

DOI: https://doi.org/10.48009/3_iis_2026_141

Clinical AI industrial solutions to data scarcity: Frameworks, challenges, and guidance

Wenli Wang, *Robert Morris University, and* wangw@rmu.edu

Chitra S. Kalla, *Robert Morris University, and* cxkst145@mail.rmu.edu

Dillon W. Shadowen, *Texas A&M University, and* dws11889@tamu.edu

Abstract

This paper investigates data scarcity as a primary challenge in clinical AI and evaluates two current industrial solutions: collaborative medical imaging AI training via the Medical Open Network for Artificial Intelligence (MONAI) and synthetic data generation via the Medical AI for Synthetic Imaging (MAISI). MONAI utilizes federated learning (FL) to enable multi-hospital distributed AI model training without transferring raw patient data outside respective institutions. MAISI generates synthetic medical imagery to train AI models, thereby mitigating privacy risks. This conceptual study identifies critical challenges associated with MONAI and MAISI, including data breaches, model inference attacks, interoperability issues, HIPAA non-compliance risks, and limitations in AI fairness and generalizability. The study proposes recommendations encompassing stronger differential privacy, enhanced data standardization, improved AI explainability, and stringent data quality validation for synthetic data. Federated learning, combined with HIPAA-compliant encryption protocols and robust explainability mechanisms, offers a promising pathway toward clinical AI deployment in hospital settings.

Keywords: clinical AI, MONAI, MAISI, HIPAA, federated learning, synthetic data generation

Introduction

Healthcare information systems have become essential components of modern healthcare delivery. These systems are responsible for storing patient records, clinical working processes, and communications among healthcare professionals. Electronic Medical Records (EMRs) have immensely enhanced the access and handling of patient data (Adler-Milstein & Jha, 2017). However, the growing amount of patient data is a double-edged sword, as data are not necessarily information or knowledge – data can both improve and hinder clinical decision making.

Artificial intelligence (AI), as one of the most powerful technological advancements today, has become an influential instrument in big data analytics to support clinical decision making. AI models can recognize patterns and predict trends in big data to assist in early diagnosis and treatment planning. AI has demonstrated effectiveness in identifying abnormalities such as tumors, fractures, and cardiovascular diseases from medical imagery (Esteva et al., 2019). AI has proven beneficial in facilitating diagnosis (e.g., in radiology, rare diseases) and improving efficiency (e.g., in summarization, medical reporting).

However, AI has not yet been widely adopted in clinical decision-making workflows. Strong evidence regarding AI model accuracy and safety is still required to make AI an integral part of clinical decision-making pipeline. Putting patients' lives in the hands of AI is a daunting task as the entire pipeline of diagnosis, severity assessment, and management of a disease, requires consistently high levels of AI model accuracy and safety throughout the clinical decision-making pipeline. AI alone probably will never become a sole clinical decision maker; however, AI has the great potential to become a competent digital assistant to medical professionals (doctors, clinicians, therapists, etc.), given sufficient rigorous evidence of AI's competency, especially in disease diagnosis.

To improve clinical AI model accuracy and safety, the first major step is to have sufficient high-quality data to train and validate AI models. This step alone is quite a challenging hurdle. For instance, there are plenty of multimodal data for training autonomous vehicles. Roads, vehicles, traffic signs, buildings, and even pedestrians, are open in the public for video recordings and image collection. The popular ResNet architecture was partially trained on cat and dog images collected from publicly available internet datasets. The large language models (LLMs) were trained literally on the entire Internet including historical books, journal publications, and social networks. Most AI frontier foundational models were trained on public data. In contrast there is a lack of publicly available medical data for clinical AI model training. Although some anonymized medical records and medical imaging datasets are publicly available such as Kaggle, these data are often limited in size and quality, hence insufficient to train clinical AI models to support clinical decision-making. This data scarcity problem in clinical AI reflects the effectiveness of existing privacy protections that restrict broad access to patient data. It implies that HIPAA compliance is working and patient data are mostly protected. Now the question is: how to solve the data scarcity problem in a highly-regulated industry for AI advancement yet still keeping data secure and private on-premises.

