BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

FEDERATED LEARNING WITH GENERATIVE AI FOR PRIVACY-PRESERVING MEDICAL DATA ANALYSIS : A Comparative Review of Recent Approaches and Open Gaps

Shreya Satyapal Bansode

Department of Artificial Intelligence and Data Science Engineering
Ajeenkya DY Patil School of Engineering, Pune, India

Prof. Neelam Jain

Assistant Professor
Department of Artificial Intelligence and Data Science Engineering
Ajeenkya DY Patil School of Engineering, Pune, India

Abstract: Hospitals possess the necessary data for training accurate diagnostic models, but privacy regulations prevent this from happening. Combining the approach of federated learning (FL) with generative models allows for expanding and balancing the training data without the need to change the location of patient data. However, the given paper describes five recent studies in the sphere of combining the above-mentioned studies: an approach of differentially private federated GAN, a review of the latest approaches in federated generative models, an empirical study of the use of federated models in creating tabular data in the field of leukaemia, a model of latent diffusion instead of using the FL approach in MRI of the brain and a one-shot approach of FL based on the use a flow model with rectified flow. In this case, the studies have been compared according to the type of federation, type of generative model, as well as the practice of privacy and evaluation. The main conclusion of the comparison is that while the way of using the FL changes from iterative one to one-shot one, the types of evidence used in privacy claims are different (e.g. formal differential privacy or any empirical or information.

Keywords: federated learning, generative models, differential privacy, synthetic data, medical imaging, diffusion models.

Introduction:

Medical ML relies heavily on big heterogeneous datasets that are limited by regulations and policies concerning the confidentiality of hospital data, thus requiring hospitals to rely on local resources. Federated learning (FL) attempts to solve this dilemma to some degree by allowing local training using local datasets

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

and sharing the model and its parameters instead of the original patient data [6]. However, FL by itself does not solve the problem of data scarcity since the data distributions at different locations may differ. Generative modelling methods, beginning with the invention of GANs [8], can increase both the size and homogeneity of the local datasets by creating new synthetic samples. Combining the two gives federated generative models (FGMs).

Based on a recent review article, the authors present an overview of how generative adversarial networks (GANs) and diffusion models were employed in medical imaging studies. In total, they screened 319 records and identified 60 articles related to FGMs. There are various factors that impact this research area. First of all, privacy is one of the aspects to be taken into account, as the data obtained from image generation through GANs can contain information about training images. Secondly, communication is an issue that must be addressed because an iterative use of Federated Learning (FL) requires a larger number of iterations, which involves spending a lot of capital for hospitals using these methods. Some of the works focusing on the impact of privacy on federated GANs are [1] and FedCSCD-GAN [16]. The federated environment is also being extended to personalised medical audio-visual models [12], interpretable medical audio models [13], and longitudinal report generation [14], while foundational models are applied to create differentially private synthetic data that counters non-IID clients [15].

Recent articles address the challenges differently, making them hard to align with each other. As such, this review tries to answer three questions. How does the literature organise federation and generation? What evidence do they use to back up their claims regarding privacy? How reliable are their evaluations of utility? The contributions of this article are as follows: i. A structured comparison of five studies; ii. An analysis of how privacy and fidelity are presented; iii. A checklist for reporting future studies. The article is a review of the literature and does not contain experiments. The figures cited are from original articles and were obtained on different benchmarks, thus making it impossible to compare them directly.

II. BACKGROUND

A. Federated Learning

The concept of federated averaging (FedAvg) [6] involves training by each client on local data and averaging the weights returned from the clients for several rounds. Thus, when the clients differ in the label distribution, image protocol, or number of samples (non-IID situation), the local updates diverge from each other, and the process of aggregation becomes impaired. One-shot FL approaches the problem by using a single exchange, which complicates convergence in a non-IID situation.

III. REVIEWED STUDIES B. Generative Models

A GAN is a type of neural network that employs a generator and a discriminator and trains the former against the latter until it is impossible to distinguish between synthetic and real examples [8]. The diffusion model is a process of adding noise to the data step-by-step and learning how to reverse the process.

