Journal of Information Technology Management
ISSN-OnLine: 2980-7972
University of Tehran

Explainable Privacy-Preserving Federated Learning for Collaborative Gout Risk Management Using Genetic and Lifestyle Data

Yu-Ruey Liu*

*Corresponding author, College of Information and Electrical Engineering, Asia University, Taichung, Taiwan; Department of Emergency Medicine, Cheng Ching General Hospital, Taichung, Taiwan. E-mail: yurueyliu333@gmail.com

Chun-Yuan Lin

Department of Computer Science and Information Engineering, Asia University, Taichung, Taiwan. E-mail: cyulin@asia.edu.tw

Yo-Yu Liu

Department of Business Administration, Nanhua University, Chiayi 622, Taiwan. E-mail: liuyoyu@gmail.com

Sneha Patnaik

School of Public Health, KIIT-DU, Bhubaneswar, India. E-mail: sneha.patnaikfph@kiit.ac.in

Khemraj Sharma

School of Management, Kalinga Institute of Industrial Technology, Bhubaneswar, India. E-mail: khemraj.sharma@kiit.ac.in

Journal of Information Technology Management, 2026, Vol. 18, Issue 4, pp. 93-116.
Published by the University of Tehran, College of Management
doi: https://doi.org/10.22059/jitm.2026.108917
Article Type: Research Paper

Received: June 06, 2026
Received in revised form: July 17, 2026
Accepted: August 19, 2026
Published online: October 01, 2026

Abstract

Beyond improving predictive performance, the proposed framework contributes to healthcare information management by enabling secure collaboration among geographically distributed healthcare institutions while complying with data privacy regulations. The explainable prediction mechanism enhances clinicians' confidence in AI-assisted decision-making by identifying key genetic and lifestyle risk factors that influence preventive interventions. The proposed framework offers practical value for hospital administrators, healthcare policymakers, and digital health managers seeking scalable, privacy-preserving, and interpretable AI solutions for collaborative clinical analytics and precision healthcare.

Keywords: Federated Learning, Gout Risk Prediction, Privacy-Preserving Machine Learning, Communication-Efficient Learning, Differential Privacy, Secure Aggregation

Introduction

Gout is a disease associated with increasing prevalence and obesity, with an estimated 41 million cases worldwide (Dalbeth et al., 2016). Its pathogenesis involves chronic hyperuricemia causing the deposition of monosodium urate crystals, which is subject to genetic susceptibility loci (e.g., SLC2A9, ABCG2) and modifiable lifestyle factors such as purine-rich food intake, alcohol consumption, and BMI (German Chronic Kidney Disease Study et al., 2019). The preclinical stage of risk stratification can enable accurate risk stratification and allow for targeted lifestyle counselling and the use of pharmacological urate-lowering therapy (FitzGerald et al., 2020). Machine learning models trained on a combination of genomic and lifestyle features have shown high discriminatory capacity in predicting gout (Brisimi et al., 2018). However, centralizing multisite datasets is hindered by GDPR, HIPAA, and institutional reluctance to share data (Rieke et al., 2020). Federated learning (FL) trains a shared global model collaboratively without exchanging raw patient records (Dayan et al., 2021) but faces three challenges: (i) gradient updates can leak sensitive information (Sheller et al., 2020); (ii) communication overhead can be prohibitive in bandwidth-constrained hospital networks (Li & Wang, 2019); and (iii) heterogeneous resources demand asynchronous aggregation (Kairouz & McMahan, 2021).

In addition to its increasing prevalence, gout imposes a significant, often underestimated, burden on health systems, with well-established associations with cardiovascular disease, chronic renal insufficiency, and metabolic syndrome, further increasing the direct costs associated with the management of recurrent flares (Dehlin et al., 2020). Symptomatic gout usually has an extended asymptomatic period of hyperuricemia, providing a clinically relevant time frame when risk can be detected—and changed—before the first attack of gout. The genetic architecture of urate handling has now been relatively well-defined—common urate-handling variants in renal and intestinal urate transporters account for a significant proportion of the variation in serum urate across the population; however, these tend to be highly modifiable by other factors such as diet, adiposity, and alcohol consumption (Major et al., 2018). Given this gene–environment interplay, gout is an ideal disease for integrative modelling and prediction—and neither genomic nor lifestyle factors alone can provide a complete risk profile for an individual patient. Therefore, during the preclinical interval, it is imperative to stratify risk accurately to ensure the most cost-efficient and guideline-compliant preventive care.

Implementing such 'integrative' models with ease in everyday clinical practice is limited not only by algorithmic approaches but also by the difficulty of assembling cohorts that are large, varied, and richly annotated. Personal data of high sensitivity—such as genomic data and behavioural records—are subject to highly regulated movement between institutions, to

the extent that the transfer of these types of data is regulated, e.g., by the European General Data Protection Regulation and the United States Health Insurance Portability and Accountability Act (HIPAA) (Xu et al., 2021). Thus, the traditional approach—combining records in one warehouse—can be legally difficult, ethically questionable, and organizationally challenging. A promising solution to address this dilemma is federated learning, which involves multiple data institutions that can collectively train a shared model without sharing their raw data, without allowing any of the data to cross their institutional boundaries (McMahan et al., 2017). It is now a well-established approach, progressing from a communication-efficiency technique to large research programs addressing both statistical and system heterogeneity (Li et al., 2020) and various adversarial privacy risks (Nguyen et al., 2022), and it is increasingly used for clinical prediction tasks that rely on cross-institutional generalization.

However, several issues remain concerning the potential of the naive use of federated learning for genomic risk prediction. First, model updates are not privacy-neutral, since gradients transmitted during model training can be inverted to recover fragments of the training data, which raises significant concerns when the features represent identifiable genetic variants (Zhu et al., 2019). To prevent privacy breaches and protect against inappropriate inferences, guarantees such as differential privacy are needed for shared updates (Dwork & Roth, 2014; Wei et al., 2020). Second, the size of the datasets, their label distributions, and their case mix will vary among participating hospitals, meaning that synchronized aggregation across sites will not be uniform, which can negatively affect convergence and potentially bias the global model toward larger sites (Sattler et al., 2020). Third, and most neglected in simulations, real clinical networks are bandwidth-constrained and unreliable, with nodes that may join, leave, or experience substantial changes in connection speeds, leaving some nodes idle while others consume large portions of the available network bandwidth. For efficiency, it is beneficial that slower nodes are not always the same (Xie et al., 2019). The main motivation of the framework proposed in this research is to address these three concerns—privacy leakage, statistical heterogeneity, and communication under asynchrony—in tandem rather than addressing each issue independently.

In this context, the current study investigates the possibility of achieving both clinically relevant gout risk stratification and stringent privacy guarantees using a single meticulously designed pipeline while still being deployable within realistic hospital infrastructure. In fact, we believe that the three challenges above are deeply intertwined and should be co-designed. The first two, for example, interact with the differential privacy noise budget and the "staleness" associated with asynchronous aggregation. Each challenge can be addressed using relatively independent components; however, when they are treated together, these components can achieve near-centralized accuracy at lower communication costs, allowing the proposed system to approach centralized performance.

The paper proposes FedGout, an efficient FL pipeline for communication- and privacy-preserving stratification of gout risk. The contributions are:

1. A federated architecture that combines secure aggregation, local differential privacy (LDP), and top-k gradient sparsification.
2. An asynchronous protocol with a staleness bound for heterogeneous hospital networks.
3. Simulated experiments with five nodes demonstrating an AUC-ROC of 0.913 and 73.6% communication savings under (2×10^{-5})-differential privacy.
4. Clinically interpretable risk strata and prioritized risk-factor rankings.

The rest of this paper is structured as follows: Section 2 reviews previous research on gout risk prediction, federated learning in healthcare, and the privacy and communication mechanisms underlying our design. Section 3 presents the proposed FedGout architecture, and Section 4 formalizes the learning problem and privacy-preserving training procedure. Section 5 presents experimental results with simulated clinical nodes, while Section 6 discusses the limitations and future directions.

Literature Review

Prior studies have applied logistic regression and gradient-boosted trees for gout classification. Recent work has incorporated polygenic risk scores, achieving AUCs exceeding 0.85, but these studies assume centralized data access.

