FL algorithms for preserving the privacy learning process.

FL Algorithm	**Advantages**	**Disadvantages**
Federated Averaging (FedAvg)	Simple to implement, scales well to large numbers of participants, and can handle non-iid data.	Can suffer from slow convergence and may require a large number of communication rounds to achieve convergence.
Federated Stochastic Gradient Descent (FedSGD)	Can achieve faster convergence than FedAvg and can handle non-iid data.	Can suffer from high variance due to small batch sizes and may require careful tuning of the learning rate and regularization.
FL with Secure Aggregation (FedSecAgg)	Provides strong privacy guarantees using homomorphic encryption and differential privacy.	Can be computationally expensive and may require specialized hardware for efficient implementation.

DOI: https://doi.org/10.54216/JCIM.080205
Received May 12, 2021 Accepted: November 11, 2021

Federated Group Knowledge Transfer (FedGKT)	Allows for transfer of knowledge across similar groups of participants while preserving privacy.	Can be computationally expensive and may require careful tuning of hyperparameters.
Robust Federated Aggregation (RFA)	Provides robustness against participant dropouts and malicious attacks by using outlier-resistant statistics.	Can be computationally expensive and may require careful tuning of hyperparameters.
Federated Meta-Learning (FedMA)	Enables faster convergence and improved generalization by learning shared model architectures across participants.	May require specialized hardware and careful tuning of hyperparameters.
Federated Distance (FedDist)	Allows for privacy-preserving similarity comparisons between data sets without sharing data directly.	Requires careful tuning of hyperparameters and may suffer from reduced accuracy compared to methods that use raw data.
Federated Distillation (FedDistil)	Allows for training a smaller and more efficient model at the edge devices by distilling knowledge from a larger, centralized model.	May suffer from reduced accuracy compared to the centralized model and may require careful tuning of hyperparameters.

3.3. Differential privacy

Differential privacy has emerged as a promising approach for protecting the privacy of patient data in precision medicine. By adding random noise to a dataset, differential privacy can ensure that individual patient records cannot be easily identified, while still enabling statistical analysis of the dataset as a whole. This is particularly important in precision medicine, where patient data can be extremely sensitive and personal, and where the use of machine learning and other advanced analytics techniques can potentially reveal even more personal information.

In precision medicine, differential privacy can be used to enable data sharing and collaborative research among healthcare providers and researchers without compromising patient privacy. By implementing differential privacy techniques, healthcare providers and researchers can ensure that they are only accessing data that has been sufficiently anonymized to prevent the identification of individual patients. This can help to build trust among patients and ensure that they are willing to participate in research studies and share their data with healthcare providers [17].

One of the main advantages of differential privacy in precision medicine is that it enables the analysis of large, complex datasets that would otherwise be too difficult to analyze due to privacy concerns. For example, differential privacy can be used to enable the analysis of genetic data from large patient cohorts, providing insights into the underlying genetic basis of complex diseases. By ensuring that patient privacy is protected, differential privacy can also enable the development of more personalized treatments and therapies, leading to better outcomes for patients. However, it is important to note that differential privacy is not a silver bullet and that additional measures may be needed to ensure that patient data is used in a responsible and ethical manner.

Formally speaking, differential privacy can be defined as follows. Let D_1 and D_2 be two neighboring datasets, where D_2 is obtained by adding or removing a single record from D_1. A randomized algorithm A satisfies epsilon-differential privacy if, for any possible output S of A, and any pair of neighboring datasets D_1 and D_2:

$$Pr[A(D_1) = S] \leq exp(\epsilon) * Pr[A(D_2) = S] \quad (1)$$

where $Pr[A(D_1) = S]$ represents the probability that algorithm A produces output S when given dataset D_1 as input. In simple terms, this definition states that a randomized algorithm satisfies differential privacy if the probability of producing a particular output when given a dataset D_1 is only slightly different from the probability of producing the same output when given a dataset D_2, which differs from D_1 by only one record. The degree of difference is controlled by the parameter epsilon, with smaller values of epsilon providing stronger privacy guarantees.