C. Privacy

Differential privacy may be defined formally with a pair (ε, δ), which tells us how much impact one record has on the output; the smaller the ε value, the more protection is available. In deep learning, the technique is often applied through gradient clipping and adding Gaussian noise, keeping track of the accumulated losses through the moment accountant [7]. Alternatively, one can take an empirical approach through

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

membership inference, attribute inference, likability or singling-out attacks against the released dataset or by arranging generated samples to resemble training samples [10]. Both paths give answers to different inquiries and yield differing outcomes.

D. Evaluating Synthetic Data and Simulating Heterogeneity

To evaluate synthetic data, fidelity (how accurately it captures real data statistics), utility (how well the models created using this data can perform real-world tasks), and privacy are typically considered [3]. Image analysis often employs the Fréchet Inception Distance (FID), where lower values imply that the distributions are closer [1], [4], [5]. In contrast, tabular data studies include correlation similarity, diversity scores, and detectability tests, where statistical classifiers are applied to the data to find out if the rows/columns come from the real or synthetic dataset; in this situation, higher detection scores demonstrate lower fidelity [3]. As real federations are difficult to access, a majority of the studies have to simulate heterogeneity by splitting an original dataset. Usually, some classes are given to each client, or the distribution of the classes in a dataset is generated by means of the Dirichlet distribution, as it was done in [1] and [3]. In study [5], sharding partitioning was applied to the two datasets and one natural federation in which each individual region works as a client. Partitioning is instrumental in drawing conclusions from research findings, and that is why it has to be included in the description.

A. Differentially Private Federated GAN [1]

A differentially private federated generative adversarial network has been developed by Li, Tong, et al., which implements five convolution layers in both the generator and discriminator of its GAN model at every participating site. Hospitals are arranged in organisations such that only some of them communicate with the server directly. Local differential privacy is enforced by truncating the discriminator's gradients and then corrupting them with Gaussian noise that decreases gradually over the rounds. The entire (ε, δ) budget is monitored using a moment accountant. The experiments utilise the MNIST dataset and a COVID CT dataset of 1070 patients. The scale of noise begins at 10 and decays by 0.02 every round; the GAN is trained for 400 global rounds using 25 local epochs per round. The experiments establish FedGAN, FedDPGAN, and PerFedGAN as baseline models. According to the authors, at an ε value of 8, the results obtained were insignificant when compared to the original model, but at an ε value of 2 the results were modestly poor both for the generated images and the corresponding classifier.

B. Survey of Federated Generative Models [2]

Mahmood et al. adhere to the PRISMA standard and categorise the audience by the generative method: GANs, diffusion techniques, encoder-decoder strategies and others. The types of GAN construction are further classified according to the nature of the hosting for the techniques, whether it is an entire technology, a single component technology, or neither. The challenges have diversified into FL-oriented, data-oriented and model-oriented. The authors point out that the majority of research makes use of simulations with benchmark datasets and that there are almost no clinical trials performed. The types of GANs offered include client-side alone (without aggregation), generator and discriminator both at the client and server, one technology being at the server, and decentralised approaches on the basis of peer-to-peer exchange. The authors also classify diffusions according to the location of synthetic data, which can be at the clients, at the main server, or in decentralised databases. The themes discussed are augmentation, MRI reconstruction, cross-modality conversion, classification and segmentation.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

C. Federated Tabular Synthesis in Leukaemia [3]

Issa et al. assess the costs involved in federating a synthetic-data generator. They train CTGAN and FedTabDiff using data from 1,540 patients with acute myeloid leukaemia, each represented by 12 clinical features, while data is distributed between 3 to 10 nodes utilising FedAvg, applying one of three data splits (balanced, imbalanced, and non-IID). They measure fidelity with correlation, diversity, distribution-distance, and classifier metrics, and evaluate privacy with four attacks. DP is not employed. Detectability of the generator with CTGAN moved from an AUC of 0.796 in the centralised setting to an AUC of 0.94–0.99 for federated generation, indicating easier detection of generated rows. FedTabDiff shows a high level of detectability at an AUC of ~0.99 even in the centralised baseline. Authors see very few consistent effects of the number of nodes and type of split, including a single dataset and absence of DP among limitations.

