IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510111

A Survey on Federated Learning for Privacy-Preserving Healthcare: Towards Differentially Private Early Sepsis Prediction

Harshil P[1], Sreehari R[2], Kesiya Shaji[3], Adivinayak M V[4], Sukanya M V[5], Girija V R[6]

Department of Computer Science and Engineering (AI & ML), College of Engineering Kottarakkara, Kerala[1-6]

Abstract: Hospitals hold huge amounts of sensitive patient data, but privacy rules and data-ownership concerns mean this data usually cannot be pooled to train machine learning models. Federated Learning (FL) gets around this by sending the model to the data instead of the data to the model. Each institution trains on its own records and shares only model updates. In this paper we survey nineteen studies on federated learning in healthcare published between 2022 and 2026, in ascending order of their citation numbers, which follows the order in which they are first mentioned in the Introduction. The studies fall into seven groups: reviews of FL in healthcare, privacy attacks on FL, cross-hospital generalisation of sepsis models, the dataset source, the base paper on federated sepsis early warning, related work on federated sepsis and Intensive Care Unit (ICU) risk prediction, and differential privacy in FL. For every study we note the objective, the method, the main results and the gap it leaves. The picture is mixed. Federated models can beat single-site models in multi-hospital ICU risk prediction, yet only 5.2% of 612 reviewed healthcare FL studies report real clinical use and only 27.7% include an explicit privacy mechanism. Sepsis models can also lose accuracy at new hospitals, and attacks on FL can reveal not only whether a record was used but which client held it. The ten sepsis studies handle privacy very differently from one another: five report no privacy mechanism in the available text, two run clipping or noise tests with no formal guarantee, one tests a single privacy level, one is centralised only, and one never states its privacy budget. As far as we could tell from the available texts, none gives a patient-level privacy guarantee with accounting across federated rounds. We end by listing the open challenges and by placing FedCare, our planned study of differentially private federated early sepsis prediction across two hospital systems in the open PhysioNet/Computing in Cardiology (CinC) 2019 data, against this gap.

Keywords: Federated Learning (FL), Healthcare, Early Sepsis Prediction, Differential Privacy (DP), Differentially Private Stochastic Gradient Descent (DP-SGD), Membership Inference, Non-Independent and Identically Distributed (Non-IID) Data, PhysioNet/Computing in Cardiology (CinC) 2019.

I. INTRODUCTION

Modern medicine leans more and more on models built from medical images, Electronic Health Records (EHR) and wearable-sensor streams. These models need large, varied datasets, but patient data is sensitive and protected by regulations such as the Health Insurance Portability and Accountability Act (HIPAA) and the General Data Protection Regulation (GDPR) [1]. Collecting records from many hospitals into one repository (a data lake) runs into legal, ethical and technical barriers [1], [2].

Federated Learning (FL) offers a way out. Rather than moving data to a central place, the participating institutions train one shared model together while the raw data stays where it is; only parameters or updates are exchanged [1], [2]. That said, FL is not private by default. A curious server can work out which client a training record came from without breaking the protocol [3], and membership-inference attacks can reveal whether a record was used in training, especially when they are judged at low false-positive rates [4]. Real deployments also have to deal with non-Independent and Identically Distributed (non-IID) data, which can make such attacks more effective [3].

Early sepsis prediction from hourly ICU records is a good test case for these ideas. Hospitals cannot freely share ICU data, sepsis is rare enough that plain accuracy can be misleading, and patient populations and recording habits differ from one hospital to the next. A model built in one hospital system can lose accuracy in another [5], although some models travel well across sites [6]. The open PhysioNet/CinC 2019 challenge data, hosted on PhysioNet [7], come from two hospital systems, so federated and cross-hospital experiments are possible without waiting for credentialed access. Recent federated sepsis and ICU risk-prediction studies [8]–[17] deal with privacy in very different ways. Some describe no mechanism, some add noise with no formal accounting, and some look at a single privacy level or only at centralised training. Differential Privacy (DP) is the standard formal tool for this problem; its use in FL has been reviewed [18] and built into practical frameworks [19].

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510111

With this in mind, we review nineteen studies on federated learning for healthcare, paying particular attention to DP and sepsis. The studies are numbered in the order in which they are first mentioned in this Introduction, and they fall into seven groups: reviews of FL in healthcare [1], [2]; privacy attacks on FL [3], [4]; cross-hospital generalisation of sepsis models [5], [6]; the dataset source [7]; the base paper [8]; related work on federated sepsis and ICU risk prediction [9]–[17]; and differential privacy in FL [18], [19]. The survey contributes a grouping of the literature into these seven categories, a short summary of each study covering what it set out to do, how, what it found and what it leaves open, a comparison of the privacy techniques with a side-by-side look at how the sepsis studies treat privacy, and a statement of the remaining research gap with a list of open challenges. Together these motivate FedCare, a reproducible study of formal patient-level differential privacy for federated early sepsis prediction.

The rest of the paper is organised as follows. Section II covers background on federated learning and differential privacy. Section III reviews the literature. Section IV compares the studies and discusses the findings. Section V states the research gap and where FedCare fits. Section VI lists open challenges and Section VII concludes.

II. BACKGROUND ON FEDERATED LEARNING AND DIFFERENTIAL PRIVACY

Federated learning is formally the problem of minimising a global loss, which is a weighted combination of the local losses on each participant's private data. In the usual hub-and-spoke form based on Federated Averaging (FedAvg) [20], a coordinating server sends the current global model to the participants, each one trains it on its own data, and the server averages the returned weights to make the next global model. Centralised communication of this kind dominates the healthcare FL literature [1], and the modular architecture proposed in [2] likewise places a central FL manager above private local models. Fig. 1 shows this workflow, in which the server uses FedAvg or Federated Proximal (FedProx) [21] aggregation and each hospital adds Differentially Private Stochastic Gradient Descent (DP-SGD) [22] noise on the client side.