Differential privacy can be classified into two main types: local differential privacy (LDP) and global differential privacy (GDP). LDP involves the addition of noise to the data at the individual level. In LDP, each user adds random noise to their data before sharing it with a central server or other parties. The noise added to the data helps to mask the specific data of each individual, thereby preserving privacy. LDP is a promising technique for ensuring privacy in scenarios where individual data needs to be protected, such as in medical studies where data contains sensitive personal information. GDP, on the other hand, involves the addition of noise to the aggregate data. GDP methods add random

DOI: https://doi.org/10.54216/JCIM.080205
Received May 12, 2021 Accepted: November 11, 2021

noise to the aggregate data to ensure that the information about any individual is not revealed in the analysis. This method is typically used for analysis on large datasets where individual-level data may not be required for analysis. GDP has been used in scenarios such as recommendation systems and market research, where the main goal is to analyze trends in the data rather than individual behavior [18].

The choice between LDP and GDP depends on the type of data and the analysis required. For example, if the analysis requires access to individual-level data, then LDP would be the best choice to ensure that the data remains private. If the analysis is only focused on the trends in the data, then GDP can be used to ensure that the data is analyzed while preserving privacy. It is important to note that both LDP and GDP have their strengths and weaknesses, and choosing the right approach requires careful consideration of the specific use case and the desired level of privacy protection.

At the algorithmic level, differential privacy can be incorporated into ML solutions for precision medicine in several ways. One common approach is to use randomized data perturbation techniques, such as adding noise to the data or using random sampling, to ensure that individual patient records cannot be easily identified. These techniques can be applied to both the input data and the output of the ML model, helping to protect patient privacy at all stages of the analysis. Another approach is to use differentially private versions of ML algorithms, such as differentially private logistic regression or differentially private decision trees. These algorithms modify the standard ML algorithms to incorporate differential privacy constraints, such as adding noise to the data or modifying the learning algorithm to limit the amount of information that can be learned about individual patients. This can help to ensure that the ML model is trained on privacy-preserving data, while still achieving high levels of accuracy and performance.

Differential privacy can also be used in combination with FL approaches for privacy-preserving analysis of distributed healthcare data. Differential privacy can be incorporated into FL in several ways. One approach is to use differentially private stochastic gradient descent (DP-SGD) to train the shared model, which adds noise to the gradient updates to protect individual patient records. This can help to ensure that the model is trained on privacy-preserving data, while still achieving high levels of accuracy and performance. Another approach is to use differential private aggregation techniques to combine the local model updates from each party. This can help to ensure that individual patient records are not leaked during the aggregation process, while still allowing for effective collaboration and model training.

One of the key challenges associated with differential privacy is finding the right balance between privacy and utility. DP introduces noise or random perturbations to the data, which can make it more difficult to extract meaningful insights from the data. As a result, achieving a high level of privacy often comes at the cost of reduced accuracy and reliability of the analysis [19]. To mitigate this challenge, it is important to carefully design privacy-preserving mechanisms that can balance privacy and utility and to tune the parameters of these mechanisms to achieve the desired level of privacy while minimizing the loss of utility.

Another challenge associated with differential privacy is the potential for privacy attacks. Even when using DP, there is always a risk that an attacker may be able to infer private information about an individual by analyzing multiple versions of a dataset. For example, an attacker could use set intersection attacks to identify individuals who appear in multiple versions of the same dataset or use auxiliary information to learn more about individuals than is revealed in the DP-protected dataset. To address these challenges, it is important to carefully evaluate the privacy guarantees provided by differential privacy mechanisms, and to implement additional privacy protections, such as data masking or access controls, where necessary. Additionally, researchers must remain vigilant and adapt to emerging privacy threats and attacks to ensure that their DP mechanisms remain effective and secure.

3.4. Homomorphic encryption

Homomorphic encryption (HE) is a cryptographic technique that allows computation to be performed directly on encrypted data, without the need to first decrypt the data. In other words, homomorphic encryption enables secure computation on encrypted data, which can help to protect sensitive data from unauthorized access or disclosure. HE is a powerful tool for privacy-preserving computation, as it allows data to be processed and analyzed without ever exposing the raw data to anyone except the data owner. For example, in precision medicine, HE could be used to enable secure analysis of genomic data across multiple institutions, without the need to share the raw genetic data. Instead, each institution could encrypt their data and send it to a central location for analysis, while retaining control over their own data [20].

