International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 4, July-Aug-2026, ISSN (Online): 2395-566X

FedBioGuard: A Privacy-Preserving and Uncertainty-Aware Multimodal Federated Framework for Explainable Antimicrobial Resistance Prediction

Deepa Shree R
Dept computational intelligence

Abstract — The rapid rise of antimicrobial resistance necessitates robust diagnostic tools that can integrate heterogeneous clinical data across decentralized healthcare institutions while ensuring patient data confidentiality. To address these requirements, this framework leverages federated learning to enable collaborative model training across institutions without necessitating data centralization (Zwiers et al., 2024), while simultaneously incorporating Bayesian inference to quantify predictive uncertainty and enhance clinical interpretability (Kumari et al., 2026). Furthermore, by synthesizing multimodal electronic health records, the proposed architecture overcomes the limitations of centralized data silos and addresses the inherent challenges of data bias and clinical validation in AMR research (Hardan et al., 2024; Narra et al., 2024). Antimicrobial resistance makes infections harder to treat, so we need faster and more reliable ways to detect resistant pathogens. Artificial intelligence can help predict resistance using genomic and other biological data (Lastra et al., 2024). However, many current methods depend on centralized datasets (Zwiers et al., 2024), offer little insight into how certain their predictions are, and are difficult to interpret in high-stakes settings (Kumari et al., 2026). In addition, healthcare and biological data are often spread across institutions and cannot be freely shared because of privacy, governance, and regulatory requirements. This paper presents FedBioGuard, a privacy-preserving framework for predicting antimicrobial resistance. It combines federated learning, multimodal data, uncertainty estimates, explainable AI, and evidence-based large language model support. The main goal is to examine whether this approach can make reliable AMR predictions across institutions with different types of data, without sharing raw patient records. Each participating institution keeps its data locally and helps train a shared model by sending model updates. The model can use genomic data, microbiome or metagenomic data when available, and structured clinical information. An uncertainty module shows how reliable each prediction may be, while explainability methods highlight the biological and clinical features that influence the results. A retrieval-augmented language model is used only after prediction to turn the results into clear summaries supported by scientific evidence. By decentralizing the training process, FedBioGuard mitigates the risks associated with data privacy while addressing the limitations of traditional cultivation-based diagnostics that are often too slow to guide immediate treatment decisions (Dayan et al., 2021; Inda-Díaz et al., 2026). By addressing technical hurdles such as data heterogeneity and the "black box" nature of deep learning, this framework fosters the development of transparent, robust clinical decision support systems (Cavallaro et al., 2023). The evaluation will compare centralized, local-only, single-modality, and federated models using simulated data from multiple institutions. The models will be assessed for predictive performance, calibration, uncertainty quality, performance under distribution shifts, and consistency of their explanations. The study aims to provide a reproducible way to examine how privacy-preserving collaboration, multimodal learning, and uncertainty estimation can work together for AMR prediction. FedBioGuard is not intended to replace laboratory susceptibility testing; instead, it is a research framework for studying trustworthy AI-assisted AMR analysis.

Keywords— antimicrobial resistance; federated learning; multimodal learning; uncertainty quantification; explainable artificial intelligence; bioinformatics; trustworthy AI; large language models

I. INTRODUCTION

Antimicrobial resistance (AMR) is a major global health challenge because microorganisms can become resistant to medicines that were previously effective against them (Tang et al., 2023). The resulting infections can become harder to treat and may increase the risks of severe disease, prolonged hospitalization, disability, and death. The World Health Organization continues to identify AMR as a major global health threat (Larkin, 2023).

The scale of the future burden remains substantial. Recent global forecasts estimate that AMR could be directly attributable to approximately 1.91 million deaths annually and associated with approximately 8.22 million deaths annually by 2050 under the reference scenario (Bakri et al., 2025). These estimates differ from the older and widely repeated projection

of 10 million annual deaths by 2050 (Kraker et al., 2016), highlighting the importance of using updated evidence when describing the AMR burden.

Traditional antimicrobial susceptibility testing remains essential for determining whether a pathogen is resistant or susceptible to a particular antimicrobial agent. However, culture-based testing can require substantial time, motivating research into computational methods that may support earlier analysis. Whole-genome sequencing and other molecular technologies have expanded the ability to identify resistance-associated genes and mutations (Olatunji et al., 2024). Nevertheless, converting complex biological information into reliable phenotype predictions remains difficult because AMR can depend on interactions among genetic variation, species-specific mechanisms, environmental factors, and other contextual variables (Hughes & Andersson, 2017).