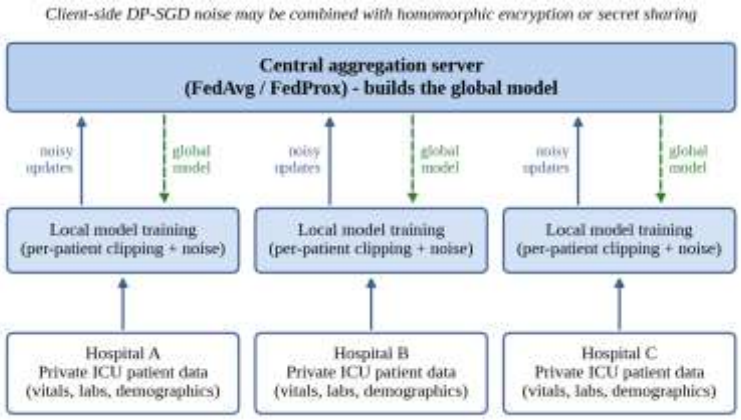


Fig. 1. General architecture of federated learning in healthcare with client-side differential privacy (illustrated with three hospitals; FedCare uses two)

Looking at how data is partitioned, horizontal FL has participants sharing the same features but holding different samples, while in vertical FL they hold overlapping samples with different features, and transfer FL covers the case where the feature spaces differ [18]. Horizontal partitioning dominates both the healthcare literature [1] and the DP-FL literature [18], and it is the setting of the ICU studies surveyed here. FedAvg [20] is the simplest aggregation rule and works well on Independent and Identically Distributed (IID) data, but it needs more rounds when data are heterogeneous. FedProx [21] adds a proximal term that keeps local models close to the global one, which helps under non-IID data. Heterogeneity between hospitals is not a side issue: in [17] differences in demographics, medication use and site characteristics changed which predictors mattered at each hospital, and in [3] more non-IID data made local models overfit more.

Attacks on FL include membership inference, which asks whether a record was in the training set [4], and source inference, in which an honest-but-curious server works out which client a record came from [3]. Source inference works when clients share gradients, model parameters or predictions on a public dataset, and its success is tied to overfitting of the local models [3]. The usual defence with a formal guarantee is differential privacy, which [18] calls the de facto standard for privacy protection in FL. Secure aggregation can be combined with it [19], and a blockchain ledger can record parameter hashes for auditability [15], although it is not itself a privacy mechanism. Other cryptographic tools such as



International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026
DOI: 10.17148/IJARCCE.2026.1510111

homomorphic encryption or secret sharing can be layered on top of client-side DP noise, as the note in Fig. 1 indicates, but they are outside the scope of FedCare.

Differential privacy (DP) puts a limit on how much any one person's data can change what a training procedure outputs. The limit is written as a pair (ε, δ) [23]: a smaller ε means stronger privacy, and δ is a small probability that the guarantee fails. For deep models the usual tool is DP-SGD [22], which clips each per-sample gradient and adds Gaussian noise, with training examples drawn by Poisson sampling. Libraries such as Opacus [24] implement it, and it was used for a sepsis transformer in [10]. Two choices decide what a reported guarantee really means. The first is the privacy unit, which says what is being protected. The review in [18] separates sample-level from client-level DP and lists user-level DP, where one person contributes many records, as an open direction. In ICU data a single patient contributes many hourly rows, so clipping per row does not protect the patient. If a whole patient sequence is treated as one training example, per-sample clipping becomes per-patient clipping. The second is the accountant, which tracks the privacy budget used up over time. Every training step spends some of it, and in FL the steps from all rounds add up. So noise added to a shared update, without a stated unit and an accountant, does not give a reported (ε, δ) guarantee. Stronger privacy also lowers accuracy [3], [19], and the convergence analysis of DP-FL is still an open direction [18].

A membership-inference attack tries to work out whether a particular record was in the training set, and it is the standard empirical check for leakage. Average-case metrics can hide real leakage, so [4] argues that attacks should be judged by the True-Positive Rate (TPR) at a very low False-Positive Rate (FPR). A convincing evaluation therefore reports the Area Under the Receiver Operating Characteristic (ROC) Curve (AUROC, also written AUC) and the TPR at a low FPR, with Confidence Intervals (CIs) over repeated runs. It should also first show, on a deliberately overfit non-private model, that the attack can detect membership at all. A result near chance only means something if the attack passed that sanity check.

III. LITERATURE SURVEY

We discuss the nineteen studies in seven groups, shown in Fig. 2: reviews of FL in healthcare, privacy attacks on FL, cross-hospital generalisation of sepsis models, the dataset source, the base paper, related work on federated sepsis and ICU risk prediction, and differential privacy in FL. The studies are presented in ascending order of their citation numbers, which is the order in which they are first mentioned in the Introduction. Table I in Section IV gives a side-by-side summary. The dataset used by FedCare [7], [25], [26] is described in Section V.

Selection criteria. A study was included if it (i) was published in 2022 or later, up to 2026; (ii) dealt with federated learning in healthcare, federated or multi-site sepsis and ICU risk prediction, differential privacy in federated learning, or privacy attacks on federated learning; (iii) reported a method, an experiment, a validation or a systematic review in enough detail to extract its objective, method, main result and limitation; and (iv) had a full text available, either as a peer-reviewed article or as a public preprint. Duplicate versions of the same work, studies with no bearing on privacy or generalisation in federated clinical prediction, and studies without an accessible full text were excluded. Within these limits we gave priority to work that used PhysioNet/CinC 2019, MIMIC or eICU data, reported a privacy mechanism or attack, or tested a model at a hospital other than the one it was trained on. The studies are therefore the work from 2022 onwards that we judged most relevant to FedCare's question, not an exhaustive search of the literature. Six of them are preprints ([3], [10], [11], [14], [15] and [18]) and are marked as such below. Seven earlier or foundational references [20]–[26], which give the standard algorithms, tools and dataset citations, support the background and methods and are not counted among the surveyed studies.