DOI: https://doi.org/10.54216/JCIM.080205
Received May 12, 2021 Accepted: November 11, 2021

There are several different types of HE, including fully homomorphic encryption (FHE), partially homomorphic encryption (PHE), and somewhat homomorphic encryption (SHE). FHE is the most powerful type of HE, as it allows arbitrary computations to be performed on encrypted data, but it is also the most computationally intensive and currently not yet practical for most real-world applications. PHE and SHE are less powerful, but more efficient and currently more practical for use in real-world applications. While HE is a promising tool for privacy-preserving computation, it is still an emerging technology and there are several challenges associated with its implementation. These include issues related to performance and scalability, as HE can be computationally expensive and slow, and may not scale well to large datasets or complex computations [10-17].

3.5. Blockchain

Blockchain technology has the potential to play a significant role in preserving the privacy of patient data in precision medicine. Precision medicine involves the use of individual-level data, such as genetic information, to tailor medical treatments and interventions to the specific needs of a patient. However, this data is sensitive and highly personal, and patients are understandably concerned about their privacy. Blockchain technology can offer a secure and decentralized way of storing and sharing patient data, which can help protect patient privacy. One of the main benefits of using blockchain technology in precision medicine is that it allows for secure and private sharing of patient data. Blockchain technology uses cryptography to secure the data, and once the data is entered into the blockchain, it cannot be altered or deleted. This means that patient data can be stored securely and accessed only by authorized parties, such as healthcare providers, researchers, and patients themselves. Additionally, because the blockchain is decentralized, there is no single point of failure or vulnerability, which further enhances the security and privacy of patient data. Another benefit of blockchain technology in precision medicine is that it can enable patients to maintain greater control over their own data. Patients can choose which data to share with whom, and they can revoke access to their data at any time. This can help patients feel more empowered and in control of their own healthcare, which can lead to better health outcomes. Moreover, patients can receive compensation for sharing their data with researchers or healthcare providers, which can incentivize them to participate in research and contribute to the advancement of precision medicine [8-7].

Blockchain technology can be classified into several types based on its accessibility and governance. These include Public Blockchain, Private Blockchain, Consortium Blockchain, and Hybrid Blockchain. Each of these types of blockchain has different characteristics and can be applied in various ways in the field of precision medicine. A Public Blockchain is a decentralized and open network that is accessible to anyone. This type of blockchain is entirely public, and anyone can participate in the network and contribute to the validation of transactions. Public blockchains are entirely transparent and offer high levels of security due to the large number of participants who verify transactions. In precision medicine, public blockchains can be used to store and share anonymized patient data, enabling researchers to analyze large datasets and develop new treatments and interventions. A Private Blockchain, on the other hand, is a closed network that is only accessible to a specific group of participants. In precision medicine, private blockchains can be used by hospitals, research institutions, or pharmaceutical companies to store and share sensitive patient data securely. Private blockchains are more centralized than public blockchains, which can increase the speed and efficiency of transactions [4-9].

A Consortium Blockchain is a hybrid between public and private blockchains. It is a decentralized network that is owned and controlled by multiple organizations. In precision medicine, a consortium blockchain can be used to enable the sharing of patient data among multiple healthcare providers or research institutions. The consortium blockchain can provide a secure and transparent way of sharing data while maintaining privacy and security. Finally, a Hybrid Blockchain is a combination of public and private blockchains. It is a flexible network that can switch between public and private modes depending on the needs of the application [9-11]. In precision medicine, a hybrid blockchain can be used to enable the sharing of anonymized patient data while maintaining the privacy of sensitive data. The hybrid blockchain can be configured to allow certain groups of participants to access private data while still enabling the public to contribute to the validation of transactions. In Table 4, we provide a comparative review of different types of Blockchains for precision medicine.