D. CATphishing [4]

In their study, Truong et al. moved away from the use of iterative federated learning (FL) and adopted a synthetic data approach. In this new method, each site trains a 3D latent diffusion model that can produce four different MRI contrast images as well as a mask for detecting the tumour, depending on the tumour type or the mutation status served as the basis of training. Once the models are developed, they are sent to a central server, which generates samples and creates a single classifier. The study received data from 2,491 patients in seven hospitals/clinical institutions, with three sites contributing data for model training and five sites reserved for testing the generalisation abilities. The autoencoder compresses 128×192×192 volumes with five channels into an 8×16×32×32 latent. The whole training process was conducted using 4 A100 GPUs. In order to heighten the diversity of the data gathered, 1,600 samples per class per site were generated, and then some samples that were too close to others were eliminated using the radiomics distances. As a result, only 2,087 mutated and 2,527 wildtype samples remained for testing.

The classifier, which is a residual U-Net combined with a Mamba module, has been trained using fivefold cross-validation. Performing poorly on separating oligodendroglioma from astrocytoma across all methods managed to yield 76.2% accuracy in the case of centralised training and 75.2% in both federated and synthetic training in tumour-type test 2.

E. One-Shot FL with FG-RF and DLKD [5]

Ma et al. propose generating features instead of pixels. Each client trains a light feature extractor and then a rectified-flow model with a transformer backbone on the extracted features. The server samples synthetic features and trains a global student by distilling both the output logits and an intermediate layer of the averaged client classifiers. Evaluation covers tuberculosis chest X-rays, brain tumour MRI and HAM10000 skin lesions with three clients. The extractor is a single 3×3 convolution with a tanh activation trained through a ResNet-18 classifier, and the flow model is a diffusion transformer. On brain tumour data, FG-RF trained in 1.87 h against 4.51 h for a pixel-level DDPM and generated one sample in 0.45 s against 22.68 s. For privacy, the authors argue that the information a generated feature holds about the input cannot exceed what the extracted feature holds, and they measure the L distance between generated and training samples: 2 feature-level generations sit farther from the training set than pixel-level generation (for example, 0.5195 against 0.3793 for one brain tumour client).

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

Ref.	Data	Generator	Federation	Privacy basis
[1]	MNIST; COVID-CT	GAN	Iterative, grouped FedAvg	Local DP + moment accountant
[2]	Survey (60 studies)	All types	Review	Review only
[3]	AML tabular (1,540)	CTGAN; FedTabDiff	Iterative FedAvg	Attack-based; no DP
[4]	Brain MRI (2,491)	3D latent diffusion	One-shot	SSIM membership attack
[5]	X-ray; MRI; skin	Rectified flow (features)	One-shot + distillation	Feature-level generation; distance check

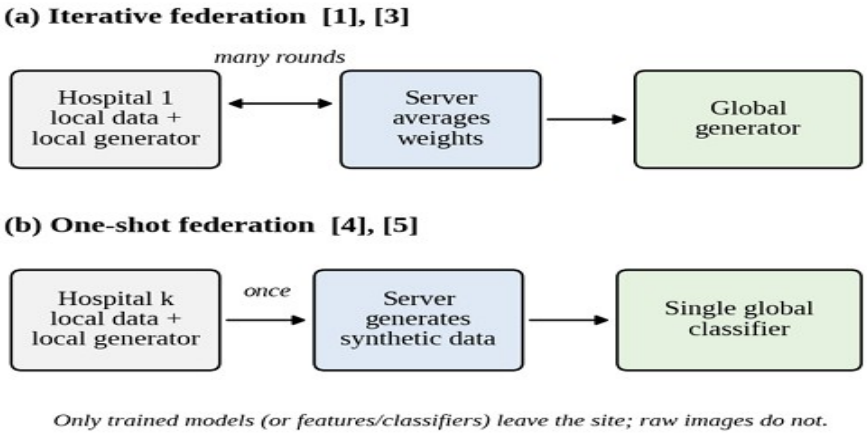


TABLE I SUMMARY OF THE REVIEWED STUDIES

Fig. 1. Iterative federation [1], [3] compared with one-shot federation [4], [5].