Currently, there are two solutions. One is to provide a secure technical architecture that allows for multiple-hospital collaboration in distributed AI model training without jeopardizing patient privacy. The other is to use synthetically generated medical data to train clinical AI models without involving any patient data, as long as the synthetic data mimic real-world clinical data. The latter solution has additional benefits -- cost-efficiently generating data that are difficult to be found in real-world yet crucial for AI model training to reduce biases and improve model accuracy.

These two solutions are not new; they are practical and are in use. However, practical adoption does not necessarily imply secure or reliable deployment. This paper identifies challenges and provides guidance how to address these challenges. While MONAI (Medical Open Network for Artificial Intelligence) and MAISI (Medical AI for Synthetic Imaging) offer promising solutions to clinical AI data scarcity, existing research has focused mainly on their technical capabilities rather than the challenges associated with real-world deployment. This study examines major challenges related to data security, patient privacy, system compatibility, AI reliability, and regulatory compliance while proposing implementation guidance for secure and HIPAA-compliant clinical AI systems.

RQ1: What governance, security, and operational challenges emerge when deploying MONAI-based federated learning and MAISI-based synthetic data systems for clinical AI development?

RQ2: What technical, organizational, and regulatory safeguards are necessary to ensure secure, HIPAA-compliant, fair, and clinically reliable deployment of MONAI and MAISI systems in healthcare environments?

This study discusses the current clinical AI industrial frameworks for solving the data scarcity problem in AI model training. It specifically shows how MONAI and MAISI can enhance clinical AI systems. Besides analyzing the existing technologies, the paper outlines some of the research opportunities that can further enhance AI model training without sacrificing security and privacy. It provides a pathway to HIPAA-compliant, stakeholder accepted clinical AI.

The paper is organized as follows: Section two introduces current clinical AI industrial frameworks by: 1) discussing a technical architecture called Federated Learning (FL) that enables secure and privacy-preserving multi-hospital collaboration in distributed clinical AI model training; 2) discussing FL-based MONAI for medical imagery AI; and 3) discussing MAISI for synthetic medical imagery generation and medical use simulation. Section three shows the conceptual answers to the research questions by: 1) identifying challenges in MONAI and MAISI; and 2) providing guidance to address these challenges. Section four has further discussions and section five concludes.

Background

The two current industrial frameworks for solving the data scarcity problem in clinical AI model training are: MONAI and MAISI. Nvidia plays a crucial role in establishing both frameworks and also complementary Clara (NVIDIA, 2023) and Holoscan (NVIDIA, 2022) technologies. For MONAI, there is also a collaborating non-profit organization that develops and manages the framework.

MONAI – Medical Imaging AI

MONAI enables multi-hospital collaboration in distributed AI model training without sacrificing patient data privacy. MONAI is probably the most famous healthcare application of federated learning. It is essential to have an understanding of federated learning as the basic technical infrastructure of MONAI. It tackles the inherent conflict between the need for AI to have a massive dataset to learn and the healthcare sector's requirement for patient privacy.

Federated Learning

Federated learning is not new; however, its application in healthcare AI training is relatively new. AI training built upon federated learning can overcome the data scarcity issue at one centralized data storage of one hospital. With healthcare data typically stored in isolated institutional silos, AI training usually takes place in one hospital with its own central repository of data. If multiple institutions deposit different dataset at one central location, this strategy poses a risk to privacy since patient records need to be moved or copied to other locations. To protect patient privacy, time-consuming data pre-processing is needed such as anonymization. Federated learning, in contrast, enables distributed AI training with collaborations among multiple institutions without moving or copying raw patient data – it is the model parameters that are transmitted between aggregation server and individual model-learning institution (e.g., hospital). Each model-learning institution provides model updates that are combined to formulate a global model at the aggregation server (Li et al., 2020). Hence, federated learning enables the training of AI models with collaboration among multiple institutions without sending sensitive patient information to the center (Li et al., 2020).