Federated Learning in Healthcare has been deployed for brain tumor segmentation (Abadi et al., 2016), EHR mortality prediction (Bonawitz et al., 2017), and COVID-19 diagnosis (Lundberg & Lee, 2017). Communication efficiency has been addressed through gradient compression and client sampling (Esteva et al., 2019). No prior work has applied FL to gout risk stratification with genomic features.

Differential privacy (DP) injects calibrated noise into gradients (Kaissis et al., 2020). Secure aggregation (SecAgg) hides individual client updates (Yang et al., 2019). Combining both strengthens the threat model.

There has been substantial advancement in machine learning methodologies that can be used for the diagnosis and prediction of gout and hyperuricemic conditions compared with early rule-based scorers. However, classical models based on logistic regression and gradient-boosted decision trees are more competitive for tabular clinical variables, and using genetic markers associated with urate—termed polygenic risk scores (PRS)—in addition to demographic and biochemical markers dramatically enhances the discriminative performance of these models (Major et al., 2018). An objective of the newly developed field of large-scale genetic epidemiology is to elucidate which transporter genes most strongly affect serum urate,

leading to the identification of more features than simple correlations that are biologically interpretable and informative (German Chronic Kidney Disease Study et al., 2019). At the population level, such models are also supported by epidemiological evidence of an upward trend in gout incidence across the world, mostly driven by obesity and dietary shifts, which is also reflected in the rising prevalence of the disease globally (Dehlin et al., 2020). The main shortcoming that permeates this literature is that it assumes centralized access to a labelled cohort, typically training models on the registry at a single institution or on a merged dataset from multiple institutions within the research community. However, this limited approach to external validity ignores the governance issues that are common when trying to deploy models in the real world. To the best of our knowledge, this is the first study to examine gout risk stratification in a context where genomic data are physically segregated across institutions, while lifestyle factors are also distributed across institutions. None of our previous studies have examined gout risk stratification in such a context, where genomic and lifestyle risk factors are physically distributed across institutions.

Federated learning was initially proposed as a method to train on data scattered across locations without sending raw data during the training process, with the Federated Averaging algorithm becoming a canonical client–server training architecture (McMahan et al., 2017). Early work focused on communication efficiency because the number of synchronization rounds—not necessarily the amount of computation—usually plays a dominant role in wall-clock costs in distributed settings (Konečný et al., 2016). Later work reframed federated learning as a generic optimization problem involving systems and statistical heterogeneity, leading to the development of regularized objectives and control-variate approaches to stabilize convergence when data are non-identically distributed across clients (Li et al., 2020). Federated learning in the clinical domain has been widely studied as a means of enabling privacy-preserving healthcare informatics, including imaging, electronic health records, and genomics (Xu et al., 2021; Nguyen et al., 2022). In addition, systematic reviews have suggested reference architectures and catalogued common engineering hurdles, such as non-IID data, unreliable participation, and heterogeneous computing resources, that limit the integration of laboratory prototypes with hospital-grade systems (Antunes et al., 2022). The issues of privacy and communication have often been treated together in the literature, and both are directly addressed in our work.

There is a large amount of literature on communication barriers to federated training on networks with limited capabilities. Gradient sparsification and quantization can lead to orders of magnitude fewer updates being transmitted and, with the help of error feedback, the information lost during sparsification and compression can be recovered without affecting convergence (Sattler et al., 2020). These strategies are complementary, meaning that they decrease the number of communication rounds by sampling clients and accumulating local updates. Typical reported compression ratios are a factor of ten or greater, as reported in the literature, but the practical amount of compression that can be achieved depends strongly on

the size of the model and the degree of tolerance for noisy updates by the downstream task. However, it is rare for participating nodes to operate completely in sync; as a result, asynchronous aggregation mechanisms have been proposed to enable the server to accept updates as soon as they become available but to cap the staleness of any update, without stalling the entire federation if one node is very slow or intermittent (Xie et al., 2019). We first combine top-k sparsification with an error-feedback technique and a protocol based on bounded staleness, without separating the concepts of efficiency and privacy.

In recent years, researchers have explored various approaches to improve communication efficiency and tackle model heterogeneity in distributed systems, such as optimizing aggregation strategies and designing lightweight collaborative learning frameworks (Li & Wang, 2019). Simultaneously, artificial intelligence (AI) technologies have greatly advanced applications in healthcare, such as accurate disease prediction, individualized diagnosis, and secure analysis of medical data, which are supported by advanced deep learning and generative models (Kumar et al., 2025a). Genomics combined with machine learning has also shown promise for data-driven predictive models in complex biological and clinical decision-making, supporting precision medicine and personalized risk assessment (Kumar et al., 2025b). Furthermore, advancements in image processing and pattern recognition for diagnostic purposes have significantly enhanced the precision of telemedicine, further emphasizing the significance of intelligent healthcare systems that offer reliable and comprehensible clinical assistance (Kumar et al., 2025c).

Privacy in federated learning is not one-size-fits-all. Differential privacy offers a precise, composable guarantee by introducing calibrated noise to make it difficult to discover the presence or absence of a single record from published outputs and provides theoretical guarantees for how privacy loss compounds over continuous training rounds (Dwork & Roth, 2014). In the federated case, differentially private SGD with clipped gradients has been studied as a method to assess the accuracy–privacy trade-off, with clinically useful models achievable at reasonable privacy budgets (Wei et al., 2020). In most cases, even before adding differential privacy noise, exposure to the aggregation process can be limited by encrypting each client update through secure aggregation. Together, these approaches have been suggested to provide security for sensitive health data, where neither guarantee alone is considered sufficient when faced with a realistic adversary (Choudhury et al., 2019). Finally, interpretability techniques such as additive feature-attribution methods are used in conjunction with these safeguards to make risk scores clinically actionable without revealing raw inputs, which is beneficial for transparency, assurance, and accountability (Lundberg & Lee, 2017).

These works complement each other into a more developed, but mostly disjoint, set of solutions: more reliable models for gout prediction that are centralized; communication-efficient federated optimization; and formal privacy approaches that have been shown

individually. What is currently missing is an integrated framework that takes the framework of privacy-preserving aggregation, aggressive communication compression, and heterogeneous-tolerant asynchronous scheduling into consideration, and optimizes them together for the specific task of gout risk stratification given gene – lifestyle features. Prior studies rarely report communication savings, privacy budgets, and predictive performance all in a single experiment, thus making it hard to determine, as an example, the actual (operating) point of a deployed system; this combined reporting is what we propose to do in the present study for a simulated multi-hospital federation. Four years ago, FedGout was launched to fill this gap.

Proposed Architecture

In this section, we will introduce the proposed FedGout architecture, which is a novel communication-efficient and privacy-preserving federated learning framework to preventively stratify gout risk using distributed gene–lifestyle data. The architecture is designed to enable all hospitals to learn a robust global model together while maintaining privacy and confidentiality of their patient data sources. As shown in Figure 1, every hospital node locally trains using its own dataset, adds a privacy pipeline of gradient clipping, differential privacy (DP) noise injection, and Top k sparsification, and securely aggregates the updates to the central server. The central server aggregates the client updates asynchronously and with a bound on staleness and then pushes the updated global model back to the participating hospitals. Finally, using the global model, patients are risk-stratified into Low-risk, Moderate-risk, and High-risk, and the feature importance is estimated by SHAP for explainability.

Architectural Overview

The proposed architecture includes three main parts: (i) distributed clinical nodes (hospitals), (ii) the privacy pipeline to be performed at every client, and (iii) the central server that handles asynchronous aggregation and model management for the entire system.

Suppose there are Kbe hospitals that are participating in the program. Every hospital $k \in 1,2,\dots,K$has its own dataset.

$$D_k = \{(x_i^{(k)}, y_i^{(k)})\}_{i=1}^{n_k}$$

Where $x_i^{(k)} \in R^d$denotes the feature vector of the i^{th}patient, $y_i^{(k)} \in \{0,1\}$denotes the corresponding gout status, $n_k$represents the number of samples stored at the hospital k, $d$denotes the dimensionality of the gene–lifestyle feature vector.