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510111

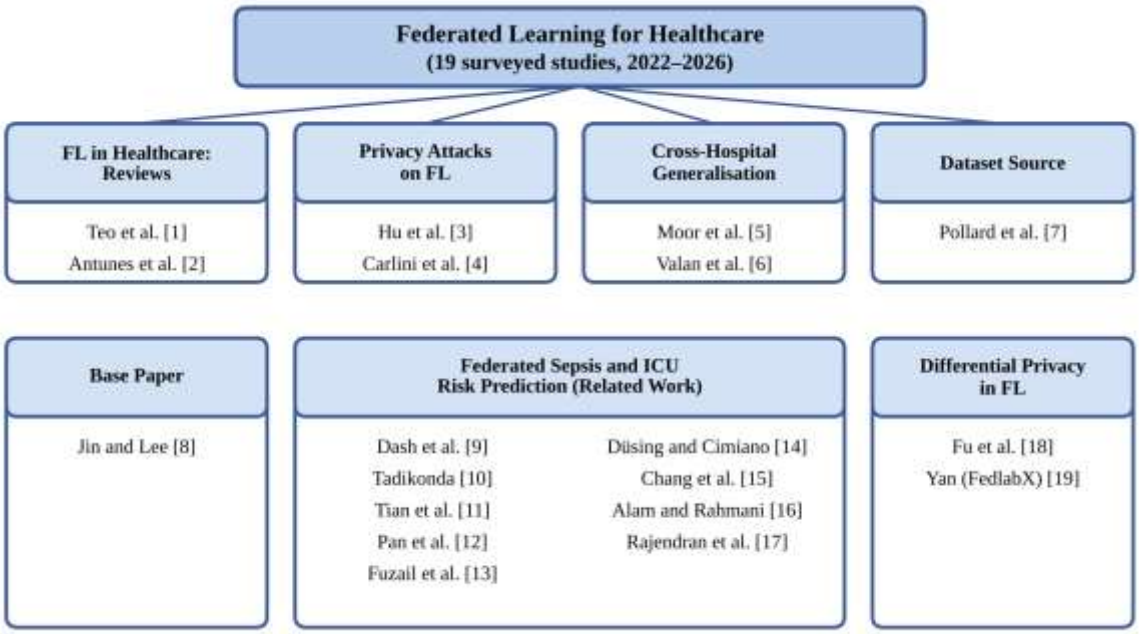


Fig. 2. Categorisation of the surveyed literature

A. Reviews of FL in healthcare

Teo et al. [1] (2024) ran a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)-guided review of 24,060 records from 11 databases and included 612 studies. Only 5.2% (32 of 612) reported real clinical use, and 65% were technical-design studies with a prototype. Medical imaging (41.7%) and EHR (23.7%) were the most common data types. Neural networks were used in 76.3% of the technical studies, horizontal partitioning and centralised communication dominated, and only 27.7% of studies had an explicit privacy mechanism. For FedCare the message is that federated training alone is not enough to claim privacy. FedCare pairs it with DP-SGD and a reported (ε, δ) guarantee, on real multi-hospital ICU data.

Antunes et al. [2] (2022) published a systematic review of federated learning on EHR. They began with 200 articles from seven databases, filtered them down to 67 studies (2016–2021) and sorted these into five topics. Deep learning dominated, appearing in about 79% of case studies, mostly medical imaging and ICU outcome prediction on Medical Information Mart for Intensive Care III (MIMIC-III) and the eICU Collaborative Research Database (eICU). The authors also propose a modular FL architecture with data owners, private local models, and modules for data confidentiality, dataset processing and algorithm verification, run by a central FL manager. It has not been implemented. For FedCare the useful part is the clear separation of local data, local training, privacy protection and central aggregation. FedCare takes that general direction but as a reproducible empirical study: a federated Long Short-Term Memory (LSTM) network for early sepsis prediction with formal differential privacy, on open data.

B. Privacy attacks on FL

Hu et al. [3] (2023, preprint) introduced the source inference attack (SIA), in which an honest-but-curious server works out which client a training record came from without deviating from the FL protocol. They tested it in three set-ups, where clients share gradients (FedSGD), model parameters (FedAvg) or predictions on a public dataset (FedMD), on Synthetic, Purchase, MNIST (Modified National Institute of Standards and Technology handwritten digits), CIFAR-10 and CIFAR-100 (Canadian Institute for Advanced Research image datasets) and FEMNIST (federated EMNIST) data with 10 clients. The attack reached a success rate of 67.1% against a 10% baseline when clients shared model parameters on the Purchase data. Success was tied to overfitting of the local models and rose with more non-IID data and more local epochs. DP-SGD lowered attack success but cost a large amount of accuracy, and in some runs the global model did not converge. Only 10 clients and fairly small models were tested. FedCare uses the same honest-but-curious server assumption, and its two-hospital, non-IID set-up is the kind of condition under which [3] found attacks most effective. It also measures what the privacy noise costs in accuracy, which is the trade-off [3] reports.

Carlini et al. [4] (2022) argued that membership-inference attacks should be judged by their true-positive rate at very low false-positive rates, not by average-case metrics such as AUC, because an attack can look weak on average and still identify some members with confidence. They proposed the Likelihood Ratio Attack (LiRA), which treats membership

DOI: 10.17148/IJARCCE.2026.1510111

inference as a hypothesis test: shadow models give the distribution of an example's scaled confidence when it is in and out of the training set, and a Gaussian likelihood ratio decides membership. On CIFAR-10 with 256 shadow models, LiRA reached 8.4% TPR at 0.1% FPR against about 1.7% for the strongest earlier attack, and a cheaper offline variant reached 7.1%. The data are images and text rather than clinical time series. FedCare follows the recommendation: its attack is shadow-model or calibrated-loss based, reports AUC and TPR at low FPR with confidence intervals, and is first checked on a deliberately overfit reference model.

C. Cross-hospital generalisation of sepsis models

Moor et al. [5] (2023) developed a deep self-attention model for sepsis prediction and validated it across five ICU databases from three countries (AUMC, eICU, MIMIC-III, Emory and HiRID), with 156,309 harmonised ICU admissions. Using only vitals and laboratory values, the model reached an AUROC of 0.847 on internal validation but 0.761 on external validation. Pooling the predictions of models trained at different sites transferred better to a new hospital than a single-site model, without moving patient data. The authors recommend on-site fine-tuning and prospective evaluation before bedside use. The study is not federated and has no privacy mechanism. For FedCare it shows that accuracy at a hospital that took no part in training is the number that matters, which is why FedCare tests transfer between the two hospital systems and asks how privacy noise changes it.