Federated learning would greatly minimize chances of breach of privacy. It allows collaboration in AI model training across organizations without compromising data privacy. Model updates rather than data are transmitted and AI model training is instead distributed across institutions and while maintaining privacy and allowing AI systems to be developed to high quality. Hospitals and research institutions can

collaboratively build more robust and less biased AI models without transferring confidential patient information outside their respective institutions.

Medical Imaging AI

Medical imaging is one of the most significant fields that have shown a solid potential of AI. Medical imaging technologies including computed tomography (CT), magnetic resonance imaging (MRI), ultrasound, and digital pathology produce large volumes of complex data which can be analyzed only with the help of sophisticated computational methods (Topol, 2019). Deep learning algorithms have been exceptionally effective in identifying the trends of medical images that otherwise cannot be readily observed with conventional analysis. With the further development of AI technologies, special frameworks that aim at helping healthcare research and development have been created to address medical imaging.

MONAI is one such framework. It is an open-source machine learning framework that offers functionality to create, train, and implement AI models for medical imaging (Cardoso et al., 2022). The framework allows researchers and clinicians to create comprehensive AI models for tumor detection, organ segmentation, and disease classification.

Imaging methods like CT and MRI scans are commonly used to produce these three-dimensional datasets and they have specialized processing methods. Also incorporated in the framework are automated model generation tools like Auto3DSeg which make the creation of segmentation models in medical imaging tasks easier by automating the model design, and training process (Cardoso et al., 2022).

MONAI Core, Label, and Model Zoo

The limitation within MONAI is addressed by providing tools that are optimized to work with healthcare imaging processes and clinical research settings (Cardoso et al., 2022). The MONAI Core has high-level medical image preprocessing, neural network training, and model evaluation. These tools assist the researcher to prepare imaging data sets through activities like normalization, image transformation, and data augmentation. Compared to other general machine learning systems, MONAI is optimally configured to process three-dimensional medical images, a frequent application of machine learning in radiology and pathology systems.

The other significant part of the MONAI ecosystem is MONAI Label that facilitates AI-assisted labeling processes. To label medical imaging with significant elements like a tumor, an organ, or an abnormal tissue, medical imaging datasets usually need manual labeling by a medical professional. This annotation procedure may prove to be very time-consuming and resource intensive and particularly when huge datasets are to be handled. MONAI Label helps to cope with this difficulty by ensuring that machine learning models are combined with annotation tools that guide experts in the labeling process. The system will be able to propose possible segmentations or mark areas of interest so that experts will be able to revise and fix annotations faster. This interactive method also saves time on the preparation of datasets and enhances productivity of the research in general.

The MONAI Model Zoo offers pre-trained AI models that researchers can utilize as the beginning point of new projects. The models are applicable to a diverse variety of medical imaging, such as disease detection and organ segmentation. Even though MONAI is a potent framework of healthcare AI development, a number of research opportunities are still open. The enhancement of integration of AI-assisted annotation tools with clinical workflows is one of the areas of research. The other opportunity is to come up with ways of enhancing model generalization in various healthcare facilities and imaging equipment. The solution of these issues will contribute to assuring that AI models will be dependable and effective in different healthcare settings.

In summary, the MONAI framework has emerged as a significant framework of formulating AI models in the domain of medical imaging. Created as a partnership among technology firms, scholars, and healthcare practitioners, MONAI offers specialized tools specifically used in medical data analysis. The framework aids in the development of deep learning applications which could be used to help in activities like disease detection, organ segmentation, and medical image classification. Since the medical imaging data are sometimes high dimensional and complex, it is possible that the traditional machine learning structures can be insufficient.

MAISI – Synthetic Imaging AI

MAISI is about the application of synthetic medical imaging in the development of machine learning models in situations where real medical datasets are inaccessible because of privacy, regulatory, or data availability issues (Guo et al., 2024). Regular medical imaging technologies like CT/MRI scans and ultrasound images produce vital diagnostic data to the physician. But these medical data for diagnosis are mostly unavailable for AI model training since the data about patients should be confidential, and institutions might not be able to share their data freely. Moreover, the creation of precise AI technologies demands extensive amounts of high-quality data, which can be very challenging to acquire. MAISI helps solve this problem by generating realistic and high-quality synthetic medical images that can be utilized in research and model training without revealing sensitive patient data.