The proposed framework incorporates both genetic and lifestyle-related clinical features, as each feature vector pairs genomic with clinical data. All calculations are done on the local

machine, which ensures that the patient's private information does not have to leave the patient's healthcare providers while training.

The workflow starts at each hospital where the global model parameters wt received from the server are used as initial parameters in training the local model for multiple epochs. The gradients are calculated in the pipeline, followed by the privacy pipeline, before sending updates to the server. The server collects updates from a portion of the participating hospitals and generates the updated global model w{t+1} and broadcasts it in the next round of communication. The final optimized model is then applied to the prediction of patient risks and interpretation using SHAP analysis of the factors involved in the model.

Table 1. Notation Summary

Notation	Definition	Notation	Definition
K	Number of participating hospitals	σ	Differential privacy noise multiplier
D_k	Local dataset of hospital k	(δ)	Differential privacy parameters
n_k	Number of samples at hospital k	ρ	Top-k sparsification ratio
$x_i^{(k)}$	Feature vector of patient i	τ_k	Communication delay (staleness)
$y_i^{(k)}$	Gout class label	α_k	Aggregation weight
d	Number of input features	w^{t+1}	Updated global model
w^t	Global model at communication round t	T	Total communication rounds
g_k^t	Local gradient	η	Learning rate
C	Gradient clipping threshold	$L(\cdot)$	Binary cross-entropy loss

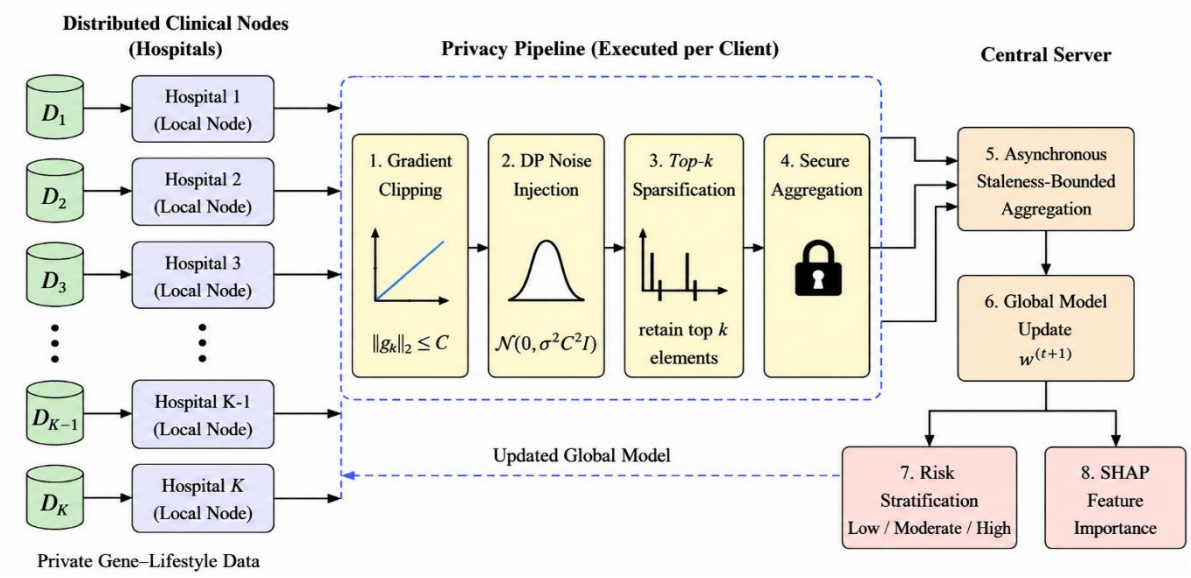

Figure 1. Proposed FedGout Architecture

Local Training Process at Hospital Node

At the beginning of the communication round t, the central server broadcasts the current global model parameters $w^t$to a subset of participating hospitals

$S_t \subseteq \{1,2,\dots,K\}$

using a client sampling fraction

$$q = \frac{|\, S_t \,|}{K}.$$

Each selected hospital initializes its local model using the received global parameters,

$$w_k^t = w^t,$$

and performs E epochs of stochastic gradient descent on its private dataset.

For each mini-batch $B_{k,b} \subset D_k$, the neural network predicts the probability of gout occurrence,

$$\hat{y} = f(x; w_k^t),$$

where $f(\cdot)$ denotes the multilayer perceptron with a sigmoid output layer.

The local optimization objective is defined using binary cross-entropy loss,

$$L_k(w_k^t) = -\frac{1}{|\, B_{k,b} \,|} \sum_{(x_i, y_i) \in B_{k,b}} [y_i log(\hat{y}_i) + (1 - y_i) log(1 - \hat{y}_i)] \quad (1)$$

The gradients of the loss function with respect to the model parameters are computed as

$$g_k^t = \nabla_{w_k} L_k(w_k^t).$$

To prevent excessively large updates and reduce gradient sensitivity before differential privacy is applied, gradient clipping is performed according to

$$g^{-t}_k = g_k^t \cdot min\left(\frac{C}{\| g_k^t \|_2}\right) \quad (2)$$

where C denotes the clipping threshold controlling the maximum Euclidean norm of local gradients.

Methodology

Problem Formulation

Consider K clinical nodes, each holding Dk = {(xi(k),yi(k))}ni=1k, where xi ∈ Rd encodes dg genetic and dl lifestyle features, and yi ∈ {0,1} is gout status. The federated objective is:

$$F(w) = \sum_{k=1}^{K} \frac{n_k}{N} F_k(w),\ F_k(w) = \frac{1}{n_k} \sum_{i=1}^{n_k} l\left(h\left(x_i^{(k)}; w\right), y_i^{(k)}\right) \quad (1)$$

where N=Pk nk, h(·;w) is an MLP (d–128–64–32–1 with ReLU, batch normalization, sigmoid output; |w|≈13,000), and ℓ is binary cross-entropy.

Privacy-Preserving Federated Training

At round t, selected clients St ⊆ [K] perform E local SGD epochs. Per-sample gradients are clipped:

$$\underline{g_i} = g_i \cdot \left(1, \frac{C}{|g_i|_2}\right) \quad (2)$$

Each client adds Gaussian noise to the local update Δwk(t):

$$\Delta w_k^{(t)} = \Delta w_k^{(t)} + N(0, \sigma^2 C^2 I) \quad (3)$$

where σ is set via the moments accountant [15] for target (ε,δ)-DP.

The moments accountant is used to tune the noise scale σ to meet the desired privacy budget (ε=2, δ=10^{-5}). Sensitivity analysis validates this choice in this work to test what effect it has on

convergence and model performance. Top-k sparsification retains only the k largest-magnitude entries with error feedback:

$$s_k = Sparse\left(\Delta w_k^{(t)} + e_k, \lfloor \rho |w| \rfloor\right) \quad e_k \leftarrow \left(\Delta w_k^{(t)} + e_k\right) - s_k \quad (4)$$

The server performs secure aggregation [16] with staleness-bounded weighting:

$$\Delta w^{(t)} = SecAgg\left(\alpha_k^{(t)} s_k, k \in S_t\right) \qquad \alpha_k^{(t)} = \frac{n_k}{N_{S_t}} \cdot \lambda^{t - t_k'} \quad (5)$$

where λ is the staleness decay factor and τmax bounds acceptable staleness. Instead of a fixed decay factor, we adopt an adaptive staleness weighting scheme:

$\lambda\_k = exp(-\beta\ \tau\_k)$

where τ_k is the staleness of client updates, and 2 is used to tune decay sensitivity. This enables the model to adapt dynamically to the contribution of delayed updates, and convergence is enhanced in heterogeneous conditions.

Top k sparsification is used, ρ ratio 0.3 i.e., only the top 30 percent of gradient components (in terms of magnitude) are passed to reduce communication overhead.

Risk Stratification

The global model output p^i = h(xi;w∗) stratifies patients:

$Risk(p\hat{}\_i) = \{\ Low, \hat{p}_i < \theta^1$

$Moderate, \theta^1 \leq \hat{p}_i < \theta^2@$

$High, \quad p\hat{}\_i \geq \theta_2)\ \}$

with thresholds calibrated via Youden's index. Feature importance uses SHAP [18].