Machine learning has therefore become an important research direction for AMR prediction (Lastra et al., 2024). Recent reviews indicate growing interest in genomic, multimodal, foundation-model, and clinically deployable AI systems for AMR (Lastra et al., 2024). At the same time, the literature identifies persistent challenges involving dataset heterogeneity, interpretability, and cross-platform normalization (Shankarnarayan et al., 2022).

A further challenge is that relevant data are often distributed across hospitals, laboratories, and research institutions. Pooling these records into a central repository may be restricted by privacy, governance, institutional, or regulatory requirements (Scheltjens et al., 2024). Federated learning offers a potential alternative by enabling participants to train models locally and exchange selected model updates rather than raw records. However, federated learning does not automatically solve all challenges associated with decentralized medical AI (Chaddad et al., 2023). Differences among clients in population characteristics, sequencing platforms, laboratory procedures, feature distributions, and resistance prevalence can produce heterogeneous, non-independent and identically distributed data (Madathil et al., 2025). Such heterogeneity can reduce stability and generalizability.

Reliability is particularly important in AMR prediction because a model should communicate not only what it predicts but also how reliable that prediction is. Conventional classifiers frequently produce a probability score, but raw confidence scores are not necessarily well calibrated (Balanya et al., 2024). A model may assign high confidence to an incorrect prediction, particularly when encountering data that differ from its training distribution. Therefore, uncertainty estimation and calibration should be evaluated explicitly rather than inferred from classification accuracy alone (Mehrtash et al., 2020),(Kompa et al., 2021).

Transparency presents another challenge. Complex deep learning models may identify predictive patterns without making their reasoning directly understandable to researchers or clinicians. Explainable AI methods can help identify influential features, although explanations themselves must be evaluated carefully and should not automatically be interpreted as causal evidence (Carriero et al., 2025).

Large language models may also assist researchers by converting technical outputs into accessible summaries. However, an unconstrained language model can produce unsupported statements. For this reason, the language-model component in the proposed framework is not used to determine AMR predictions. Instead, it operates as a post-prediction interpretation layer that receives structured model outputs and retrieves relevant evidence before generating a summary. Evidence grounding is important because current research on LLM factuality continues to identify hallucination and factual inconsistency as significant challenges (Huang et al., 2024).

This paper proposes FedBioGuard, a research framework designed to investigate these challenges jointly. Its central research question is:

Can uncertainty-aware multimodal federated learning improve the reliability and calibration of antimicrobial resistance prediction across heterogeneous datasets while preserving the locality of participating datasets?

The study has four objectives:

- Develop a multimodal predictive architecture for integrating available genomic, microbiome, and structured metadata.
- Train and compare local, centralized, and federated versions of the predictive system.
- Quantify and evaluate predictive uncertainty and calibration under heterogeneous client distributions.
- Provide feature-level explanations and evidence-grounded natural-language summaries without allowing the LLM to alter the underlying prediction.

The main contribution of this work is therefore not the claim that federated learning, uncertainty estimation, multimodal AI, explainability, or RAG are individually new. Instead, the research investigates a clearly defined integration of these components and evaluates their interaction in AMR prediction under simulated decentralized conditions. This distinction is

important because recent AMR-AI literature already identifies federated learning, multimodal architectures, uncertainty-aware modeling, interpretability, and evidence-grounded AI as active research directions.

II. RELATED WORK

1. AI for AMR Prediction

AI-based AMR prediction has increasingly used genomic information, resistance-associated markers, k-mer representations, metadata, and other biological features (García et al., 2025). Classical machine-learning approaches, including logistic regression and random forests, remain important baselines because of their comparatively straightforward interpretation (Moradigaravand et al., 2018). Deep learning has expanded the ability to learn representations from complex and high-dimensional data (López-Cortés et al., 2024).

Despite this progress, recent reviews identify inconsistent evaluation practices and limited cross-site validation as barriers to clinical translation (Ardila et al., 2025),(López-Cortés et al., 2024). Models that perform strongly on internal test sets may not generalize to new populations, laboratories, or geographic regions.