Synthetic data are those datasets that are created artificially following the characteristics and the statistical trends of real patient data. In medical imaging, synthetic images can imitate anatomy, disease patterns, and changes in imaging while preserving the privacy of real patients. Generative AI models play an important role in producing such data. Generative models, including generative adversarial networks (GANs) and diffusion models, can learn the patterns of real medical images and then generate synthetic images that have the same features, covering a range of clinical situations such as disease stages, imaging conditions, and anatomy differences. Through producing such data, MAISI enables researchers to produce better balanced datasets that enhance the training of AI models, since skewed or insufficient datasets can result in wrong predictions when models are implemented in actual clinical settings.

MAISI generates synthetic medical imaging data that can be used to train machine learning models that segment, classify, and detect diseases in images. Image segmentation refers to the process of recognizing and isolating various structures in medical images, e.g., organs, tissues, or tumors. MAISI helps this process by producing synthetic images with anatomical structures labeled, which can then be utilized to train a segmentation algorithm. One of the aims is for segmentation models to automatically detect tumors in radiology scans with high accuracy in real clinical settings, assisting and enabling clinicians to assess disease progression or treatment results.

Besides facilitating the training of models, MAISI is significant in solving the issue of data scarcity in healthcare studies. Large medical imaging datasets are not readily available in many healthcare institutions, particularly in the case of a rare disease or a specialized medical condition. The artificial creation of data enables researchers to scale up the size of data they have and model scenarios that are not necessarily representative of real-patient data. Synthetic imaging also helps safeguard the privacy of patients, since it enables researchers to train models without exposing identifiable patient information and without the lengthy regulatory approvals and data management that using actual patient data requires.

Nvidia Clara and Holoscan

Besides MONAI and MAISI frameworks, there are other technologies assisting the development and deployment of clinical AI. For instance, Nvidia Clara and Holoscan offer computing infrastructure and deployment environments to AI models.

The Nvidia Clara healthcare AI platform is created to also enhance the creation of medical imaging. The platform offers its users the power of computing with the help of GPUs, which enables researchers to train more intricate deep learning models on large medical datasets (Nvidia, 2023) as medical imaging systems produce data in large volumes which must be processed by high performance computing.

Another technology that can support real-time AI processing in the clinical setting is Holoscan. This platform provides an opportunity to deploy AI models directly into healthcare systems where rapid analysis is needed. Real-time imaging analysis can help the surgeons during medical procedures and make AI-assisted clinical decision-making possible.

Table 1 provides a comparison of key clinical AI industrial practices including technology, purpose, key benefits, and challenges. The following section will further discuss these and other related challenges.

Table 1. Comparison of Key Clinical AI Industrial Practices

Technology	Purpose	Key Benefits	Challenges
MONAI	Framework supporting distributed, collaborative AI model training for medical imaging AI	Protect patient data privacy and reduce centralized storage risks Specialized tools for medical image processing and segmentation	Require large datasets and computational resources Communication overhead and model synchronization issues
MAISI	Synthetic medical imaging data generation	Address data scarcity issue, especial on rare diseases No risk to patient privacy	Risk of artificial patterns Require clinical validation
Nvidia Clara	Healthcare AI computing platform	Accelerate model training using GPU computing	High infrastructure cost
Nvidia Holoscan	Real-time clinical AI deployment	Enable rapid image analysis during clinical procedures	Integration with existing systems can be complex

Results

RQ1: What governance, security, and operational challenges emerge when deploying MONAI-based federated learning and MAISI-based synthetic data systems for clinical AI development?

The technologies discussed above can be combined to form potent healthcare AI systems. From a governance perspective, these technologies reveal a broader set of organizational and operational challenges that extend beyond technical AI development and require systematic evaluation before large-scale clinical adoption.