DOI: https://doi.org/10.54216/JCIM.080205
Received May 12, 2021 Accepted: November 11, 2021

Table 1: Summary of different types of Blockchain for precision medicine

Type	Access Restrictions	Transparency	Cost	Security	Privacy	Transaction rate	Incentive for mining	Architecture	Example	Use sector
Public Blockchain	Permissioned	Yes	High	High	Med	Slower	Yes	Decentralized	Bitcoin, Ethereum, Litcoin	medical research projects, medical record
Private Blockchain	Permissioned to access network	No	Med	Med	High	Faster	No	Partially decentralized	R3, Corda, B3i	medical supply chains, clinical trial
Consortium Blockchain	Permissioned	Little	Low	Med	High	Fastest	No	Partially decentralized or centralized	Monax, Multichain	shared patient record system
Hybrid Blockchain	Permissioned	Little	Med	Med	High	Fastest	No	Partially decentralized	Dragonchain, XinFin's	All of the above

4. Outlooks

Privacy preservation for AI in precision medicine is an important area of research that is constantly evolving. In this section, we introduce some future outlooks for privacy preservation in AI and precision medicine:

- Greater patient control over data: Patients are becoming increasingly aware of the importance of privacy and control over their health data. Future developments in precision medicine are likely to give patients greater control over their data, including the ability to choose who can access their data, how it is used, and for what purposes.
- Improved data-sharing agreements: As precision medicine research becomes more collaborative, there will be a need for better data-sharing agreements that balance privacy concerns with the need for sharing data. Future developments in this area are likely to involve the development of standardized agreements that protect patient privacy while enabling the sharing of data for research purposes.
- Increased use of blockchain technology: Blockchain technology is already being used in precision medicine to protect patient privacy by enabling secure and transparent sharing of data. Future developments in blockchain technology are likely to make it even more effective in preserving patient privacy and enabling the secure sharing of data among different stakeholders.
- More stringent regulations: Governments around the world are introducing more stringent regulations to protect patient privacy in the context of AI and precision medicine. As AI and precision medicine continue to evolve, policymakers may develop novel regulatory frameworks to govern the use of patient data in these fields. These frameworks could incorporate privacy-preserving measures and establish standards for data use and sharing that prioritize patient privacy.
- Emphasis on transparency and explainability: As AI becomes more prevalent in precision medicine, there may be an increased emphasis on transparency and explainability. This could help to ensure that patients understand how their data is being used and can make informed decisions about whether or not to share their data with researchers or clinicians.
- Using case studies in future research related to privacy preservation for AI in precision medicine can help to shed light on important issues and provide actionable insights for stakeholders. By using this approach, researchers

DOI: https://doi.org/10.54216/JCIM.080205
Received May 12, 2021 Accepted: November 11, 2021

can help to promote the responsible and ethical use of AI in precision medicine while ensuring that patient privacy and confidentiality are protected.

5. Conclusions

This review explores some of the key challenges and risks associated with the use of AI in precision medicine, as well as some of the strategies and techniques that can be used to mitigate these risks. We seek to highlight the importance of transparency, explainability, and patient control over their own data, as well as the potential benefits of different types of blockchain in preserving privacy. More, we explore different types of privacy-preservation methods in precision medicine along with the specifications associated with each of them in taxonomized fashion. While much work remains to be done in this field, the future direction for privacy preservation in AI and precision medicine is promising. As stakeholders across the healthcare ecosystem work together to protect patient privacy and confidentiality, we can ensure that the benefits of AI in precision medicine are realized while minimizing the potential risks to patient privacy. By continuing to explore these issues and develop effective strategies for privacy preservation, we can help to promote the responsible and ethical use of AI in precision medicine for the benefit of patients and society.

References

[1]. Yu, K. H., Hart, S. N., Goldfeder, R., Zhang, Q. C., Parker, S. C., & Snyder, M. (2017). Harnessing big data for precision medicine: infrastructures and applications. In *PACIFIC SYMPOSIUM ON BIOCOMPUTING 2017* (pp. 635-639).

[2]. Ziller, A., Passerat-Palmbach, J., Trask, A., Braren, R., Rueckert, D., & Kaissis, G. (2020). Artificial Intelligence in Medicine and Privacy Preservation. *Artificial Intelligence in Medicine*, 1-14.

[3]. Huang, L. C., Chu, H. C., Lien, C. Y., Hsiao, C. H., & Kao, T. (2009). Privacy preservation and information security protection for patients' portable electronic health records. *Computers in Biology and Medicine*, *39*(9), 743-750.