Valan et al. [6] (2025) tested Sepsis Watch, a model first built at Duke University, on adult emergency-department encounters from four sites of Summa Health in Ohio (2020–2021; 205,005 encounters from 101,584 patients, first 36 hours of each stay). AUROC ranged from 0.906 to 0.960 across the four sites, but the Area Under the Precision-Recall Curve (AUPRC) ranged only from 0.177 to 0.252, and the authors judged the model portable across academic and community settings. The setting is an emergency department rather than an ICU, the sites belong to one health system, and the study is neither federated nor private. The wide gap between AUROC and AUPRC is a reminder that AUROC alone flatters a rare-event model, so FedCare reports AUPRC.

D. Dataset source

Pollard et al. [7] (2026) is a Nature Health paper reflecting on 25 years of PhysioNet as a platform for biomedical research. It describes how PhysioNet has supported tens of thousands of research papers and names secure data sharing, community engagement and scalable research infrastructure as its priorities. It is not a sepsis or FL study. We cite it because PhysioNet hosts the PhysioNet/CinC 2019 challenge data that FedCare uses, and it is the standard citation the dataset page asks users to include.

E. Base paper

Jin and Lee [8] (2026) is the base paper for FedCare. They studied leakage-aware federated learning for ICU sepsis early warning on the official A and B partitions of PhysioNet/CinC 2019. They compared local training, FedAvg and FedProx with LSTM and Transformer models, tested both transfer directions between the two hospital systems and prediction horizons of 3, 6 and 12 hours, removed the ICU Length of Stay (ICULOS) variable to limit time-to-event leakage, and used a fixed alert-rate evaluation. They also ran membership-inference and clipping-noise stress tests. The authors say plainly that these tests give no formal DP guarantee. No privacy accounting is reported, and the membership-inference AUC was near chance. FedCare uses the same dataset and also removes ICULOS, but adds DP-SGD with per-patient clipping and a reported accountant.

F. Federated sepsis and ICU risk prediction

Dash et al. [9] (2026) proposed a privacy-aware framework for early sepsis detection that combines personalised FL with differential privacy on Medical Information Mart for Intensive Care IV (MIMIC-IV), using five simulated clients. They used Extreme Gradient Boosting (XGBoost), a Bidirectional LSTM (Bi-LSTM) and a stacked ensemble at a single privacy setting ($\varepsilon = 3$, $\delta = 10^{-5}$). Membership-inference accuracy was 64.5% for standard FL and 51.1% for their method, and they also ran a cross-dataset check on eICU. Noise is applied to the shared update, but the privacy unit and the accounting across rounds are not described, only one ε is tested, and the attack is described only in words. FedCare includes $\varepsilon = 3$ so results can be compared, but sweeps ε from 1 to 8 with a stated privacy unit and a fully specified attack.
Tadikonda [10] (2024, preprint) built a centralised, interpretable transformer for early sepsis prediction from ICU time series and trained it with DP-SGD in Opacus on MIMIC-IV, at $\varepsilon = 1$, 3 and 5 and without DP. The model is not federated. It shows DP-SGD can be applied to sepsis time series but says nothing about multi-hospital training. FedCare covers both cases: centralised training to isolate the cost of DP alone, and federated training.

Tian et al. [11] (2026, preprint) studied multi-centre sepsis early prediction with privacy preservation, using 648 patients from three Chinese hospitals. They used FedAvg with Laplace noise whose sensitivity is estimated, and tested a reconstruction attack. The dataset is small and tabular, the noise scale is estimated from the update itself, and no round-

DOI: 10.17148/IJARCCE.2026.1510111

to-round accounting was seen. The attack-based evaluation is a good idea. FedCare applies it to a larger open time-series dataset and fixes the privacy unit and accountant in advance.

Pan et al. [12] (2024) proposed an adaptive FL framework for clinical risk prediction (sepsis and Acute Kidney Injury (AKI)) on the seven largest eICU hospitals. It uses no differential privacy, eICU needs credentialed access, and some hospitals have very few sepsis cases. It shows that multi-hospital ICU federation is workable and that low prevalence at small sites is a real problem. FedCare uses open data and reports AUPRC because sepsis is rare.

Fuzail et al. [13] (2026) combined hybrid federated learning with Explainable AI (XAI) for sepsis mortality prediction, using federated tree ensembles on a Kaggle copy of the CinC 2019 data. Differential privacy is left as future work. Oversampling before splitting, and the use of an ICU length-of-stay feature, raise leakage concerns. For this reason FedCare splits by patient before any preprocessing, fits scalers and imputers on training data only, uses a weighted loss instead of oversampling, and removes ICULOS.

Düsing and Cimiano [14] (2025, preprint) proposed a federated attention-enhanced LSTM for sepsis onset prediction with variable prediction horizons. Its abstract does not mention differential privacy. It supports the use of a federated LSTM for sepsis, and FedCare adds a formal privacy layer on top.

Chang et al. [15] (2026, preprint) proposed a sepsis early-warning framework for multi-centre ICUs that combines FL with a medical knowledge graph, a temporal transformer and model-agnostic meta-learning for local adaptation, with a blockchain ledger that records parameter hashes for auditability. They split MIMIC-IV (76,540 ICU stays) and eICU (200,859 stays) across 5 to 20 simulated hospital nodes and report an AUC of 0.956 six hours before onset, against 0.781 for a centralised LSTM and 0.848 for standard FL. Each client clips its gradient and adds Gaussian noise (DP-FedAvg), but the privacy budget (ε, δ) is never given as a number in the text we read, and no confidence intervals are reported. The hospitals are simulated and the gain over centralised training is large enough to need independent checking. FedCare reports every ε it uses, with a stated privacy unit, and keeps the model simple enough to isolate the cost of privacy.

Alam and Rahmani [16] (2023) built FedSepsis, a federated multimodal system for early sepsis detection on MIMIC-III that combines vitals, laboratory values and demographics with natural-language processing of clinical notes. It was run on Raspberry Pi and Jetson Nano edge devices, with CPU, memory, temperature and network load recorded. The best multimodal set-up with generative adversarial networks reached an AUPRC of 96.55%, an AUROC of 99.35% and an earliness of 4.56 hours, and with a moderate number of devices the federated set-up performed like a single server. The abstract mentions no differential privacy or encryption, so privacy rests on keeping data on the devices. FedCare shares the federated sepsis aim but adds a formal privacy guarantee and is a simulation, not an edge deployment.