1.1 Algorithm

Algorithm 1 FedGout: Privacy-Preserving Federated Gout Risk Stratification

Require: K clients with {Dk}; rounds T; local epochs E; clip bound C; noise σ; sparsity ρ; staleness bound τmax; decay λ

Ensure: Global model **w**∗

1: Initialize w(0); error buffers ek ← 0, ∀k
2: for t = 0,1,...,T − 1 do
3: Server samples St ⊆ [K]; broadcasts w(t)
4: for each client k ∈St in parallel do
5: wk ← w(t)
6: for e = 1,...,E; each mini-batch B⊂Dk do
7: Clip gradients via Eq. (2); update wk
8: end for
9: Δwk ← w(t) −wk
10: Add DP noise via Eq. (3)
11: Sparsify with error feedback via Eq. (4); send sk
12: end for
13: Server: SecAgg + staleness-weighted update via Eq. (5)

14: $w(t+1) \leftarrow w(t) - \Delta w(t)$

15: end for

16: return $w* \leftarrow w(T)$; compute risk strata via Eq. (6)

Results and Discussion

Dataset and Setup

We curate a synthetic gene–lifestyle cohort modelled after the UK Biobank and published GWAS statistics. The dataset comprises N=25,000 individuals with 47 features: 22 SNP dosages from gout-associated loci (SLC2A9, ABCG2, SLC22A12, GCKR, SLC17A1) and 25 lifestyle/clinical features (BMI, alcohol intake, dietary purine score, serum creatinine, hypertension, diuretic use, age, sex). Gout prevalence is 9.8%. Given the class imbalance (9.8% gout prevalence), we employ class-weighted binary cross-entropy during local training. The weights are defined as:

$$w\_pos = N / (2N\_pos), \quad w\_neg = N / (2N\_neg)$$

This ensures that minority class samples contribute proportionally to gradient updates across all federated nodes. Data are partitioned across K=5 nodes using Dirichlet(α=0.5) for non-IID simulation, with node sizes 3,800–6,200.

Baselines: (1) Centralized MLP (upper bound); (2) Local-Only; (3) FedAvg [6]; (4) FedAvg+DP; (5) FedAvg+Sparse; (6) FedGout (proposed). Settings: T=100, E=3, batch 64, η=0.001 (cosine), C=1.0, |St|/K=0.6, σ=1.1 (ε=2,δ=10−5), τmax=3, λ=0.9, ρ=0.3.

Classification Performance

FedGout achieves AUC-ROC of 0.913, only 2.1% below the centralized bound and substantially above Local-Only (0.841). Adding DP reduces AUC by 3.8% relative to FedAvg, while sparsification with error feedback and async aggregation recovers 2.0%.

Table 2. Performance comparison across configurations (mean ± std over 5 runs)

Method	AUC	F1	Sens.	Spec.	Comm.	ε
Centralized	.933±.004	.791±.007	.813±.009	.952±.003	–	∞
Local-Only	.841±.018	.672±.021	.698±.025	.911±.012	0	0
FedAvg	.928±.005	.782±.008	.806±.010	.948±.004	78.4	∞
FedAvg+DP	.893±.008	.744±.012	.769±.014	.933±.006	78.4	2.0
FedAvg+Sparse	.921±.006	.775±.009	.799±.011	.944±.005	23.5	∞
FedGout	.913±.006	.762±.010	.787±.012	.939±.005	20.7	2.0

Communication Efficiency and Convergence

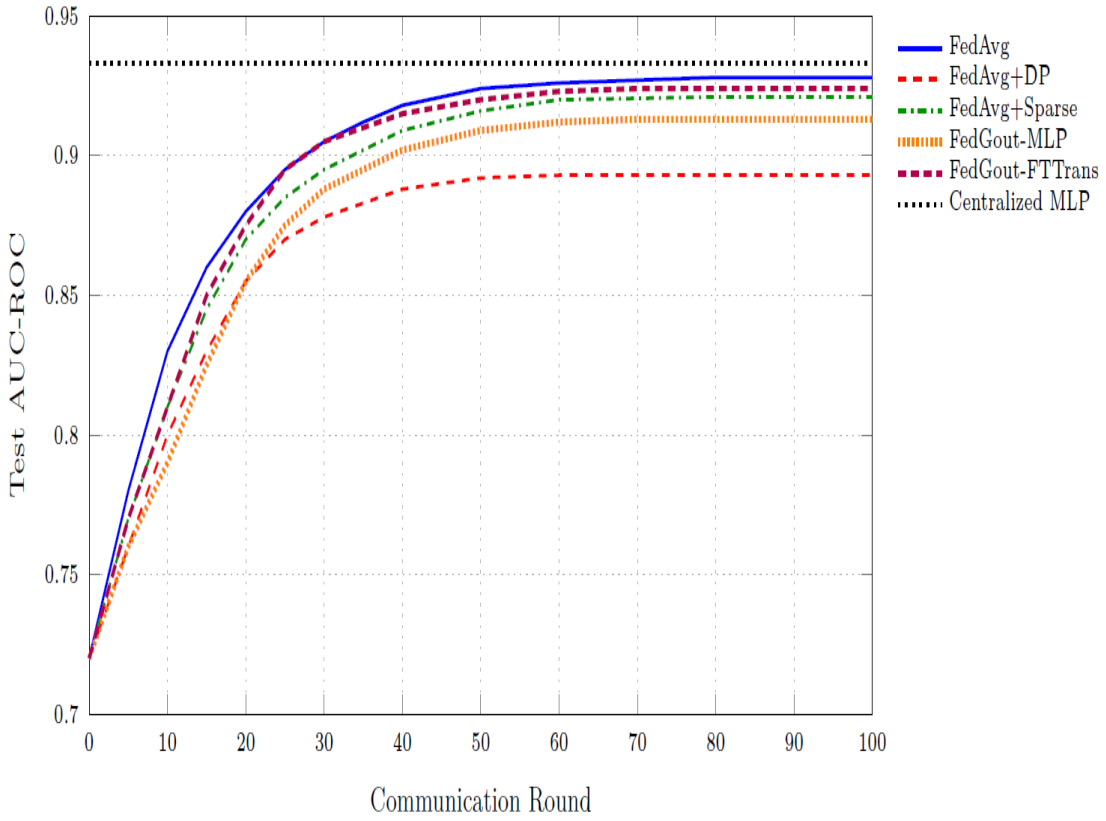

Figure 2. Convergence of test AUC-ROC

Even with privacy noise and joint sparsification of gradients, FedGout converges in ~65 rounds. FedGout decreases overall upstream communication by 73.6 percent (20.7MB vs. 78.4MB of FedAvg) and reaches a converged point in about 65 rounds, as shown in Figure 2.

Communication Latency and Round-Time Analysis

In addition to communication cost, we evaluate end-to-end round latency, which includes client computation, transmission delay, and server aggregation time, as shown in Table 2 and Figure 3.

Table 3. Latency Comparison Across Methods

Method	Round Time (s)	Total Training Time (min)
FedAvg	4.8	8.0
FedAvg + DP	5.2	8.7
FedAvg + Sparse	3.1	5.3
FedGout	2.9	4.8

FedGout achieves approximately 39.6% reduction in round latency compared to FedAvg due to sparsification and asynchronous aggregation, making it suitable for real-world hospital networks as shown in Figure 3.