2. Multimodal AMR Modeling

Different biological data modalities may capture complementary aspects of resistance (Žitnik et al., 2018). Genomic data can identify known or potentially relevant resistance-associated variation, while microbiome data may describe community composition and associated ecological patterns (Melo et al., 2021),(Chetty & Blekhman, 2024). Structured metadata may provide additional contextual information (Soenksen et al., 2022).

However, multimodal integration introduces challenges including missing modalities, differing dimensionalities, small sample sizes, and modality imbalance (Farhadizadeh et al., 2025). FedBioGuard therefore treats multimodal integration as conditional on actual data availability rather than assuming that every participating client possesses every modality.

3. Federated Learning for Healthcare and Biological Data

Federated learning enables decentralized optimization by training local models and aggregating selected updates into a global model (Fang et al., 2024). Its central advantage is that raw records can remain within participating sites.

However, keeping raw data local should not be described as complete privacy protection. Model updates may themselves reveal information under certain attack scenarios (Rieke et al., 2020). Therefore, privacy must be evaluated as a design objective rather than assumed solely from the use of federated learning.

4. Uncertainty and Calibration

Predictive uncertainty is especially relevant when a model encounters ambiguous or unfamiliar examples (Kompa et al., 2021). Two predictions with the same predicted class may carry very different levels of reliability.

FedBioGuard distinguishes:

- Predictive performance: whether the predicted class is correct (Flach, 2019).
- Calibration: whether predicted confidence corresponds to observed correctness (Kompa et al., 2021).
- Uncertainty: how much ambiguity the model associates with the prediction (Kompa et al., 2021).

These properties must be measured separately (Scalia et al., 2020).

5. Explainable AI and Evidence-Grounded LLMs

Feature attribution methods can identify which model inputs were influential for a particular output (Baniecki & Biecek, 2024). Depending on the final model architecture, appropriate techniques may include SHAP, Integrated Gradients, permutation importance, or modality-specific attribution methods.

The LLM layer is deliberately separated from the predictive layer. Its purpose is to organize the following information:

- predicted phenotype;
- uncertainty score;
- confidence category;
- influential features;
- retrieved scientific evidence.

The generated explanation should cite retrieved sources and clearly distinguish model output from established biological evidence.

III. RESEARCH GAP

The proposed work addresses the following combined gap:
Existing research has extensively explored AI and machine learning for AMR prediction (Bilal et al., 2025). However, the interaction among decentralized data heterogeneity, multimodal prediction, calibration, and uncertainty-aware decision support requires more systematic evaluation.

The gap will be tested rather than merely asserted through comparison of the following settings:

- local-only learning;
- centralized pooled-data learning, where permitted for experimental comparison;
- standard federated learning;
- uncertainty-aware federated learning;
- single-modality and multimodal variants.

Proposed Framework: FedBioGuard

FedBioGuard contains six layers.

Local Data Layer

Each participating client retains its own dataset (Guo et al., 2023).

Possible inputs include:

$$Xi=\{Gi,Mi,Ci\}X_i = \{G_i, M_i, C_i\}$$

where:

- GiG_i = genomic features;
- MiM_i = microbiome or metagenomic features, when available;
- CiC_i = structured contextual metadata.

The outcome variable is:

$$Yi∈\{0,1\}Y_i \in \{0,1\}$$

where 0 represents susceptible and 1 represents resistant for a predefined species-drug task.

2. Modality-Specific Encoders

Separate encoders transform each modality:

$$zG=fG(G)z_G=f_G(G)zM=fM(M)z_M=f_M(M)zC=fC(C)z_C=f_C(C)$$

These representations are combined through a fusion function:

$$z=F(zG,zM,zC)z = F(z_G,z_M,z_C)$$

The study should compare at least two fusion strategies, such as:

- early concatenation;
- attention-based or gated fusion.

3. AMR Prediction Layer

The fused representation is passed to a classifier:

$$y^=h(z)\hat{y}=h(z)$$

The output estimates the probability of resistance.

4. Federated Training Layer

For client kk, local training produces parameters:

$$θk(t+1)\theta_k^{(t+1)}$$

A weighted aggregation baseline can be expressed as:

$$θ(t+1)=∑k=1Knk∑j=1Knjθk(t+1)\theta^{(t+1)} = \sum_{k=1}^{K} \frac{n_k}{\sum_{j=1}^{K}n_j} \theta_k^{(t+1)}$$

where nkn_k is the local sample size.