Rajendran et al. [17] (2023) studied data heterogeneity in FL on eICU data from seven hospitals, for acute kidney injury and sepsis risk. They compared local, pooled and federated training with logistic regression and a multilayer perceptron, and used SHapley Additive exPlanations (SHAP) values to compare which features mattered at each site. Federated models generally outperformed local ones, while pooled models often did not beat local ones. Demographic and medication features differed most across sites, and lab tests and vital signs were more consistent. The study uses no privacy mechanism, and eICU needs credentialed access. For FedCare it supports comparing local, centralised and federated training, and treating differences between real hospitals as genuine heterogeneity.

G. Differential privacy in FL

Fu et al. [18] (2024, preprint) wrote a systematic review of differentially private FL covering more than 70 articles. They organise the field by FL scenario (horizontal, vertical and transfer) and by DP model: central DP with sample-level or client-level neighbouring datasets, local DP, and the shuffle model. Horizontal FL dominates, local DP struggles in high-dimensional models because the budget is split across dimensions, and DP combined with secure aggregation often needs discrete noise mechanisms. They list six open directions, including convergence analysis, user-level DP in horizontal FL, and privacy auditing through practical attacks. FedCare takes up two of these on a clinical task: patient-level (user-level) clipping with a stated unit, and auditing the guarantee with a membership-inference attack.

Yan et al. [19] (2024) introduced FedlabX, which combines asynchronous messaging, Diffie-Hellman (DH) based secure aggregation, gradient differential privacy and role-based anonymous authentication. They also study how differential privacy behaves under different data distributions. Experiments use MNIST and CIFAR-10. It lines up closely with FedCare's goals of federated training, protected updates and differential privacy, and it can guide the privacy design while the implementation stays focused on sepsis prediction.

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510111

IV. COMPARATIVE ANALYSIS AND DISCUSSION

Table I summarises the surveyed studies. The discussion that follows pulls out themes that run across all seven groups. Across the multi-hospital studies, federation helps. In [17] federated models generally beat models trained at a single hospital on acute kidney injury and sepsis risk, while models pooled from all sites often did not beat local ones. The edge-device study [16] found that, with a moderate number of devices, federated training performed like a single server, and [15] reports a federated AUC of 0.956 against 0.781 for a centralised LSTM, although its hospitals are simulated and no confidence intervals are given. These results are the main motivation for FedCare: several hospitals can contribute to a shared model without pooling raw patient records. The non-federated studies add a warning. In [5] a sepsis model fell from an AUROC of 0.847 internally to 0.761 at new sites, although pooling models from several sites transferred better than a single-site model, and in [6] AUROC stayed high across four emergency departments but AUPRC stayed low. We will test the federated case for early sepsis prediction by comparing local-only, centralised and federated training.

Privacy mechanisms nearly always cost something in accuracy, efficiency or convergence. In [3], DP-SGD lowered source-inference success but caused large accuracy drops, and in some runs the global model did not converge. FedlabX [19] shows the effect of gradient DP on model accuracy at strict ε. In the sepsis literature, [10] reports centralised DP-SGD at three values of ε, [9] reports one federated setting at $\varepsilon = 3$, and [15] describes DP-FedAvg without giving the privacy budget in the text we read. None of the reviewed sepsis studies gives a multi-ε curve in a federated setting. For FedCare this trade-off is the main thing to measure: we compare ε of 1, 3, 5 and 8, and no DP, for the same federated LSTM, and report AUPRC to account for sepsis prevalence. Table II lists the privacy techniques found in the surveyed studies with their reported benefits and limits.

Non-IID data is a central difficulty in [3] and [17]. In [3] it raised the success of source-inference attacks, and in [17] it changed which features mattered at each hospital. The sepsis studies respond with adaptive strategies [12], meta-learning [15] and local-versus-federated comparisons [17]. This matters because hospitals differ in patient populations, clinical practice and recorded variables. In the CinC 2019 data the difference between the two hospital systems is real, and [8] tests transfer in both directions. Splitting one hospital into random shards is different: it adds clients but not real heterogeneity. FedCare says so explicitly.

Leakage testing is uneven. Among the sepsis studies only [8], [9] and [11] run an attack, and [4] shows why that is not enough: average-case numbers such as AUC can look harmless while an attack still identifies some members with confidence at a very low false-positive rate. Source inference [3] adds a second question that, as far as we could tell, none of the sepsis studies asks, namely which client a record came from, and it matters most when clients are as different as two real hospitals.

Communication and energy cost matter when clients exchange updates again and again. Most of the healthcare literature uses a centralised communication architecture [1], [16] recorded CPU, memory, temperature and network load on Raspberry Pi and Jetson Nano devices, and frameworks such as FedlabX [19] deal with deployment concerns like authentication, messaging and network isolation. FedCare is a single-machine simulation, so communication cost is not a main outcome. It uses the common server-client set-up with FedAvg and FedProx, where each client trains locally with DP-SGD and keeps its own privacy accountant.

TABLE I SUMMARY OF THE SURVEYED STUDIES

Ref.	Authors (Year)	Focus / Technique	Data / Setting	Key Findings
[1]	Teo et al. (2024)	PRISMA systematic review of FL in healthcare	612 studies from 24,060 records	Only 5.2% real clinical application; 27.7% use an explicit privacy mechanism.
[2]	Antunes et al. (2022)	Systematic review of FL on EHR; modular FL architecture	67 studies (2016–2021) filtered from 200	About 79% of case studies use deep learning; architecture is conceptual only.