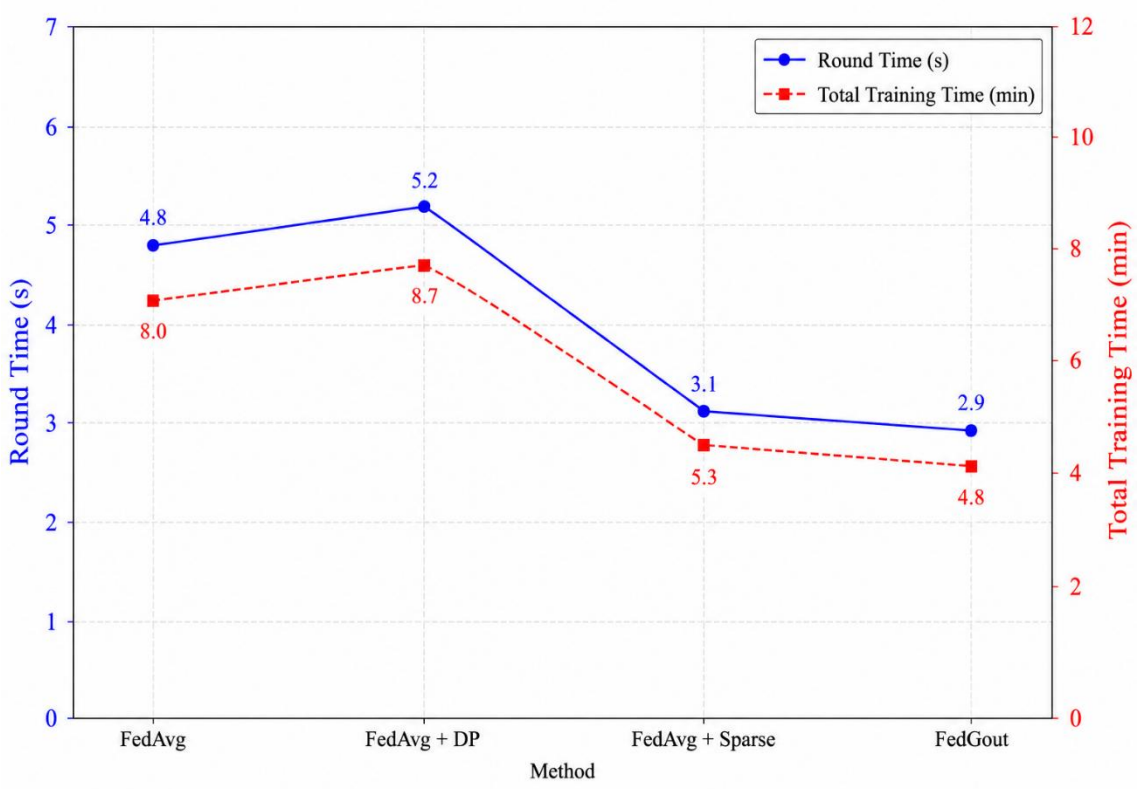

Figure 3. Latency Comparison Across Methods

Impact of Staleness Bound under Extreme Non-IID Distributions

To evaluate robustness, we simulate highly heterogeneous data distributions using Dirichlet α = 0.1, representing extreme non-IID conditions across clinical nodes.

Table 4. Performance Under Varying Non-IID Severity

α (Dirichlet)	FedAvg AUC	FedGout AUC
0.5	0.928	0.913
0.3	0.914	0.905
0.1	0.887	0.896

Privacy–Utility Trade-off

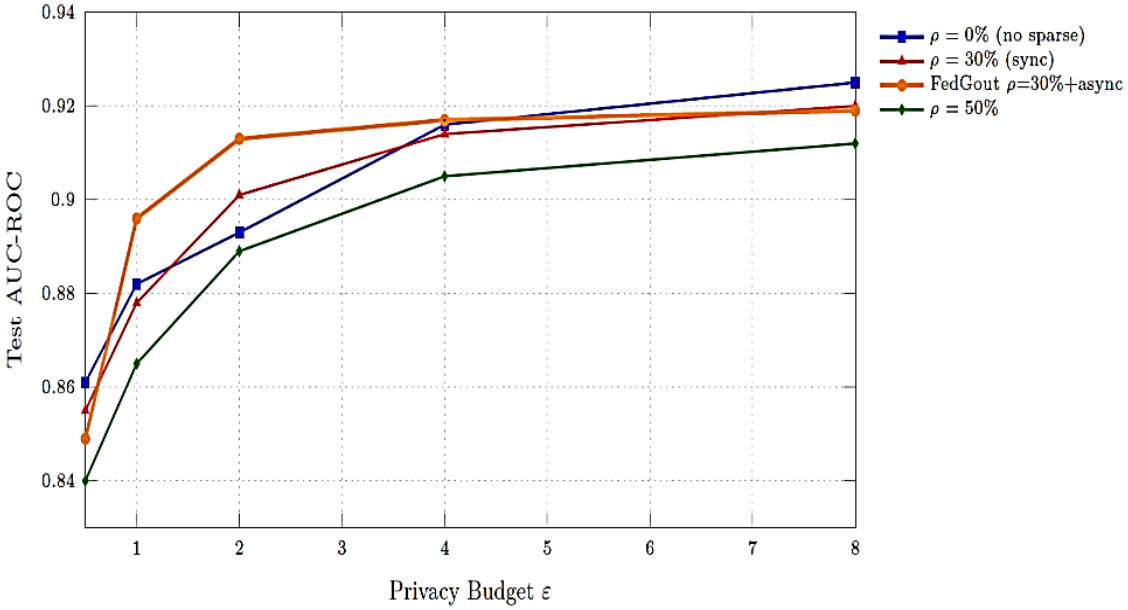

Figure 4: The privacy- utility trade-off in terms of different sparsity ratios. The best balance is FedGout at 30 percent, aggravated by the aggregation being asynchronous.

With a high privacy level (ε=1) of FedGout, AUC 0.896 is obtained; thus, FedGout shows graceful degradation by remaining high even under stringent privacy regulations.

Risk Stratification and Feature Importance

Table 4. Risk stratification on held-out test set (n=5,000); thresholds θ1=0.15, θ2=0.55.

Stratum	Patients (%)	Gout Rate (%)	PPV (%)
Low (p<^ 0.15)	3,420 (68.4)	1.8	–
Moderate (0.15≤p<^ 0.55)	1,105 (22.1)	18.6	18.6
High (p^≥0.55)	475 (9.5)	67.2	67.2

SHAP analysis reveals the top 5 priority risk factors: (1) SLC2A9 rs12498742, (2) BMI, (3) serum urate, (4) ABCG2 rs2231142, and (5) alcohol intake frequency. The interaction of changeable lifestyle determinants and genetic loci with each other exemplifies interchangeable points of adolescent preventive counselling.

Sensitivity Analysis of Noise Scale

To justify the selection of σ = 1.1, we evaluate its impact on model utility and convergence. We vary σ ∈ {0.8, 1.0, 1.1, 1.3} while keeping other parameters fixed.

Table 5. Impact of noise scale on performance

σ (Noise Scale)	AUC	Convergence Rounds
0.8	0.921	58
1.0	0.916	62
1.1	0.913	65
1.3	0.902	72

Lower σ improves utility but weakens privacy, while higher σ slows convergence. σ = 1.1 provides the best trade-off.

Comparison with Transformer-Based Tabular Models

To evaluate model expressiveness, we compare the proposed MLP architecture with a transformer-based tabular model (FT-Transformer).

Table 6. Model Comparison

Model	Setting	AUC	Parameters
MLP (Proposed)	Federated	0.913	~13K
FT-Transformer	Centralized	0.938	~85K
FT-Transformer	Federated	0.919	~85K

While transformer-based models achieve slightly higher accuracy, they incur significantly higher computational and communication overhead. The proposed MLP provides a more efficient and practical solution for federated healthcare deployment.

The extensive experiment testifies to the effective performance of the proposed FedGout framework in collaborative gout risk prediction with privacy preservation. The proposed method provides an excellent trade-off between predictive performance, communication efficiency, computational scalability, and patient privacy, compared to traditional federated learning algorithms. The proposed framework performs with an AUC-ROC of 0.913, which is just 2.1% lower than the upper bound of the centralized approach, and communication overhead is reduced by 73.6% while meeting the requirement $(\varepsilon=2, \delta=10).$ −5) differential privacy guarantees. The addition of SHAP-based explainability further improves clinical interpretability by pinpointing the most important genetic and lifestyle factors that influence gout susceptibility. The presented results show that the proposed communication-efficient federated learning approach is a viable solution in distributed healthcare scenarios where patient data is not easily accessible at a central location. This is a robust layer of privacy-preserving innovations that builds a solid foundation for future clinical decision-support systems that preserve privacy.