Each technology independently presents its own challenges (see Table 1 in the above section). In addition, the application of modern computing systems also presents new security and privacy issues (see Table 2) which require new guidance. For instance, despite the fact that privacy-preservation architecture like federated learning and the frameworks like MONAI and MAISI have made strides in clinical AI, new risks

associated with these frameworks must be managed before they become a common practice in real-world healthcare settings.

Table 2. Challenges in Clinical AI Industrial Practices that Address Data Scarcity Problem

Challenges	Description	Possible Impact
Data breaches	Unauthorized access to healthcare datasets and model parameters	Exposure of sensitive patient information and/or proprietary AI model parameters
Data quality and interoperability risks	Inconsistent data formats among institutions	Reduced model performance due to non-interoperable data
AI model inference risks	Biased AI models and unfair model inference	Reduced model inference fairness and applicability
Regulatory non-compliance	Failure to meet HIPAA security requirements	Legal penalties and loss of patient trust

Data Breaches

In medical institutions, data breaches can reveal very sensitive data, e.g., medical history of the patients, diagnostic reports and personal identification information that can lead to severe legal, financial, and ethical repercussions. AI model training requires large volumes of data and data storage, which often have vulnerabilities in security and privacy. The increasing application of AI to the healthcare sector poses more privacy and security threats. Big amounts of healthcare data are needed to train and validate AI models, which expose sensitive patient information to a greater extent. Unless these datasets are secured accordingly, they can be an enticing target of cyberattack or unauthorized access.

Extensive research has demonstrated that deep learning models perform well in medical imagery analysis. However, there is still limited research on how these models can be securely integrated into healthcare systems. Despite AI advancements in functionalities, security weaknesses remain, especially in collaborative AI frameworks. Even though federated learning will reduce the necessity to exchange raw patient data, attackers can still seek to obtain sensitive information through the shared AI model updates.

One significant concern is that attackers may extract sensitive information from machine learning models. Attackers have been demonstrated to be capable of inferring or reconstructing some of the training data based on a model output or parameter, particularly when the models are publicly available or unprotected (Shokri et al., 2017). Such attacks also point out weak points that AI systems might have the necessity of more solid security measures in their development and implementation.

Data Quality and Interoperability Risks

Data quality and interoperability are the key to the successful AI model training. Healthcare data are usually gathered through a variety of channels such as hospitals, diagnostic laboratories, imaging centers, and wearable gadgets. Various imaging devices, data standards, and storage systems are frequently used in hospitals and research institutions. These differences may impact federated learning model performance and cause interoperability issues.

From a governance perspective, interoperability limitations create cross-institutional coordination challenges that undermine collaborative AI development. Since federated learning depends on cooperation between multiple healthcare organizations, inconsistent data standards and incompatible information systems can weaken both technical performance and organizational trust among participating institutions.

The discrepancy in records and norms may obstruct successful data integration and review (Kellermann and Jones, 2013). The quality of data may also have a harmful effect on the performance of AI models resulting in incorrect predictions and unreliable outcomes.

Synthetic medical imaging generation raises new issues regarding data quality. It is important to assess its data quality. The production of synthetic images would be of less value if the images are not reflective of real medical conditions. When synthetic images are not realistic, machine learning models trained on them could do poorly when they are applied on real patient datasets.

AI Model Inference Risks

AI models also need improvement even during inference. Most AI models are trained on datasets in a small subset of geographic scope or of limited number of healthcare institutions. Consequently, such models may not work well with other populations of patients. These limitations can lead to biased AI models, with subsequent unfairness during inference.

Ethical concerns should also be addressed in the creation of AI technology in healthcare. Synthetic data contributes to patient privacy protection, but still, there is a need to ensure that AI models trained from synthetic data are responsible and do not add bias to the medical decision-making process.

Regulatory Non-compliance

The Health Insurance Portability and Accountability Act (HIPAA) in the United States is an act that sets special requirements of safeguarding patient health data in the U.S. It develops the rules of patient confidentiality of their health information, as well as the privacy and security of the management of the healthcare data (Office for Civil Rights, 2013). The objectives of HIPAA regulation are to ensure the confidentiality, integrity, and availability of electronic health information and at the same time guarantee access to healthcare providers to needed data to care about patients.