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510111

Ref.	Authors (Year)	Focus / Technique	Data / Setting	Key Findings
[3]	Hu et al. (2023)	Source inference attack by an honest-but-curious server (FedSGD, FedAvg, FedMD)	Synthetic, Purchase, MNIST, CIFAR-10, FEMNIST, CIFAR-100; 10 clients	Up to 67.1% attack success (10% baseline); tied to overfitting and non-IID data; DP-SGD lowers it at a large accuracy cost.
[4]	Carlini et al. (2022)	LiRA: likelihood-ratio membership inference judged by TPR at low FPR	CIFAR-10, CIFAR-100, WikiText-103 and others	About 10 times more powerful at low FPR; 8.4% TPR at 0.1% FPR on CIFAR-10.
[5]	Moor et al. (2023)	Deep self-attention sepsis model; external validation across sites	Five ICU databases, three countries; 156,309 admissions	AUROC 0.847 internal, 0.761 external; pooling site models transfers better than one site.
[6]	Valan et al. (2025)	External validation of Sepsis Watch	4 emergency departments; 205,005 encounters	AUROC 0.906–0.960 but AUPRC only 0.177–0.252; not federated, no privacy mechanism.
[7]	Pollard et al. (2026)	Overview of PhysioNet as a research platform	PhysioNet (hosts the CinC 2019 data)	25 years of open data sharing; standard citation for the platform.
[8]	Jin and Lee (2026)	Leakage-aware FL for ICU sepsis (base paper); local, FedAvg, FedProx; LSTM and Transformer	CinC 2019 official sets A/B; horizons 3/6/12 h	Both transfer directions; ICULOS removed; membership-inference AUC near chance; no formal DP or accounting.
[9]	Dash et al. (2026)	Personalised FL + DP; XGBoost, Bi-LSTM, stacked ensemble	MIMIC-IV, 5 simulated clients; eICU check	One setting ($\varepsilon = 3$, $\delta = 1e\text{-}5$); membership-inference accuracy 64.5% vs 51.1%. Privacy unit and accounting not described.
[10]	Tadikonda (2024)	Centralised transformer with DP-SGD (Opacus)	MIMIC-IV	ε of 1, 3, 5 and no DP; not federated. Shows DP-SGD on sepsis time series.
[11]	Tian et al. (2026)	FedAvg with Laplace noise; reconstruction attack	648 patients, 3 Chinese hospitals	Small tabular data; noise scale estimated from the update; no round-to-round accounting seen.
[12]	Pan et al. (2024)	Adaptive FL for sepsis and AKI risk prediction	7 largest eICU hospitals	No DP; credentialed data; very few sepsis cases at some hospitals.
[13]	Fuzail et al. (2026)	Federated tree ensembles with XAI	CinC 2019 (Kaggle copy)	DP left as future work; oversampling before splitting and an ICU length-of-stay feature raise leakage concerns.
[14]	Düsing and Cimiano (2025)	Federated attention-enhanced LSTM, variable horizons	Sepsis onset prediction	Abstract does not mention differential privacy. Supports federated LSTM for sepsis.
[15]	Chang et al. (2026)	FL with knowledge graph, temporal transformer and meta-learning; DP-FedAvg	MIMIC-IV and eICU; 5–20 simulated nodes	AUC 0.956 vs 0.781 centralised LSTM; (ε, δ) not given as numbers; no confidence intervals.
[16]	Alam and Rahmani (2023)	FedSepsis: multimodal FL on edge devices	MIMIC-III; Raspberry Pi and Jetson Nano	AUPRC 96.55%, AUROC 99.35%, earliness 4.56 h; no DP or encryption mentioned.

 This work is licensed under a Creative Commons Attribution 4.0 International License



International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026
DOI: 10.17148/IJARCCE.2026.1510111

Ref.	Authors (Year)	Focus / Technique	Data / Setting	Key Findings
[17]	Rajendran et al. (2023)	Data heterogeneity in FL: local vs pooled vs federated	eICU, 7 hospitals; AKI and sepsis	Federated generally beat local; pooled often did not; demographic and medication features differed most across sites.
[18]	Fu et al. (2024)	Systematic review of differentially private FL; taxonomy of DP models	More than 70 articles	Horizontal FL dominates; local DP struggles in high dimensions; open: user-level DP, privacy auditing, convergence.
[19]	Yan et al. (2024)	FedlabX: Kafka messaging, DH aggregation, gradient DP	MNIST, CIFAR-10	98.4% (MNIST-IID), 80.3% (CIFAR-10-IID); accuracy falls at strict ε. Closely relates FL, DP and secure update protection.

TABLE II PRIVACY-ENHANCING TECHNIQUES IN THE SURVEYED STUDIES

Technique	Studies	Reported Benefit	Reported Limitation
Differential privacy	[3], [9]–[11], [15], [18], [19]	Formal privacy guarantee [18]; DP-SGD applied to a sepsis model at ε of 1, 3, 5 [10]; clipping and Gaussian noise inside FedAvg [15]; DP lowers source-inference success [3]	Large accuracy cost [3]; CIFAR-10 about 44% at ε = 1 [19]; single ε [9]; ε not stated [15]; no round-to-round accounting seen [11]
Empirical leakage testing	[3], [4], [8], [9], [11]	Measures practical leakage: membership-inference AUC near chance [8]; accuracy 64.5% vs 51.1% [9]; reconstruction attack [11]; source inference up to 67.1% [3]; TPR at low FPR shows what averages hide [4]	No formal guarantee [8]; attack described only in words [9]; success depends on overfitting [3]
Secure aggregation	[19]	Diffie-Hellman based secure aggregation combined with gradient DP and anonymous authentication [19]	Evaluated under a semi-honest threat model [19]
Blockchain audit ledger	[15]	Records parameter hashes and versions for traceability [15]	Not itself a privacy mechanism [15]
Federated training only	[12], [13], [14], [16], [17]	Raw data stay at each hospital or device	No formal privacy guarantee; leakage remains possible [3], [4]

Most of the evidence is still at prototype level. Only 5.2% of the reviewed healthcare FL studies report real clinical application, and only 27.7% include an explicit privacy mechanism [1]. There is a clear gap between research ideas and practical privacy-preserving healthcare systems. FedCare aims to contribute at the prototype level with careful measurement rather than a new algorithm, looking at both predictive performance and the privacy trade-offs.

Table III compares the ten sepsis-related studies on the points that decide what a privacy claim is worth. Five add clipping or noise [8]–[11], [15], and five report no privacy mechanism, or leave it for future work, in the text we read [12]–[14], [16], [17]. Of the first five, only [10] uses DP-SGD, and it is centralised, while [15] describes DP-FedAvg but never states ε in the text we read. Study [9] reports one ε without a stated privacy unit, [8] says its tests give no formal guarantee, and in [11] the noise scale is estimated from the update itself. Membership-inference results appear in [8] and [9], and [11] tests a reconstruction attack, but a result near chance only means something if the attack is shown to work on a model known to leak [4]. That is why FedCare first runs a deliberately overfit reference model.