FedAvg should be treated as a baseline, not automatically as the final federated algorithm (Zhao et al., 2018). Additional methods may be compared if non-IID heterogeneity substantially affects performance.

5. Uncertainty Layer

The final uncertainty method should be selected before experimentation and not changed after observing results.

A practical primary approach is (Xia et al., 2021):

Deep Ensembles

Train MM independently initialized models:

$$pm(y|x),m=1,\dots,Mp_m(y|x), \quad m=1,\dots,M$$

The mean prediction is:

$$p¯(y|x)=1M∑m=1Mpm(y|x)\bar{p}(y|x)= \frac{1}{M} \sum_{m=1}^{M}p_m(y|x)$$

Predictive disagreement can be used as an uncertainty signal (Schirmer et al., 2023).

Calibration will be evaluated using (Dani et al., 2026; Mehrtash et al., 2020):

- Expected Calibration Error;
- Brier Score;
- reliability diagrams.

The paper should avoid treating arbitrary probability intervals as clinically validated categories (initiative et al., 2019; Zheng et al., 2026). If High/Medium/Low labels are displayed, thresholds must be predefined and reported.

6. Explanation and LLM Layer

The explanation process is:

$$z= S(z, \mathcal{I}, \mathcal{E}) \text{explanation}$$

, where denotes the LLM-driven synthesis mapping the fused embeddings and influential features against a curated knowledge base of antibiotic resistance evidence .

The LLM must not modify:

- predicted class;
- predicted probability;
- uncertainty score;
- feature attribution values.

It only converts structured outputs and retrieved evidence into a readable explanation.

IV. MATERIALS AND METHODS

1. Study Design

This study uses publicly available AMR datasets and constructs a simulated decentralized environment (Kempter et al., 2026). No private patient records are required for the initial study.

The primary analysis will focus on predefined pathogen-antibiotic prediction tasks selected according to:

- sufficient sample size;
- reliable phenotype labels;
- biologically meaningful feature availability;
- class balance sufficient for evaluation.

2. Data Selection

Before implementation, each dataset should be documented with (Chindelevitch et al., 2023):

- source;
- organism;
- antimicrobial;
- phenotype definition;
- number of samples;
- feature type;
- preprocessing steps;
- missingness;
- license or access conditions.

Do not fabricate this table. Populate it only after final dataset selection.

Table 1. Dataset Description

Dataset	Organism	Drug	Samples	Genomic Data	Microbiome Data	Metadata
Dataset A	[Actual dataset name]	[Actual organism]	[Actual drug]	[Actual number]	Yes/No	Yes/No
Dataset B	[Actual dataset name]	[Actual organism]	[Actual drug]	[Actual number]	Yes/No	Yes/No

3. Preprocessing

The preprocessing pipeline will include:

- quality checking;
- removal of invalid or duplicate records where justified;
- phenotype label harmonization;
- feature normalization where required;
- categorical encoding;
- missing-value handling;
- prevention of train-test leakage.

All preprocessing parameters must be fitted only on training data and transferred to validation or test data (Παπουτσόγλου et al., 2023).

4. Simulated Federated Clients

Clients will be created using at least two scenarios.

Scenario A: Approximately IID partitioning

Samples are divided approximately randomly (Rodríguez-Barroso et al., 2020).

Scenario B: Non-IID partitioning

Clients differ deliberately in one or more characteristics (Lu et al., 2024).

- class prevalence;
- geographic source;
- data modality availability;
- feature distribution.

This enables explicit testing of robustness under heterogeneity (Chai et al., 2024).

5. Baseline Models

The study should compare:

Classical

- Logistic Regression
- Random Forest

Deep Learning
- Single-modality neural network
- Multimodal neural network

Training Strategies
- Local-only
- Centralized experimental upper-bound baseline
- Federated baseline
- Proposed uncertainty-aware federated approach

The centralized baseline is used only to understand the performance trade-off and does not represent the intended deployment architecture.

V. EVALUATION PROTOCOL

1. Predictive Performance
Evaluate:

Accuracy=TP+TNTP+TN+FP+FNAccuracy=\frac{TP+TN}{TP+TN+FP+FN}

Also report:
- Precision
- Recall
- F1-score
- AUROC
- AUPRC
- Sensitivity
- Specificity

AUPRC is particularly important when resistance classes are imbalanced (Saito & Rehmsmeier, 2015).