Healthcare organizations should make sure that AI systems do not violate regulatory provisions associated with auditing, transparency, and accountability. Despite healthcare AI model development, the challenging question of HIPAA compliance has not been fully solved by either MONAI or MAISI, although both have provided pathways to accomplish HIPAA compliance.

Healthcare organizations should put in place administrative, technical, as well as physical protection measures that would prevent unauthorized access, disclosure, and misuse of patient data (Office for Civil Rights, 2013). These protection measures involve secure storage of data, access control measures, employee training, and constant supervision of information systems.

RQ2: What technical, organizational, and regulatory safeguards are necessary to ensure secure, HIPAA-compliant, fair, and clinically reliable deployment of MONAI and MAISI systems in healthcare environments?

Healthcare AI systems should ensure that a clear record is kept of the usage of data, model training, and decision generation. It may be challenging to ensure that these regulatory standards are met, especially when training AI models on data that represent more than one institution or distributed system, like federated learning environments. Consequently, designing safe AI systems and effective policies related to governing AI should be considered as the key to retaining patient trust as well as allowing the responsible use of AI technologies in healthcare.

Strong Security Posture and Privacy Preservation

Data breach risks require more powerful security measures to prevent data spillage and AI model parameter manipulation or loss, as well as to safeguard the integrity of data and the AI models. Stronger differential privacy, multi-party distributed computation security, and AI model encryption techniques are all pathways to HIPAA compliance.

Differential privacy applies a statistical noise to avoid identification of individual records in datasets (Dwork and Roth, 2014). Secure multi-party computation enables two or more parties to mutually compute with encrypted information without letting the information be known to them. Integrating federated learning with such methods with strong encryption for data at rest and in transit would be able to develop safe AI solutions that do not violate healthcare policies.

Standardization for Data Quality and Interoperability

There is a continuing need for the development of standard healthcare data frameworks that facilitate the smooth exchange of data maintenance of security and regulatory compliance. Better interoperability standards and data governance models will also be needed to support large-scale healthcare AI research. Creation of standardized models can also enable the federated learning systems to effectively operate in a wide range of healthcare systems.

For instance, standardized healthcare data formats like HL7 FHIR (Fast Healthcare Interoperability Resources) can significantly decrease integration friction between institutions in federated learning. It is a next-generation technical standard created by Health Level Seven International (HL7) designed to enable the rapid, secure, and efficient exchange of healthcare information. It bridges the gap between disparate EHR systems using modern web technologies (ONC, 2021).

As for synthetic data generation, investigators should consider the realistic resemblance, heterogeneity, and clinical utility of artificial datasets. Synthetic datasets should reflect a variety of patient populations to ensure that AI models can work properly with populations that are different. Evaluation methods may include visual inspection by medical professionals, statistical comparison of the synthetic and real data, and the performance testing of the AI models trained with the synthetic data. Such assessment methods can be used to guarantee that the synthetic data produced by MAISI is effective and can be used in medical studies.

Synthetic data may enhance the development of AI models, but it should be done cautiously to eliminate the artificial patterns that are absent in clinical data. The researchers should thus be cautious about the generation models and have their models tested by clinical expertise. The incorporation of synthetic data into healthcare research necessitates the cooperation of computer scientists, clinicians, and data engineers to make sure that the produced data resembles the medical realities.

AI Model Inference Performance Monitoring

To address the AI model inference risks, better generalizability and fairness of healthcare AI systems should be involved to make sure that they are able to deliver reliable results to different groups of demographics. AI model inference performance needs to be closely monitored so that proper “circuit breaker” is in place in case there is a detection of bias and unfairness in AI model inference. There is also a need for better evaluation methods of AI performance in healthcare settings.

