



CONFERENCE

—NORTH AMERICA 2026—

Privacy-Preserving AI Across Institutions

Federated Learning Lessons for Medical Imaging
Sudhanshu Shrivastava
ECE graduate student, UC Davis

Public-data federation for retinal segmentation

5,550

fundus photographs

9 public datasets

7 countries

Optic disc and optic cup segmentation

Retrospective benchmark using public images

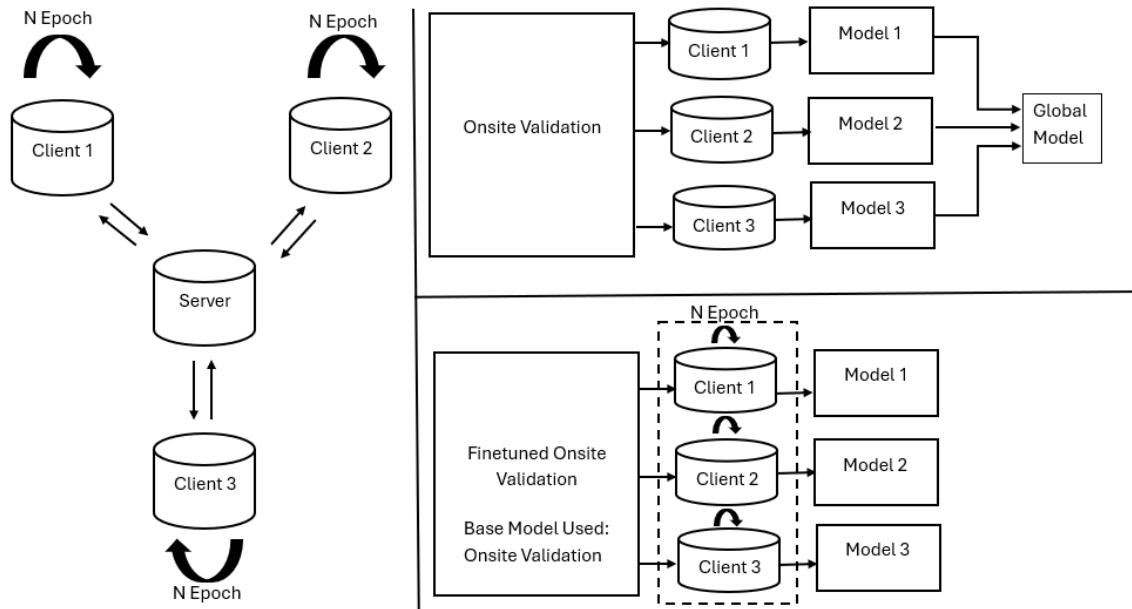
Study results come from a preprint

Model comparisons

Approach	Training or adaptation
Local model	One dataset
Pooled central model	All training datasets together
Global Validation	Federated averaging
Weighted Global Validation	Sample-weighted aggregation
Onsite Validation	Local training with early stopping
Fine-tuned Onsite Validation	Federation, then adaptation at each site

Personalization produces a model for each participating site

Federated training and local fine-tuning



Original study, Figure 1(ii)