IV. COMPARATIVE ANALYSIS

A. Federation Style

In the studies [1] and [3], the classical iterative style remains intact, with a total of 400 and 500 rounds globally, respectively. In studies [4] and [5], the models are exchanged once. These one-shot designs allow for communication efficiency, but they lead to the generator being a single point of failure, as later training can't remedy a poorly performing generator. Additionally, they send the trained model to the server – in study [4] it is the generator and in study [5] it is the flow model and classifier.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

B. Utility Against Centralised Training

Ref.	Reported result
[1]	Accuracy on synthetic COVID-CT at (ε, δ) = (8, 10^{-5}): 0.8982 with 1 client, 0.9254 with 15 clients; FID 210.63 to 192.76.
[3]	Federation lowers fidelity significantly (P<.001): up to 21% for CTGAN and 62% for FedTabDiff; privacy mostly unchanged.
[4]	IDH accuracy: 96.2% centralised real, 95.8% FL, 95.5% synthetic; AUC 0.979, 0.980, 0.966.
[5]	Accuracy (centralised / FedAvg / FG-RF): brain tumour 75.64 / 36.95 / 60.15; chest X-ray 86.11 / 54.88 / 86.32; HAM10000 73.43 / 55.08 / 70.01.

TABLE II HEADLINE RESULTS AS REPORTED (DIFFERENT BENCHMARKS)

The most impressive difference can be seen between [3] and [4]. For example, on the small tabular data, the generator performance in the federation process is decreased [3], while for 3D MRI, the performance of the synthetic data-trained classifier is only about 1% lower than the performance that was achieved with the real data [4]. These methods use different metrics, data and tasks, so the reason of the gap is not only the modality. he research in [5] achieves results that are almost equal to the centralised ones for chest X-rays, while it has performance that is 15 -15% lower among the results for brain tumour MRI. The importance of the privacy budget is important as well, also regarding the results in [1]. For example, the two federated approaches in [4] achieved results that are within 1% of those of centralised training, whereas the multi-round FedAvg approach in [5] lost 18-39%. It is possible to note that the suggested model has compensated for most of the loss from the two datasets and reached the same level of accuracy as the centralised accuracy of chest X-ray data.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

COVID-CT at five clients rises from 0.7982 at ε = 2 to 0.9226 at ε = 10.

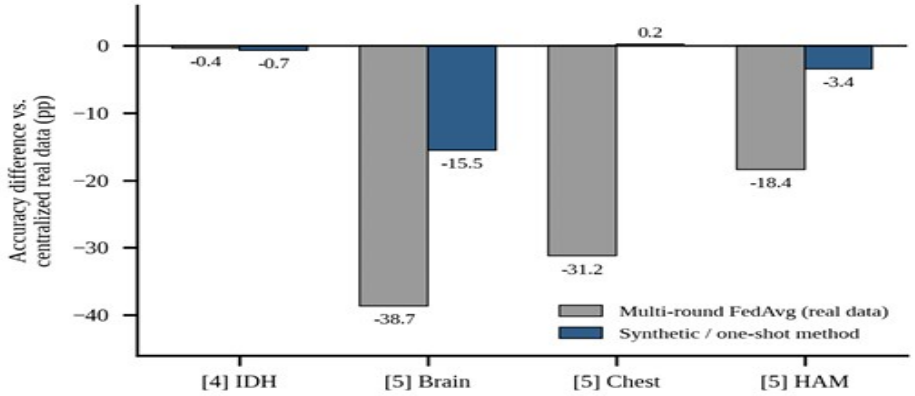


Fig. 2. Accuracy difference from centralized training on real data, in percentage points, computed from the values reported in [4] and [5] (higher is better).

C. What Type of Evidence for Privacy?

Only study [1] gives a formal assurance, which clearly indicates the trade-off involved: a smaller ε translates into a decreased quality of the generated images. Other studies either give empirical evidence or use argumentation. Study [3] holds that there is no membership inference and a lower risk of singling out after the federation, although the output of the federated CTGAN turns out to be more distinguishable from the real examples, so many of the gains in privacy may derive from the significance of reduced fidelity. Study [4] demonstrates that a membership inference resulting in random estimates (AUC 0.493) is possible using the image similarity attack method, which is quite a simple method. Study [5] relies on the argument of mutual information as well as on the result of the memorisation check and achieves an L2 distance threshold of 0.1, according to [10]; although this is interesting, it is not a possible guarantee. Neither of the studies [4] and [5] employs differential privacy; hence, they cannot provide the privacy-utility curve.