2. Calibration
Evaluate (Mehrtash et al., 2020):
- Brier Score;
- Expected Calibration Error;
- reliability diagrams.

3. Uncertainty Quality
The uncertainty system should be tested by determining whether uncertain cases are more likely to be incorrect (Dai et al., 2024).

Possible analyses include (Ding et al., 2020),(Zhou et al., 2024):
- error rate across uncertainty quantiles;
- risk-coverage curves;
- selective prediction performance;
- distribution-shift experiments.

4. Federated Robustness
Compare model performance across:
- IID clients;
- heterogeneous clients;
- varying client sample sizes;
- varying participation rates;
- unseen or held-out distributions where possible.

5. Explainability Evaluation
Explanations should not be judged solely by whether they "look reasonable." (Nauta et al., 2023)

Evaluate:
- stability under small perturbations;
- agreement with known resistance-associated features where appropriate;
- sparsity or concentration;
- expert assessment, if qualified experts are available.

Biological agreement should be described as supportive evidence rather than proof of causality.

6. LLM Evaluation
The LLM component should be evaluated separately from the retriever in a RAG system (Friel et al., 2024).

Measure
- factual consistency with retrieved evidence;
- citation support;
- completeness;
- unsupported-claim rate;
- human readability.

A useful failure analysis should classify outputs into:
- fully supported;
- partially supported;
- unsupported;
- contradicted by retrieved evidence.

VI. STATISTICAL ANALYSIS

Results should be reported across multiple random seeds or repeated experimental runs (Ferrer et al., 2024).

For each metric, report:
mean±standard deviation\text{mean} \pm \text{standard deviation}

Where appropriate, confidence intervals should also be provided.

Statistical comparisons must be selected based on the final experimental design. The manuscript should not claim significance merely because one model has a numerically higher score (Reimers & Gurevych, 2018).

VII. RESULTS

Important
This section must contain real experimental findings only. I will not invent results because that would make the paper scientifically unreliable.

Use the following structure after experiments.

1. Predictive Performance

Table 2. Comparison of predictive models

Model	AUROC	AUPRC	F1	Sensitivity	Specificity
Logistic Regression	Actual	Actual	Actual	Actual	Actual
Random Forest	Actual	Actual	Actual	Actual	Actual
Centralized DL	Actual	Actual	Actual	Actual	Actual
Standard FL	Actual	Actual	Actual	Actual	Actual
FedBioGuard	Actual	Actual	Actual	Actual	Actual

2. Calibration
Report:
- ECE;
- Brier score;
- reliability curve.

3. Heterogeneity Analysis
Compare IID and non-IID settings.

4. Uncertainty Analysis
Demonstrate whether the model assigns greater uncertainty to:
- incorrect predictions;
- ambiguous examples;
- shifted distributions.

5. Explainability Analysis
Show selected examples where feature attribution identifies influential features.

Do not claim a feature causes resistance unless independently supported by appropriate biological evidence.

6. LLM Grounding Analysis
Report the percentage of explanations that are:
- fully supported;
- partially supported;
- unsupported.

VIII. DISCUSSION

FedBioGuard is designed to address a central challenge in trustworthy AMR AI: high predictive performance alone is not sufficient for responsible deployment. A model may achieve strong discrimination while remaining poorly calibrated, unstable across institutions, or difficult to interpret (Chen et al., 2024).

The federated component investigates whether institutions can collaboratively improve models while retaining local records. However, federated learning should not be presented as a guarantee of complete privacy (Li et al., 2020). The results must be interpreted alongside the chosen threat model and any privacy-enhancing mechanisms.

The uncertainty component is important because it changes the system from a simple predictor into a model that can potentially distinguish between relatively reliable and ambiguous predictions. The central question is not whether every prediction can be labeled "confident," but whether uncertainty estimates meaningfully correspond to the probability of error (Bhatt et al., 2021).

Multimodal learning may improve prediction when different modalities contain complementary information. However, more modalities do not automatically produce better results. Performance may decrease when modalities are noisy, highly incomplete, or poorly aligned (Choi et al., 2025). The experiments should therefore determine whether multimodal integration produces measurable benefit.