The sepsis studies also differ in how carefully they keep test data separate from training data. Study [8] removes ICULOS, a time counter that can reveal time-to-event information. In [13], oversampling before splitting and an ICU length-of-stay feature raise leakage concerns. The CinC 2019 data deliberately keeps missing values and errors, and some columns

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510111

are missing systematically in one hospital, so imputation and scaling must be fitted on training patients only. FedCare therefore splits by patient before any preprocessing, handles imbalance with a weighted loss instead of oversampling, and treats one patient's whole ICU sequence as one training example.

V. RESEARCH GAP AND POSITIONING OF FEDCARE

The survey leaves a narrow but defensible gap. Among the reviewed studies, none combines all of the following, as far as we could tell from the available texts. The first is DP-SGD with per-patient clipping and a reported accountant across federated rounds, with the privacy unit stated; the review in [18] lists user-level DP as an open direction. The second is a multi-ε sweep in a federated setting on open sepsis data, since related work uses one ε, gives none, or is centralised. The third is the effect of the privacy level on accuracy at an external hospital that took no part in training; [5] shows that non-private models lose accuracy at new sites, but privacy noise is not part of that picture. The fourth is a reproducible attack description with confidence intervals that reports TPR at low FPR as argued in [4], including an honest report if the attack finds no signal.

FedCare is a reproducible empirical study aimed at this gap. It does not claim to be the first differentially private federated sepsis study, and it is not a new algorithm. What it offers is a more careful privacy treatment and a cross-hospital measurement on an open dataset, with the code released. It uses the PhysioNet/CinC 2019 data [7], [25], [26]: 40,336 ICU records from two hospital systems (set A with 20,336 patients, set B with 20,000), hourly rows with 40 variables, and a label that turns positive six hours before the adjudicated sepsis time. A federated LSTM is trained with FedAvg [20], using FedProx [21] as a comparison, and DP-SGD [22] is implemented with Opacus [24].

TABLE III PRIVACY TREATMENT IN SEPSIS-RELATED FEDERATED LEARNING STUDIES

Study	Data	Federated	Privacy mechanism	ε / δ	Accounting	Attack evaluation
[8] Jin and Lee	CinC 2019 (A/B)	Yes	Clipping and noise stress tests (no formal DP)	None	None	Membership inference; AUC near chance
[9] Dash et al.	MIMIC-IV; eICU check	Yes (5 simulated clients)	DP noise on shared update	ε = 3, δ = 1e-5	Not specified	Membership inference; 64.5% vs 51.1%
[10] Tadikonda	MIMIC-IV	No (centralised)	DP-SGD (Opacus)	ε = 1, 3, 5	Not specified	Not specified
[11] Tian et al.	648 patients, 3 hospitals	Yes	Laplace noise, estimated sensitivity	Not specified	None seen	Reconstruction attack
[12] Pan et al.	eICU, 7 hospitals	Yes	None	Not specified	Not specified	Not specified
[13] Fuzail et al.	CinC 2019 (Kaggle copy)	Yes	None (future work)	Not specified	Not specified	Not specified
[14] Düsing and Cimiano	Sepsis onset data	Yes	Not mentioned in abstract	Not specified	Not specified	Not specified
[15] Chang et al.	MIMIC-IV and eICU	Yes (5–20 simulated nodes)	DP-FedAvg: client clipping and Gaussian noise	Not given as numbers	Not specified	Not specified
[16] Alam and Rahmani	MIMIC-III	Yes (edge devices)	None mentioned	Not specified	Not specified	Not specified

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510111

Study	Data	Federated	Privacy mechanism	ε / δ	Accounting	Attack evaluation
[17] Rajendran et al.	eICU, 7 hospitals	Yes	None	Not specified	Not specified	Not specified
FedCare (planned)	**CinC 2019 (A/B)**	**Yes**	**DP-SGD: per-patient clipping, Gaussian noise, Poisson sampling**	**ε = 1, 3, 5, 8 and no DP; fixed δ**	**Per-client accountant over all rounds**	**Shadow-model or calibrated-loss attack; AUC and TPR at low FPR, with CIs**

Note: "Not specified" means that the information is not given in the available text of the study. For [14] and [16] only the abstract was checked.

The study asks four Research Questions (RQs). RQ1 asks how much accuracy an LSTM loses at each level of formal patient-level privacy (ε from 1 to 8), both centralised and federated. RQ2 asks how privacy noise changes accuracy on a held-out hospital, when we train on one hospital system and test on the other. RQ3 asks whether more clients and different numbers of rounds change the privacy–utility trade-off, which the review in [18] lists as an open convergence question. RQ4 asks whether a well-specified membership-inference attack, judged by TPR at low FPR [4], succeeds against non-private and private models, and how that changes with ε.

We use three client set-ups: the two real hospitals as clients; four random shards of one hospital with the other hospital as an external test set (in both directions); and four shards of each hospital. Results are reported as AUROC and AUPRC with 95% confidence intervals over five seeds. The limits are worth stating up front. Only two real hospitals are available, the shards inside a hospital are simulated, the guarantee covers training-data patients only and not hyperparameter search, the server is assumed honest-but-curious, and secure aggregation, personalisation, live deployment and legal-compliance claims are out of scope.

VI. OPEN CHALLENGES AND FUTURE DIRECTIONS

The literature points to several challenges that bear directly on how FedCare is designed and evaluated. Differences in patient populations and data distributions across hospitals can affect aggregation and make attacks more effective [3], [17], and performance can drop at hospitals that were not part of the original training group [5]. Only the difference between real hospitals is genuine heterogeneity, so simulated shards should not be presented as such, and the effect of privacy noise on an unseen hospital is not documented for sepsis.

FL does not fully prevent leakage: membership inference [4] and source inference [3] both remain possible. Keeping raw data local is not enough; differential privacy with a stated unit and accounting, plus a reproducible attack evaluation, is needed. Stronger privacy also lowers accuracy [3], [19], and sepsis is rare, so privacy noise could leave models nearly useless. Even non-private models can show a large gap between AUROC and AUPRC [6]. This has to be measured with AUPRC and reported even if the result is negative.