The other research opportunity is human-AI collaboration within the clinical workflow. Even though the AI systems are capable of performing complex diagnostic tasks, AI should always be supervised by humans to prevent AI hallucinations leading to unfavorable consequences. In addition, the medical staff should be able to understand and interpret the results of AI systems to ensure proper supervision.

Explainable AI for Transparency, Trust, and HIPAA-compliance

To establish trust within healthcare professionals and patients, it is necessary to ensure the healthcare AI systems are fair and transparent. The study of explainable AI methods can be used to enhance transparency and trust in healthcare AI products.

Explainable AI systems provide rationales for AI predictions, which enables clinicians to see how an AI model arrives at a specific conclusion. Such transparency is essential in order to develop trust in AI-based diagnostic systems, as well as ensure accountability in medical decision-making.

The adoption of robust governance frameworks of healthcare AI systems helps establish data access policies, model development policies, system auditing policies and regulatory compliance policies. With the increased adoption of AI technologies in clinical settings, policymakers need to work out policies that would make the use of AI safe, ethical, and responsible and foster innovation.

Discussion

The technologies of AI, including MONAI and MAISI, can provide healthcare systems with significant opportunities for improvement and promote medical research. Such technologies can revolutionize the medical imaging analysis processes, as they can facilitate the speedy processing of large datasets and enhance the precision of the disease identification. AI models can assist medical professionals in identifying patterns in complex medical images, and this could aid in the early diagnosis and better patient outcomes.

Besides that, special AI frameworks like MONAI include specialized tools that enable the development and deployment of medical imaging models more effectively. When it is applied alongside distributed learning technologies like federated learning, healthcare organizations will be able to cooperate in creating superior AI systems without violating patient privacy. Among the most significant benefits of these technologies, there is the possibility of supporting the large-scale collaborative research. Federated learning enables various hospitals and research centers to engage in the construction of machine learning models but retain their sensitive data in the institutional contexts. This is because it minimizes the necessity to share confidential information about patients across organizations and assists institutions to adhere to privacy rules like the HIPAA.

Consequently, federated learning may help healthcare institutions to create more robust and precise AI models using a variety of datasets gathered in a range of clinical environments. Although these technologies have these benefits, they still face a number of challenges that need to be resolved before these technologies are extensively adopted in healthcare systems. Lack of interoperability between healthcare information systems is one of the greatest barriers. Different electronic healthcare record systems, data storage formats and imaging standards are often used in healthcare organizations. The differences complicate the process of sharing and integrating datasets between institutions. In the absence of standard data formats and communication protocols, joint work in AI research is more complicated and inefficient. Future studies must center on coming up with better data standardization models and data interoperability models that will enable healthcare systems to share data in a safe and efficient manner.

The other problem is associated with the generalizability of AI models. Most of the machine learning models are being trained on the data gathered in particular hospitals or areas. Consequently, these models might perform poorly when used in new settings with other populations of patients or imaging equipment. Indicatively, differences in imaging equipment, scanning procedures, and patient demographics can affect

the accuracy of AI predictions. Researchers should come up with models that can adapt to different datasets and healthcare settings in order to solve this problem.

The problem of security and privacy is also critical in healthcare AI systems. Federated learning eliminates the necessity to transmit raw data; however, it does not exclude the security risks. There is a danger of malicious actors trying to interfere with the updates to the models or obtain sensitive data out of the trained models. Privacy preservation methods like differential privacy and encrypted communication protocols are already incorporated into federated learning models. However, for clinical care and HIPAA compliance, these privacy and security measures need to be carefully examined and improved.

Besides the technical issues, human-AI collaboration is an essential factor that needs to be considered when implementing AI systems in clinical healthcare. To continue building trust between healthcare providers and patients, it will be necessary to improve the security of AI systems. AI technologies cannot be regarded as substitutes of medical workers but can be regarded as an aid to clinical decision-making. Physicians and other healthcare professionals must remain actively involved in interpreting AI-generated outputs and making final treatment decisions. This is why the explainable AI models are a significant field of research.