D. Evaluation Practice

The findings can only be trusted to a certain degree due to recurrent problems. Firstly, in [3], the ten synthetic datasets utilised in each t-test appear to be generated from the same pre-trained model, which means that sampling variation is being judged rather than the variation between different training processes, meaning that it can be overestimated. Secondly, in [4], the conclusion that synthetic and actual training yield similar results is made due to the absence of significance in McNemar tests; however, having found no difference does not imply having proved equivalence. In addition, it could serve as a weaker argument than having an established margin for equivalence. Thirdly, in [5], the multi-round FL baseline's performance is below the centralised top performance (for example, 36.95% compared to 75.64% for brain tumours), meaning that the results presented are somehow related to the benchmark difficulty for FedAvg. Lastly, only a few clients were involved (three in [4] and [5]), and most non-IID cases are connected to label skew.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

E. Handling Heterogeneity

There are different approaches used for the treatment of non-IID data in the literature. The study [1] discusses the classification of institutions by geographic proximity and investigates the performance of different label-skewed and Dirichlet partitions. In study [3], the researchers examine the performance of non-IID splits based on the balance and age of different partitions, concluding that the performance of all three types of splits is similar. In study [4], the authors demonstrate that the generator for each site is specific to that site. The FID value for the site's synthetic data relative to its real data is about 35 to 66, while the value when comparing the data with the other sites is usually above 200. In study [5], divergence is performed for a time of aggregation using dual-layer distillation. The authors illustrate this idea with a t-SNE plot indicating that the final features of the clients are merged after training. Heterogeneity is considered in the multi-site MRI approach of study [4] and the naturally federated tuberculosis data of study [5].

F. Threat Surface

Federation affects the information that leaves hospitals and, thus, that which attackers can exploit. The information summarised in Table III comes from the four key studies reviewed in this paper, with the term "exposure" representing the authors' interpretations of the content rather than their claims. A single-shot methodology minimises the time of exposure, and this is pointed out in [4] while not necessarily limiting the extent of exposure: the entire generator becomes visible at once so that one attack can expose more than one upgrade of the iterative FL. Iterative techniques provide observation opportunities through many snapshots of the model. Which risk is more significant is an empirical question since none of the studies reviewed carries out a comparison between the two approaches.

TABLE III. WHAT LEAVES THE SITE AND WHERE RISK SITS

Ref.	Leaves the site	Main exposure	Mitigation reported
[1]	Noised discriminator updates, each round	Gradient leakage over many rounds	Clipping, Gaussian noise, moment accountant
[3]	Model weights, each round	Weight updates; released synthetic rows	Empirical attack tests; no DP
[4]	Trained generator, once	Memorisation in the generator	SSIM attack; augmentation, dropout, clipping
[5]	Flow model and classifier, once	Memorisation in the flow model	Feature-level generation; distance check

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

V. OPEN GAPS

Lack of consistent benchmark. Every study uses its own datasets, splits and metrics, thus making it impossible to conclude anything about the rankings from the research [2].

Privacy-accuracy trade-offs for unique methods. Since the experiments in [4] and [5] are not based on DP, it is unclear how much accuracy would have been sacrificed if a proper guarantee had been applied.

Faithful heterogeneity. Unlike label skew, the differences between scanners, protocols and populations are hardly ever modelled, while [2] emphasises that issues of dealing with non-IID input data are one of the main problems.

Clinical validation. According to the survey [2], most of the existing work is based on simulations. The work [4] represents an exception, as it includes a part with real multi-institution MRI**.**

Inconsistent cost evaluation. Communication is described as round complexity in [1] and as training and inference time in [5], making it impossible to make comparisons**.**

Impartiality and prejudice. The study [2] points out that the use of artificial data generated from biased datasets could lead to the presence of bias in various applications; however, the studies conducted do not refer to the performance in each subgroup of the population.

Clarity and confidence. The reasons why a generator produces a certain output are significant for the implementation of this technique in clinics [2], but the methods analysed in the paper only produce general results**.**

Rewards and control. Publication [1] states that it is challenging for various organisations to agree on collaboration; thus, it suggests combining FL with blockchain technology, an idea that is still untested.

Computational work. Building a three-dimensional model requires four A100 graphics processing units according to the article [4], which are not available in many medical institutions; the designers in [5] aim to resolve this problem through the use of light models.

VI. A REPORTING CHECKLIST

To make future comparisons possible, studies of federated synthetic data should report the items in Table IV. The checklist is a proposal derived from the gaps above, not a validated standard.