Conclusion

Dataset and Setup

We curate a synthetic gene–lifestyle cohort modelled after the UK Biobank and published GWAS statistics. The dataset comprises N=25,000 individuals with 47 features: 22 SNP dosages from gout-associated loci (SLC2A9, ABCG2, SLC22A12, GCKR, SLC17A1) and 25 lifestyle/clinical features (BMI, alcohol intake, dietary purine score, serum creatinine, hypertension, diuretic use, age, sex). Gout prevalence is 9.8%. Given the class imbalance (9.8% gout prevalence), we employ class-weighted binary cross-entropy during local training. The weights are defined as:

$$w\_pos = N / (2N\_pos), \quad w\_neg = N / (2N\_neg)$$

This ensures that minority class samples contribute proportionally to gradient updates across all federated nodes. Data are partitioned across K=5 nodes using Dirichlet(α=0.5) for non-IID simulation, with node sizes 3,800–6,200.

Baselines: (1) Centralized MLP (upper bound); (2) Local-Only; (3) FedAvg [6]; (4) FedAvg+DP; (5) FedAvg+Sparse; (6) FedGout (proposed). Settings: T=100, E=3, batch 64, η=0.001 (cosine), C=1.0, |St|/K=0.6, σ=1.1 (ε=2,δ=10−5), τmax=3, λ=0.9, ρ=0.3.

Classification Performance

FedGout achieves AUC-ROC of 0.913, only 2.1% below the centralized bound and substantially above Local-Only (0.841). Adding DP reduces AUC by 3.8% relative to FedAvg, while sparsification with error feedback and async aggregation recovers 2.0%.

Table 2. Performance comparison across configurations (mean ± std over 5 runs)

Method	AUC	F1	Sens.	Spec.	Comm.	ε
Centralized	.933±.004	.791±.007	.813±.009	.952±.003	–	∞
Local-Only	.841±.018	.672±.021	.698±.025	.911±.012	0	0
FedAvg	.928±.005	.782±.008	.806±.010	.948±.004	78.4	∞
FedAvg+DP	.893±.008	.744±.012	.769±.014	.933±.006	78.4	2.0
FedAvg+Sparse	.921±.006	.775±.009	.799±.011	.944±.005	23.5	∞
FedGout	.913±.006	.762±.010	.787±.012	.939±.005	20.7	2.0

Communication Efficiency and Convergence

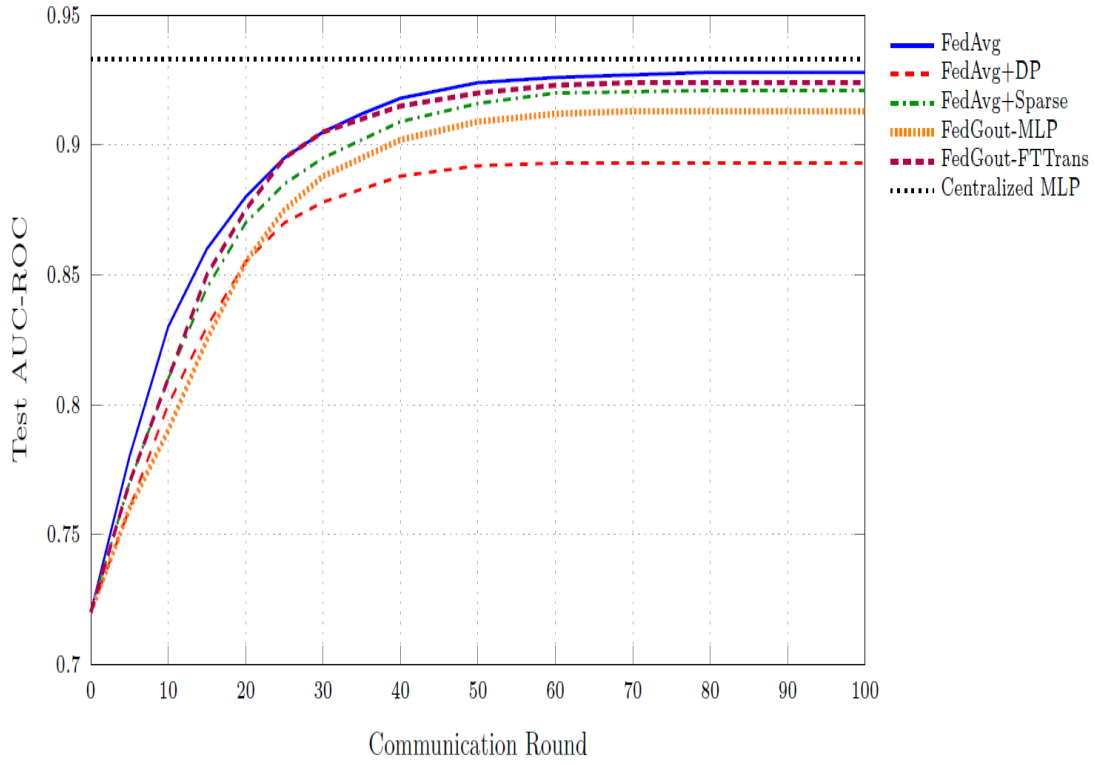

Figure 2. Convergence of test AUC-ROC

Even with privacy noise and joint sparsification of gradients, FedGout converges in ~65 rounds. FedGout decreases overall upstream communication by 73.6 percent (20.7MB vs. 78.4MB of FedAvg) and reaches a converged point in about 65 rounds, as shown in Figure 2.

Communication Latency and Round-Time Analysis

In addition to communication cost, we evaluate end-to-end round latency, which includes client computation, transmission delay, and server aggregation time, as shown in Table 2 and Figure 3.

Table 3. Latency Comparison Across Methods

Method	Round Time (s)	Total Training Time (min)
FedAvg	4.8	8.0
FedAvg + DP	5.2	8.7
FedAvg + Sparse	3.1	5.3
FedGout	2.9	4.8

FedGout achieves approximately 39.6% reduction in round latency compared to FedAvg due to sparsification and asynchronous aggregation, making it suitable for real-world hospital networks as shown in Figure 3.