Moreover, the application of AI technologies in clinical practice presupposes both modifications in organizational practices and employee training. The medical workers should be educated to learn about the abilities and the shortcomings of AI systems. Hospitals may also need to upgrade existing infrastructure to support AI applications to meet their computational needs. The mentioned organizational aspects provide a clear relevance to interdisciplinary cooperation between healthcare practitioners, engineers, information scientists, and policy-makers.

In summary, the combination of AI architectures like MONAI and MAISI, together with other technologies such as Nvidia Clara and Holoscan, can enhance the quality of diagnoses, facilitate the collective medical research, and aid the progress of the precision medicine. Nonetheless, further research is still needed in order to resolve issues connected to interoperability, model reliability, security, and human-AI collaboration. Addressing these challenges with suggested recommendations, healthcare organizations can create AI systems that are innovative and responsible, which will eventually enhance the quality and efficiency of healthcare provision.

This study contributes to healthcare AI research by examining the challenges that arise when AI systems are used in real clinical settings. By focusing on security, privacy, interoperability, AI performance, and regulatory compliance, the study provides guidance on the clinical use of healthcare AI systems beyond their technical capabilities

Conclusion

The field of artificial intelligence is reshaping clinical healthcare by enabling the analysis of complex medical data at unprecedented speed and sophistication. Frameworks like MONAI and MAISI offer feasible solutions to train and validate AI models capable of supporting medical research and clinical decision-making.

Nonetheless, there is a strong privacy and security concern in AI usage in the healthcare field. There are still challenges related to HIPAA compliance. Future research needs to continue enhancing the scalability, interoperability, and security of clinical healthcare AI systems. Partnerships among healthcare

organizations, researchers, and technology developers will further strengthen the effective, responsible, and safe AI applications in clinical healthcare.

Acknowledgements

This work was conducted as part of the NAIRR Pilot project NAIRR250299 ("Integrating Video AI Analytics in Symani Robotics Assisted Microsurgery"), informing the project's technical-architecture and federated-learning design for multi-site model development.

References

- Adler-Milstein, J., & Jha, A. K. (2017). HITECH Act drove large gains in hospital electronic health record adoption. *Health Affairs*, 36(8), 1416–1422.
- Cardoso, M. J., et al. (2022). MONAI: An open-source framework for deep learning in healthcare imaging. *Nature Methods*, 19, 365–368.
- Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science*, 9(3–4), 211–407.
- Esteva, A., et al. (2019). A guide to deep learning in healthcare. *Nature Medicine*, 25(1), 24–29.
- Guo, P., Zhao, C., Yang, D., Xu, Z., Nath, V., Tang, Y., Simon, B., Belue, M., Harmon, S., Turkbey, B., & Xu, D. (2024). *MAISI: Medical AI for synthetic imaging*. arXiv. <https://arxiv.org/abs/2409.11169>
- Kellermann, A. L., & Jones, S. S. (2013). What it will take to achieve the as-yet-unfulfilled promises of health information technology. *Health Affairs*, 32(1), 63–68.
- Li, T., Sahu, A., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, 37(3), 50–60.
- NVIDIA. (2023). NVIDIA Clara: Healthcare AI platform overview. Retrieved on July 9th, 2026 from <https://www.nvidia.com/en-us/clara/>
- NVIDIA. (2022). NVIDIA Clara Holoscan: Real-time AI computing for medical devices. Retrieved on July 9th, 2026 from <https://www.nvidia.com/en-us/edge-computing/holoscan/>
- Office for Civil Rights. (2013). Summary of the HIPAA privacy rule. U.S. Department of Health & Human Services.
- ONC (2021) What is HL7 FHIR? Office of the National Coordinator of Healthcare Information Technology. Retrieved on May 24, 2026 from <https://healthit.gov/wp-content/uploads/2021/04/What-Is-FHIR-Fact-Sheet.pdf>
- Shokri, R., Stronati, M., Song, C., & Shmatikov, V. (2017). Membership inference attacks against machine learning models. *IEEE Symposium on Security and Privacy*, 3–18.

Topol, E. (2019). Deep medicine: How artificial intelligence can make healthcare human again. Basic Books.