TABLE IV. PROPOSED REPORTING CHECKLIST

Dimension	Report
Fidelity	A distribution metric (e.g., FID for images) and a detectability test, per site and pooled.
Utility	Downstream accuracy against centralised real data and standard FL, with a pre-specified equivalence margin and runs repeated across training seeds.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

Privacy	A formal (ε, δ) where DP is used, plus at least one calibrated attack and a memorisation check on any shared model.
Cost	Rounds, bytes transferred and wall-clock time for clients and server.
Setting	Client count, heterogeneity type and the weakest baseline tested.

Table V provides an application of the checklist for all four key studies. The entries show the author's own interpretation of the literature; Y means there is a reference, P means there is some mention in the article or supplementary material, and N means there is no mention at all. Not every research study cannot make it through all its requirements. Typically, the requirements may be missing particulars of equivalence testing, and the most concerning one for patients is the absence of a proper privacy guarantee in one-shot diffusion-based approaches.

TABLE V. CHECKLIST COVERAGE OF PRIMARY STUDIES

Item	**[1]**	**[3]**	**[4]**	**[5]**
Formal (ε, δ) guarantee	Y	N	N	N
Empirical attack	N	Y	P	P
Centralized baseline	Y	Y	Y	Y
Equivalence test	N	N	N	N
Cost reported	P	N	P	P

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

VII. A PROPOSED EXPERIMENTAL PROTOCOL

The comparison shown provides a basis for concrete research that no existing studies carry out. Train the same generator on one publicly available database of medical data using three approaches: centralised training, iterative federation with FedAvg protocol, and one-shot federation, in which the clients send their generator just once before receiving any data. Do each of the approaches with a differential privacy mechanism at different privacy budgets, for example, with ε = 2, 8, and no privacy protection. Use at least two types of data partitioning, such as label skew and Dirichlet sampling. The researchers must use some randomisation methods so that the variations come from training and not merely from random data selection. All required values must be reported in Table IV. The utility of the results must be evaluated using some test of equivalence in which the threshold for success should be decided before the experiment and at least two percentage points must be proven to have clinical backing.

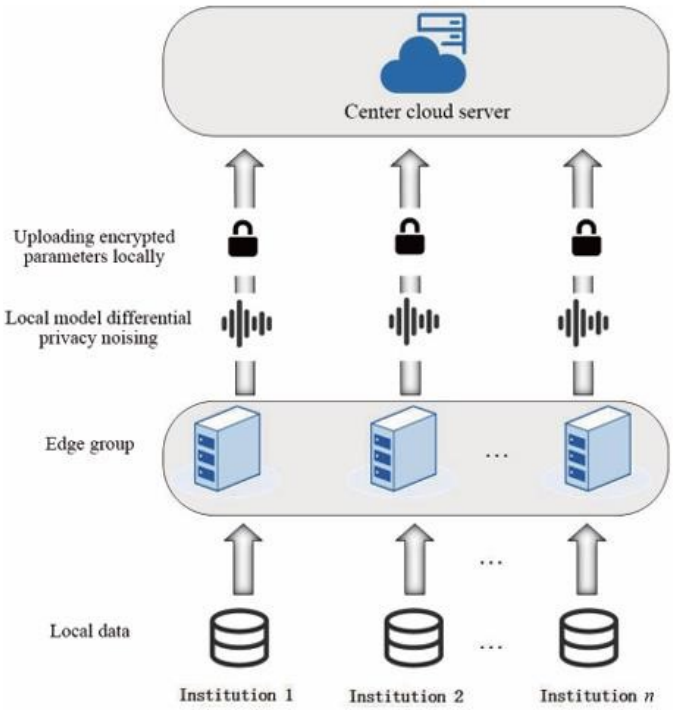


Fig. 3. Schematic framework of the local differential privacy for the selected edge group's local model.