The XAI component provides information about influential inputs, but feature attribution must be interpreted cautiously. An influential feature is not automatically a causal mechanism (Blöbaum & Shimizu, 2017). The system's explanations should instead be considered hypothesis-supporting tools.

The evidence-grounded LLM is intentionally separated from the AMR classifier. This architecture reduces the risk that generative text directly alters the predictive decision. Its purpose is to assist interpretation and literature synthesis, not to replace microbiological expertise or susceptibility testing (Schwartz et al., 2023).

Limitations

This study has several anticipated limitations.

- First, simulated federated environments cannot fully reproduce real institutional differences (Chen et al., 2026).
- Second, public datasets may contain biases related to geography and sampling practices (Gangavarapu et al., 2023).
- Third, the current evaluation lacks integration with privacy-preserving techniques such as differential privacy or homomorphic encryption, and future work should explore their practical use (Peng et al., 2024).
- Fourth, federated learning reduces raw-dat (Ovadia et al., 2019)a movement but does not by itself eliminate every privacy risk.
- Fifth, uncertainty estimates depend on the selected methodology and may not transfer perfectly to all forms of distribution shift (Ovadia et al., 2019).
- Sixth, XAI methods describe model behavior rather than establishing biological causality (Blöbaum & Shimizu, 2017).
- Finally, the LLM may still produce unsupported statements despite retrieval grounding (Amugongo et al., 2025). Therefore, automated explanations should be evaluated and should not be treated as independent clinical evidence (Gambetti et al., 2025).

IX. CONCLUSION

This paper presents FedBioGuard, a research framework for investigating trustworthy antimicrobial resistance prediction through the integration of multimodal learning, federated optimization, uncertainty quantification, explainable AI, and evidence-grounded language-model assistance. The central contribution is the proposed evaluation of whether uncertainty-aware multimodal federated learning can provide reliable AMR prediction under heterogeneous decentralized conditions.

Rather than treating privacy, accuracy, uncertainty, and interpretability as independent requirements, FedBioGuard evaluates them as interacting properties of a single analytical system. The framework also separates predictive modeling from generative explanation so that the LLM functions as an evidence-grounded communication layer rather than a component that determines the resistance prediction.

The study does not propose replacing established laboratory diagnostics or clinical judgment. Its purpose is to develop and evaluate computational methods that may support future research into privacy-conscious, transparent, and reliability-aware AMR analytics. Future work should extend the framework to real multi-institutional collaborations, formal privacy attacks and defenses, prospective validation, and expert evaluation of explanations.

REFERENCES

1. World Health Organization, Global Research Agenda for Antimicrobial Resistance in Human Health. Geneva, Switzerland: WHO, 2025.
2. T. Bhardwaj and K. Sumangali, “Secure healthcare data management using federated learning, blockchain, and explainable artificial intelligence: A systematic review,” Frontiers in Digital Health, vol. 8, 2026, Art. no. 1871960.
3. M. Mohammadi, M. Vejdanihemmat, M. Lotfinia, M. Rusu, D. Truhn, A. Maier, and S. Tayebi Arasteh, “Differential privacy for medical deep learning: Methods, tradeoffs, and deployment implications,” npj Digital Medicine, vol. 9, 2026, Art. no. 93.

4. "Privacy preservation for federated learning in health care," Patterns, vol. 5, no. 7, 2024, Art. no. 100974.
5. "Federated learning in healthcare: A comprehensive survey on privacy, scalability and clinical applications," ICT Express, vol. 12, no. 4, pp. 825–839, 2026.
6. "A survey: Federated learning in healthcare," Computers & Electrical Engineering, vol. 135, 2026, Art. no. 111195.
7. "Federated learning in healthcare: Recent progress and challenges," Computers & Electrical Engineering, vol. 131, 2026, Art. no. 110924.
8. "Federated Learning in Healthcare: From Research to Real-World Deployment," Annual Review of Biomedical Engineering, 2026.
9. T. Bhardwaj and K. Sumangali, "Secure healthcare data management using federated learning, blockchain, and explainable artificial intelligence: A systematic review," Frontiers in Digital Health, vol. 8, 2026, Art. no. 1871960.
10. "Privacy preservation for federated learning in health care," Patterns, 2024. This work is particularly useful for discussing the fact that federated model updates can still create information-leakage risks and therefore require additional privacy mechanisms.