Choices like oversampling before splitting or keeping time-counter features can inflate results [8], [13], so patient-level splits and training-only preprocessing are needed. Repeated exchange of updates can also get costly as the number of clients grows, particularly on limited edge devices [16], and real deployment faces barriers such as infrastructure, explainability, standardisation and the need for reliable healthcare operation [1]. FedCare is a research study and makes no clinical or legal-compliance claims.

Future work includes secure aggregation, personalisation, source-inference testing [3], privacy auditing [18], bringing hyperparameter tuning inside the privacy guarantee, dynamic client participation and more hospitals, including the challenge's hidden third hospital system, which is not available.

VII. CONCLUSION

The reviewed studies show that federated learning lets healthcare institutions collaborate without directly sharing patient records, and in multi-hospital ICU risk prediction federated models often do better than models trained at a single hospital. They also show that FL is not private by default, that differential privacy and other security techniques add protection but cost some utility, and that sepsis models can lose accuracy at new hospitals. Recent federated sepsis studies handle

 This work is licensed under a Creative Commons Attribution 4.0 International License

DOI: 10.17148/IJARCCE.2026.1510111

privacy very differently: some report no mechanism, some add noise without formal accounting or a stated budget, and some evaluate only one privacy level or only centralised training.

FedCare addresses this gap as a reproducible research study. It trains a federated LSTM for early sepsis prediction on the open PhysioNet/CinC 2019 data with DP-SGD, a stated patient-level privacy unit and a reported (ε, δ) accountant. It then measures what each privacy level costs in accuracy, how privacy affects an external hospital, and how well a clearly specified membership-inference attack works. Secure aggregation and personalisation can be explored once this study is complete.

REFERENCES

[1] Z. L. Teo et al., "Federated machine learning in healthcare: A systematic review on clinical applications and technical architecture," *Cell Reports Medicine*, vol. 5, no. 2, Art. no. 101419, 2024, doi: 10.1016/j.xcrm.2024.101419.

[2] R. S. Antunes, C. A. da Costa, A. Küderle, I. A. Yari, and B. Eskofier, "Federated learning for healthcare: Systematic review and architecture proposal," *ACM Trans. Intell. Syst. Technol.*, vol. 13, no. 4, Art. no. 54, 2022, doi: 10.1145/3501813.

[3] H. Hu, X. Zhang, Z. Salcic, L. Sun, K.-K. R. Choo, and G. Dobbie, "Source inference attacks: Beyond membership inference attacks in federated learning," arXiv:2310.00222, 2023 (accepted, *IEEE Trans. Dependable Secure Comput.*).

[4] N. Carlini, S. Chien, M. Nasr, S. Song, A. Terzis, and F. Tramèr, "Membership inference attacks from first principles," in *Proc. IEEE Symp. Security and Privacy (SP)*, 2022, arXiv:2112.03570.

[5] M. Moor et al., "Predicting sepsis using deep learning across international sites: A retrospective development and validation study," *eClinicalMedicine*, vol. 62, Art. no. 102124, 2023, doi: 10.1016/j.eclinm.2023.102124.

[6] B. Valan et al., "Evaluating Sepsis Watch generalizability through multisite external validation of a sepsis machine learning model," *npj Digital Medicine*, vol. 8, Art. no. 350, 2025, doi: 10.1038/s41746-025-01664-5.

[7] T. Pollard et al., "PhysioNet as a global platform for biomedical research," *Nature Health*, vol. 1, no. 8, pp. 792–795, 2026, doi: 10.1038/s44360-026-00096-z.

[8] H. Jin and H. Lee, "Leakage-aware federated learning for ICU sepsis early warning: Fixed alert-rate evaluation on PhysioNet/CinC 2019 and MIMIC-IV," *Applied Sciences*, vol. 16, no. 6, Art. no. 2735, 2026, doi: 10.3390/app16062735.

[9] J. Dash, P. K. Sahu, and S. Satpathy, "Privacy-aware artificial intelligence framework for early sepsis detection in sustainable healthcare systems," *Discover Computing*, vol. 29, Art. no. 508, 2026, doi: 10.1007/s10791-026-10438-w.

[10] D. S. Tadikonda, "A privacy-preserving and interpretable transformer-based model for early prediction of sepsis in ICU patients using time-series data," SSRN preprint, 2024.

[11] X. Tian et al., "Federated learning for multi-center sepsis early prediction with privacy-preserving," arXiv:2606.04338, 2026.

[12] W. Pan, Z. Xu, S. Rajendran, and F. Wang, "An adaptive federated learning framework for clinical risk prediction with electronic health records from multiple hospitals," *Patterns*, vol. 5, no. 1, Art. no. 100898, 2024, doi: 10.1016/j.patter.2023.100898.

[13] M. Z. Fuzail et al., "Optimizing sepsis mortality prediction using hybrid federated learning and explainable AI framework," *Scientific Reports*, vol. 16, Art. no. 5218, 2026, doi: 10.1038/s41598-026-36245-3.

[14] C. Düsing and P. Cimiano, "Improving early sepsis onset prediction through federated learning," arXiv:2509.20885, 2025.

[15] Y. Chang, G. Lin, J. J. Chuang, S. Liu, X. Li, and Y. Li, "A federated learning framework with knowledge graph and temporal transformer for early sepsis prediction in multi-center ICUs," arXiv:2603.15651, 2026.

[16] M. U. Alam and R. Rahmani, "FedSepsis: A federated multi-modal deep learning-based Internet of Medical Things application for early detection of sepsis from electronic health records using Raspberry Pi and Jetson Nano devices," *Sensors*, vol. 23, no. 2, Art. no. 970, 2023, doi: 10.3390/s23020970.

[17] S. Rajendran, Z. Xu, W. Pan, A. Ghosh, and F. Wang, "Data heterogeneity in federated learning with electronic health records: Case studies of risk prediction for acute kidney injury and sepsis diseases in critical care," *PLOS Digital Health*, vol. 2, no. 3, Art. no. e0000117, 2023, doi: 10.1371/journal.pdig.0000117.

[18] J. Fu et al., "Differentially private federated learning: A systematic review," arXiv:2405.08299, 2024.

[19] Y. Yan et al., "FedlabX: A practical and privacy-preserving framework for federated learning," *Complex & Intelligent Systems*, vol. 10, pp. 677–690, 2024, doi: 10.1007/s40747-023-01184-3.