VIII. LIMITATIONS OF THIS REVIEW

This research paper has some restrictions. It discusses five selected studies on the basis of appropriateness and availability, without a systematic search, while the report [2] serves as the best reference in the field of study. The data is provided in its original form, with a lack of replication and examination of additional notes and annexes, resulting in disregard for some points mentioned there. The measurements and datasets vary among the studies, hindering the ability to conduct a comparative analysis effectively. Finally, terms like "weak benchmark" or "little partition" are the authors' points of view that can be debated.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

IX. CONCLUSION

According to recent research, synthetic data can be effective for iterative federation, especially in certain scenarios. Studies reveal that when generating feature values (instead of images), concerns about information loss can be minimised. Notably, while some work relies on formal privacy guarantees, others make use of heuristics. As evaluations are based on small client numbers and might contain weak reference points, which in the latter can also be missing altogether. Thus, further progress can occur only in connection with the introduction of benchmarks, privacy-utility graphs for one-shot techniques, and particular reporting procedures amending the checklist proposed here. The next logical step appears to be a study involving the same generator trained under BYO and one-shot methods and/or considering both cases of the technology changing DP status.

REFERENCES

[1] S. Li, L. Hu, C. Sun, H. Li, and J. Hu, "Enhancing medical assistance through secure federated edge data augmentation with local differential privacy in non-IID scenarios," *Tsinghua Sci. Technol.*, vol. 31, no. 2, pp. 726–744, Apr. 2026, doi: 10.26599/TST.2024.9010179.

[2] H. Mahmood, Z. Alamgir, S. T. Javed, S. Karim, and M. Awais, "Federated generative models in medical imaging: Current advances, challenges, and future directions," *IEEE Access*, vol. 14, pp. 5197–5217, 2026, doi: 10.1109/ACCESS.2026.3650810.

[3] I. Isasa, M. Catalina, G. Epelde, N. Aginako, and A. Beristain, "Synthetic tabular data generation under horizontal federated learning environments in acute myeloid leukaemia: Case-based simulation study," *JMIR Med. Informat.*, vol. 13, Art. no. e74116, 2025, doi: 10.2196/74116.

[4] N. C. D. Truong *et al.*, "Categorical and phenotypic image synthetic learning as an alternative to federated learning," *Nature Commun.*, vol. 16, Art. no. 9384, 2025, doi: 10.1038/s41467-025-64385-z.

[5] Y. Ma, H. Zhang, Q. Yang, G. Luo, and Y. Zhu, "A new one-shotfederated learning framework for medical imaging classification with feature-guided rectified flow and knowledge distillation," in *Proc. Eur. Conf. Artif. Intell. (ECAI)*, 2025, pp.185–192, doi: 10.3233/FAIA250805.

[6] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas,"Communication-efficient learning of deep networks from decentralised data," in *Proc. 20th Int. Conf. Artif. Intell. Statist.*, 2017, pp. 1273–1282.

[7] M. Abadi *et al.*, "Deep learning with differential privacy," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2016, pp. 308–318.

[8] I. Goodfellow *et al.*, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.

[9] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer,"High-resolution image synthesis with latent diffusion models," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2022, pp. 10684–10695.

[10] N. Carlini *et al.*, "Extracting training data from diffusion models," in *Proc. 32nd USENIX Secur. Symp.*, 2023, pp. 5253–5270.

[11] X. Liu, C. Gong, and Q. Liu, "Flow straight and fast: Learning to generate and transfer data with rectified flow," arXiv:2209.03003, 2022.

[12] H. Zhu, R. Togo, T. Ogawa, and M. Haseyama, "Personalized federated learning for medical vision-language models via efficient fine-tuning and uncertainty-aware disentanglement," *J. Biomed. Informat.*, vol. 178, Art. no. 105014, 2026.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

[13] D. Babuc and T. F. Fortiş, "MedFusion-LM: Explainable large language model for transforming medical outcomes in federated learning with neural architecture search blueprints," in *Proc. ECML PKDD Workshop Responsible Healthcare Using Mach. Learn.*, Porto, Portugal, Sep. 2025.

[14] H. Zhu *et al.*, "Personalised longitudinal medical report generation via temporally-aware federated adaptation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2026.

[15] F. Abacha, S. G. Teo, L. C. Cordeiro, and M. A. Mustafa, "Synthetic data-aided federated learning using foundation models," in *Federated Learning in the Age of Foundation Models—FL 2024 Int. Workshops* (Lecture Notes in Computer Science, vol. 15501). Cham, Switzerland: Springer, 2025.

[16] A. Rehman *et al.*, "FedCSCD-GAN: A secure and collaborative framework for clinical cancer diagnosis via optimised federated learning and GAN," *Biomed. Signal Process. Control*, vol. 89, Art. no. 105893, 2024.

.