04

Personalized models versus pooled training

Fine-tuned Onsite Validation versus the pooled central model

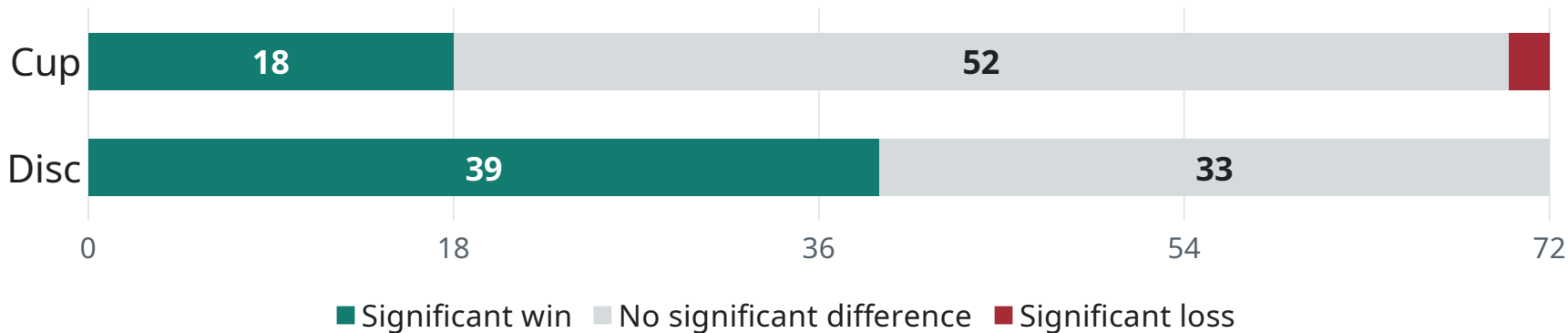
Task	No significant difference	Significant win / loss
Cup	9 of 9 sites	0 wins / 0 losses
Disc	7 of 9 sites	1 win / 1 loss

Reported Dice comparisons with multiple-testing correction

No significant difference does not establish equivalence

Cross-site tests favor disc personalization

Fine-tuned federated model versus a model trained only at its source site



72 directed comparisons per task across the nine datasets

All target datasets also participate in federated training

Training and evaluation

Local training

Each site trains locally and sends model weights to the server

Onsite early stopping

Local validation loss controls training over multiple epochs

Site-specific fine-tuning

The final federated model is adapted separately at nine sites

Per-site evaluation

Disc and cup Dice scores are evaluated on each site's test set

Methods, pp. 8–11. Global validation/test-set separation needs clarification (p. 11).

Local adaptation deserves its own evaluation

A shared model can be a starting point for a site-specific model

Disc and cup results require separate evaluation

Raw data staying local does not guarantee privacy

Next evidence: unseen institutions and an explicit threat model

Preprint DOI: 10.64898/2025.12.02.25341485

Gains over standard federated variants

Task	Significant wins	Share of comparisons
Disc	14 of 27	About 52%
Cup	7 of 27	About 26%

9 personalization sites × 3 federated comparators = 27

These percentages count wins across comparisons

Collaborators and study source

Sudhanshu Shrivastava¹ Upasana Thakuria¹
Scott Kinder¹ Giacomo Nebbia¹ Nazlee Zebardast²
Sally L. Baxter^{3,4} Benjamin Xu⁵ Saif Aldeen Alryalat¹
Malik Kahook¹ Jayashree Kalpathy-Cramer¹
Praveer Singh¹

¹ Ophthalmology, University of Colorado Anschutz Medical Campus

² Massachusetts Eye and Ear Infirmary, Harvard Medical School

³ Ophthalmology Informatics and Data Science, UC San Diego

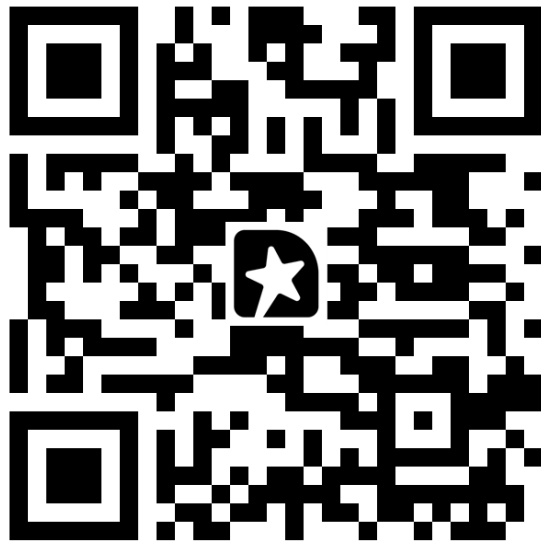
⁴ Biomedical Informatics, Department of Medicine, UC San Diego

⁵ Roski Eye Institute, Keck School of Medicine, University of Southern California

medRxiv preprint, December 4, 2025

DOI: 10.64898/2025.12.02.25341485

Thank you



Privacy-Preserving AI Across Institutions: Federated Learning
Lessons for Medical Imaging