[4]. Zhang, Aiqing, and Xiaodong Lin. "Towards secure and privacy-preserving data sharing in e-health systems via consortium blockchain." *Journal of medical systems* 42, no. 8 (2018): 140.

[5]. Sahi, M. A., Abbas, H., Saleem, K., Yang, X., Derhab, A., Orgun, M. A., ... & Yaseen, A. (2017). Privacy preservation in e-healthcare environments: State of the art and future directions. *Ieee Access*, *6*, 464-478.

[6]. Pascual, D., Amirshahi, A., Aminifar, A., Atienza, D., Ryvlin, P., & Wattenhofer, R. (2020). Epilepsygan: Synthetic epileptic brain activities with privacy preservation. *IEEE Transactions on Biomedical Engineering*, *68*(8), 2435-2446.

[7]. Zhao, J., Chen, Y., & Zhang, W. (2019). Differential privacy preservation in deep learning: Challenges, opportunities and solutions. *IEEE Access*, *7*, 48901-48911.

[8]. Saleh, Y. N., Chibelushi, C. C., Abdel-Hamid, A. A., & Soliman, A. H. (2020). Privacy preservation for wireless sensor networks in healthcare: State of the art, and open research challenges. *arXiv preprint arXiv:2012.12958*.

[9]. Deebak, B. D., Al-Turjman, F., Aloqaily, M., & Alfandi, O. (2019). An authentic-based privacy preservation protocol for smart e-healthcare systems in IoT. *IEEE Access*, *7*, 135632-135649.

[10]. Sharma, N., & Bhatt, R. (2018). Privacy preservation in WSN for healthcare application. *Procedia computer science*, *132*, 1243-1252.

[11]. Wang, W., Chen, L., & Zhang, Q. (2015). Outsourcing high-dimensional healthcare data to cloud with personalized privacy preservation. *Computer Networks*, *88*, 136-148.

[12]. Chenthara, Shekha, Khandakar Ahmed, Hua Wang, Frank Whittaker, and Zhenxiang Chen. "Healthchain: A novel framework on privacy preservation of electronic health records using blockchain technology." *Plos one* 15, no. 12 (2020): e0243043.

[13]. 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.

[14]. Sengupta, J., Ruj, S., & Bit, S. D. (2020). A comprehensive survey on attacks, security issues and blockchain solutions for IoT and IIoT. *Journal of Network and Computer Applications*, *149*, 102481.

[15]. Mabkhot, M. M., Al-Ahmari, A. M., Salah, B., & Alkhalefah, H. (2018). Requirements of the smart factory system: A survey and perspective. *Machines*, *6*(2), 23.

[16]. Wen, Q., Gao, Y., Chen, Z., & Wu, D. (2019, May). A blockchain-based data sharing scheme in the supply chain by IIoT. In *2019 IEEE International Conference on Industrial Cyber Physical Systems (ICPS)* (pp. 695-700). IEEE.

DOI: https://doi.org/10.54216/JCIM.080205
Received May 12, 2021 Accepted: November 11, 2021

[17]. Zhao, S., Li, S., & Yao, Y. (2019). Blockchain enabled industrial Internet of Things technology. *IEEE Transactions on Computational Social Systems*, *6*(6), 1442-1453.
[18]. Liu, M., Yu, F. R., Teng, Y., Leung, V. C., & Song, M. (2019). Performance optimization for blockchain-enabled industrial Internet of Things (IIoT) systems: A deep reinforcement learning approach. *IEEE Transactions on Industrial Informatics*, *15*(6), 3559-3570.
[19]. Ghadge, A., Weiß, M., Caldwell, N. D., & Wilding, R. (2020). Managing cyber risk in supply chains: A review and research agenda. *Supply Chain Management: An International Journal*, *25*(2), 223-240.
[20]. Helo, P., & Hao, Y. (2019). Blockchains in operations and supply chains: A model and reference implementation. *Computers & Industrial Engineering*, *136*, 242-251.
[21]. García-Villarreal, E., Bhamra, R., & Schoenheit, M. (2019). Critical success factors of medical technology supply chains. *Production Planning & Control*, *30*(9), 716-735.

DOI: https://doi.org/10.54216/JCIM.080205
Received May 12, 2021 Accepted: November 11, 2021