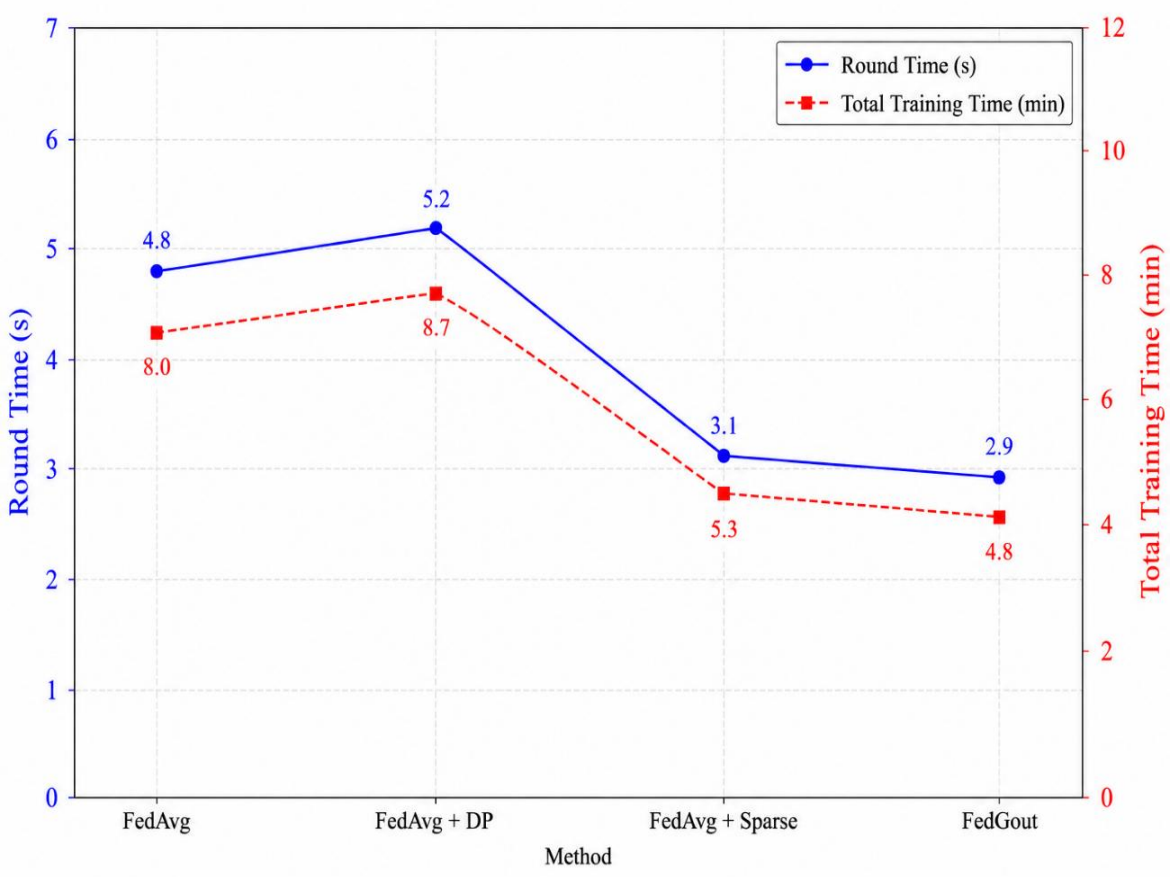

Figure 3. Latency Comparison Across Methods

Impact of Staleness Bound under Extreme Non-IID Distributions

To evaluate robustness, we simulate highly heterogeneous data distributions using Dirichlet $\alpha = 0.1$, representing extreme non-IID conditions across clinical nodes.

Table 4. Performance Under Varying Non-IID Severity

α (Dirichlet)	FedAvg AUC	FedGout AUC
0.5	0.928	0.913
0.3	0.914	0.905
0.1	0.887	0.896

Privacy–Utility Trade-off

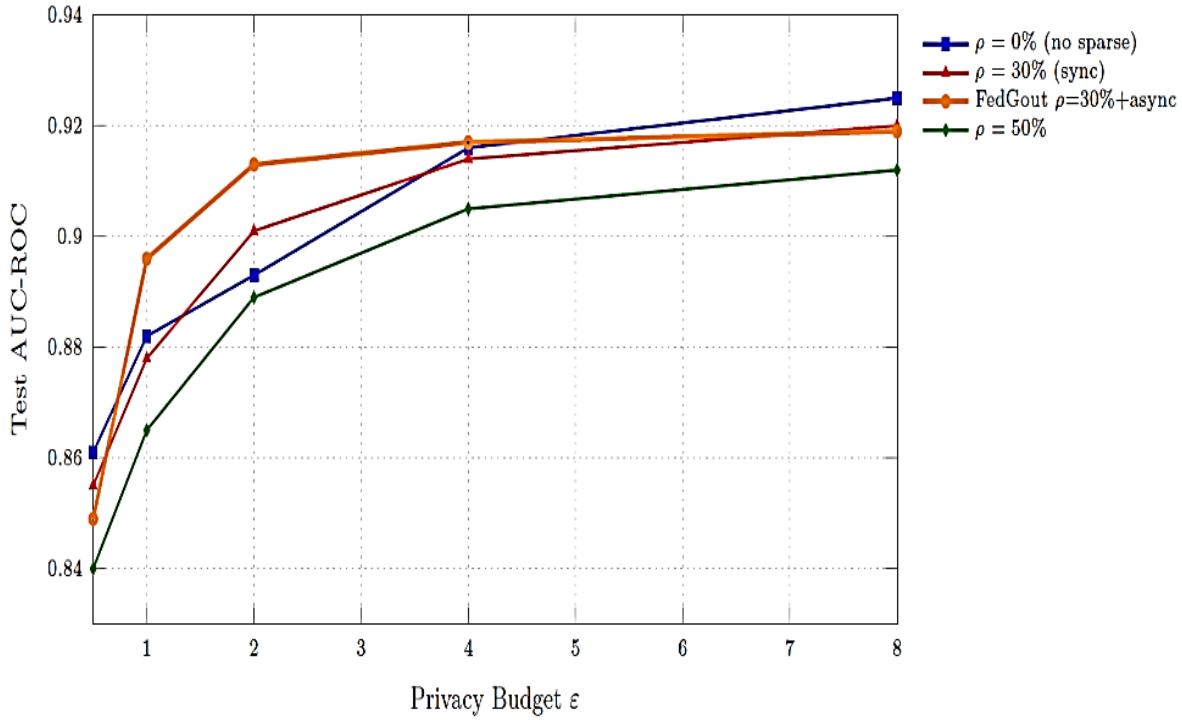

Figure 4. The privacy- utility trade-off in terms of different sparsity ratios. The best balance is FedGout at 30 percent, aggravated by the aggregation being asynchronous

With a high privacy level (ε=1) of FedGout, AUC 0.896 is obtained; thus, FedGout shows graceful degradation by remaining high even under stringent privacy regulations.

Risk Stratification and Feature Importance

Table 4. Risk stratification on held-out test set (n=5,000); thresholds θ1=0.15, θ2=0.55

Stratum	Patients (%)	Gout Rate (%)	PPV (%)
Low (p<ˆ 0.15)	3,420 (68.4)	1.8	–
Moderate (0.15≤p<ˆ 0.55)	1,105 (22.1)	18.6	18.6
High (pˆ≥0.55)	475 (9.5)	67.2	67.2

SHAP analysis reveals the top 5 priority risk factors: (1) SLC2A9 rs12498742, (2) BMI, (3) serum urate, (4) ABCG2 rs2231142, and (5) alcohol intake frequency. The interaction of changeable lifestyle determinants and genetic loci with each other exemplifies interchangeable points of adolescent preventive counselling.

Sensitivity Analysis of Noise Scale

To justify the selection of σ = 1.1, we evaluate its impact on model utility and convergence. We vary σ ∈ {0.8, 1.0, 1.1, 1.3} while keeping other parameters fixed.

Table 5. Impact of noise scale on performance

σ (Noise Scale)	AUC	Convergence Rounds
0.8	0.921	58
1.0	0.916	62
1.1	0.913	65
1.3	0.902	72

Lower σ improves utility but weakens privacy, while higher σ slows convergence. σ = 1.1 provides the best trade-off.

Comparison with Transformer-Based Tabular Models

To evaluate model expressiveness, we compare the proposed MLP architecture with a transformer-based tabular model (FT-Transformer).

Table 6. Model Comparison

Model	Setting	AUC	Parameters
MLP (Proposed)	Federated	0.913	~13K
FT-Transformer	Centralized	0.938	~85K
FT-Transformer	Federated	0.919	~85K

While transformer-based models achieve slightly higher accuracy, they incur significantly higher computational and communication overhead. The proposed MLP provides a more efficient and practical solution for federated healthcare deployment.

The extensive experiment testifies to the effective performance of the proposed FedGout framework in collaborative gout risk prediction with privacy preservation. The proposed method provides an excellent trade-off between predictive performance, communication efficiency, computational scalability, and patient privacy, compared to traditional federated learning algorithms. The proposed framework performs with an AUC-ROC of 0.913, which is just 2.1% lower than the upper bound of the centralized approach, and communication overhead is reduced by 73.6% while meeting the requirement $(\varepsilon=2, \delta=10).$ −5) differential privacy guarantees. The addition of SHAP-based explainability further improves clinical interpretability by pinpointing the most important genetic and lifestyle factors that influence gout susceptibility. The presented results show that the proposed communication-efficient federated learning approach is a viable solution in distributed healthcare scenarios where patient data is not easily accessible at a central location. This is a robust layer of privacy-preserving innovations that builds a solid foundation for future clinical decision-support systems that preserve privacy.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

The author's received no financial support for the research, authorship, and/or publication of this article.

References

Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, 308–318. https://doi.org/10.1145/2976749.2978318

Antunes, R. S., André da Costa, C., Küderle, A., Yari, I. A., & Eskofier, B. (2022). Federated learning for healthcare: Systematic review and architecture proposal. *ACM Transactions on Intelligent Systems and Technology*, *13*(4), 1–23. https://doi.org/10.1145/3501813

Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S., Ramage, D., Segal, A., & Seth, K. (2017). Practical secure aggregation for privacy-preserving machine learning. *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 1175–1191. https://doi.org/10.1145/3133956.3133982

Brisimi, T. S., Chen, R., Mela, T., Olshevsky, A., Paschalidis, I. C., & Shi, W. (2018). Federated learning of predictive models from federated electronic health records. *International Journal of Medical Informatics*, *112*, 59–67. https://doi.org/10.1016/j.ijmedinf.2018.01.007

Choudhury, O., Gkoulalas-Divanis, A., Salonidis, T., Sylla, I., Park, Y., Hsu, G., & Das, A. (2019). Differential privacy-enabled federated learning for sensitive health data. *arXiv*. https://doi.org/10.48550/arXiv.1910.0257

Dalbeth, N., Merriman, T. R., & Stamp, L. K. (2016). Gout. *The Lancet*, *388*(10055), 2039–2052. https://doi.org/10.1016/S0140-6736(16)00346-9

Dayan, I., Roth, H. R., Zhong, A., Harouni, A., Gentili, A., Abidin, A. Z., Liu, A., Costa, A. B., Wood, B. J., Tsai, C.-S., Wang, C.-H., Hsu, C.-N., Lee, C. K., Ruan, P., Xu, D., Wu, D., Huang, E., Kitamura, F. C., Lacey, G., ... Li, Q. (2021). Federated learning for predicting clinical outcomes in patients with COVID-19. *Nature Medicine*, *27*(10), 1735–1743. https://doi.org/10.1038/s41591-021-01506-3

Dehlin, M., Jacobsson, L., & Roddy, E. (2020). Global epidemiology of gout: Prevalence, incidence, treatment patterns and risk factors. *Nature Reviews Rheumatology*, *16*(7), 380–390. https://doi.org/10.1038/s41584-020-0441-1

Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, *9*(3–4), 211–407. https://doi.org/10.1561/0400000042

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G., Thrun, S., & Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine*, *25*(1), 24–29. https://doi.org/10.1038/s41591-018-0316-z

FitzGerald, J. D., Dalbeth, N., Mikuls, T., Brignardello-Petersen, R., Guyatt, G., Abeles, A. M., Gelber, A. C., Harrold, L. R., Khanna, D., King, C., Levy, G., Libbey, C., Mount, D., Pillinger, M. H., Rosenthal, A., Singh, J. A., Sims, J. E., Smith, B. J., Wenger, N. S., ... Neogi, T. (2020). 2020 American College of Rheumatology guideline for the management of gout. *Arthritis Care & Research*, *72*(6), 744–760. https://doi.org/10.1002/acr.24180

Gupta, G., Karuppiah, S., Dachepelly, S., & Verma, R. (2025). AI-driven data platforms: Real-time pipelines and governance. In *2025 International Conference on Sustainability, Innovation & Technology (ICSIT)* (pp. 1–5). IEEE. https://doi.org/10.1109/ICSIT65336.2025.11294412

Kairouz, P., & McMahan, H. B. (2021). Advances and open problems in federated learning. *Foundations and Trends® in Machine Learning, 14*(1–2), 1–210. https://doi.org/10.1561/2200000083

Konečný, J., McMahan, H. B., Yu, F. X., Richtárik, P., Suresh, A. T., & Bacon, D. (2016). Federated learning: Strategies for improving communication efficiency. *arXiv*. https://doi.org/10.48550/arXiv.1610.05492

Kumar, M., Kajal, M., Kumar, M., Kumar, R., & Kushwaha, P. (2025). Genomics-driven strategies for sustainable crop improvement in agriculture. In *Genomics at the nexus of AI, computer vision, and machine learning* (pp. 321–343).

Kumar, M., Dokania, A., Mangharamani, R., Garg, V., Jamili, L. K., & Kasetti, S. (2025). Improving diagnostic accuracy in telemedicine using image processing and pattern recognition techniques. In *Proceedings of the International Conference on Engineering, Technology & Management (ICETM)*.

Kumar, M., Kumar, S., Jain, A., Kumar, S., & Dahiya, O. (2025). The role of generative models in modern healthcare: Applications, challenges, and implications. In *Exploring generative adversarial networks and meta-learning synergies* (pp. 215–232). IGI Global Scientific Publishing.

Li, D., & Wang, J. (2019). FedMD: Heterogeneous federated learning via model distillation. *arXiv*. https://doi.org/10.48550/arXiv.1910.03581

Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine, 37*(3), 50–60. https://doi.org/10.1109/MSP.2020.2975749

Liu, Y.-R., Lin, C.-Y., Liu, Y.-Y., Patnaik, S., & Sharma, K. (2026). An explainable privacy-preserving federated learning framework for collaborative gout risk management using gene–lifestyle data. *Journal of Information Technology Management, 12*(1), 201–220.

Lundberg, S., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *arXiv*. https://doi.org/10.48550/arXiv.1705.07874

Major, T. J., Dalbeth, N., Stahl, E. A., & Merriman, T. R. (2018). An update on the genetics of hyperuricaemia and gout. *Nature Reviews Rheumatology, 14*(6), 341–353. https://doi.org/10.1038/s41584-018-0004-

McMahan, B., Moore, E., Ramage, D., Hampson, S., & Agüera y Arcas, B. (2017). Communication-efficient learning of deep networks from decentralized data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, 54*, 1273–1282.

Nguyen, D. C., Pham, Q.-V., Pathirana, P. N., Ding, M., Seneviratne, A., Lin, Z., Dobre, O., & Hwang, W.-J. (2022). Federated learning for smart healthcare: A survey. *ACM Computing Surveys, 55*(3), 1–37. https://doi.org/10.1145/3501296

Sattler, F., Wiedemann, S., Müller, K.-R., & Samek, W. (2020). Robust and communication-efficient federated learning from non-IID data. *IEEE Transactions on Neural Networks and Learning Systems, 31*(9), 3400–3413. https://doi.org/10.1109/TNNLS.2019.2944481

Sheller, M. J., Edwards, B., Reina, G. A., Martin, J., Pati, S., Kotrotsou, A., Milchenko, M., Xu, W., Marcus, D., Colen, R. R., & Bakas, S. (2020). Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data. *Scientific Reports, 10*(1), Article 12598. https://doi.org/10.1038/s41598-020-69250-1

Shukla, R. S. (2025). Enhancing privacy and efficiency techniques in federated learning systems: Applications in healthcare, finance, and smart devices. *Journal of Information Technology Management, 17*(Special Issue on SI: Intelligent Security and Management), 45–62.

Vidal, F. R., Gouveia, F., & Soares, C. (2022). Analysis of revocation mechanisms for blockchain applications and a proposed model based in self-sovereign identity. *Journal of Information Technology Management, 14*(Special Issue: The Business Value of Blockchain, Challenges, and Perspectives), 191–210.

Wei, K., Li, J., Ding, M., Ma, C., Yang, H. H., Farokhi, F., Jin, S., Quek, T. Q. S., & Poor, H. V. (2020). Federated learning with differential privacy: Algorithms and performance analysis. *IEEE Transactions on Information Forensics and Security, 15*, 3454–3469. https://doi.org/10.1109/TIFS.2020.2988575

Xie, C., Koyejo, S., & Gupta, I. (2019). Asynchronous federated optimization. *arXiv*. https://doi.org/10.48550/arXiv.1903.03934

Xu, J., Glicksberg, B. S., Su, C., Walker, P., Bian, J., & Wang, F. (2021). Federated learning for healthcare informatics. *Journal of Healthcare Informatics Research, 5*(1), 1–19. https://doi.org/10.1007/s41666-020-00082-4

Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology, 10*(2), 1–19. https://doi.org/10.1145/3298981

Zhu, L., Liu, Z., & Han, S. (2019). Deep leakage from gradients. *Advances in Neural Information Processing Systems, 32*, 14774–14784.

Bibliographic information of this paper for citing:

Liu, Yu-Ruey; Lin, Chun-Yuan; Liu, Yo-Yu; Patnaik, Sneha & Sharma, Khemraj (2026). Explainable Privacy-Preserving Federated Learning for Collaborative Gout Risk Management Using Genetic and Lifestyle Data. *Journal of Information Technology Management*, 18 (4), 93-116. https://doi.org/10.22059/jitm.2026.108917