Benchmarking **federated learning** approaches against **siloed** and **mega-analysis** regimes

Michelle Wang
PhD candidate, NeuroDataScience-ORIGAMI Lab (JB Poline)

2025 June 23

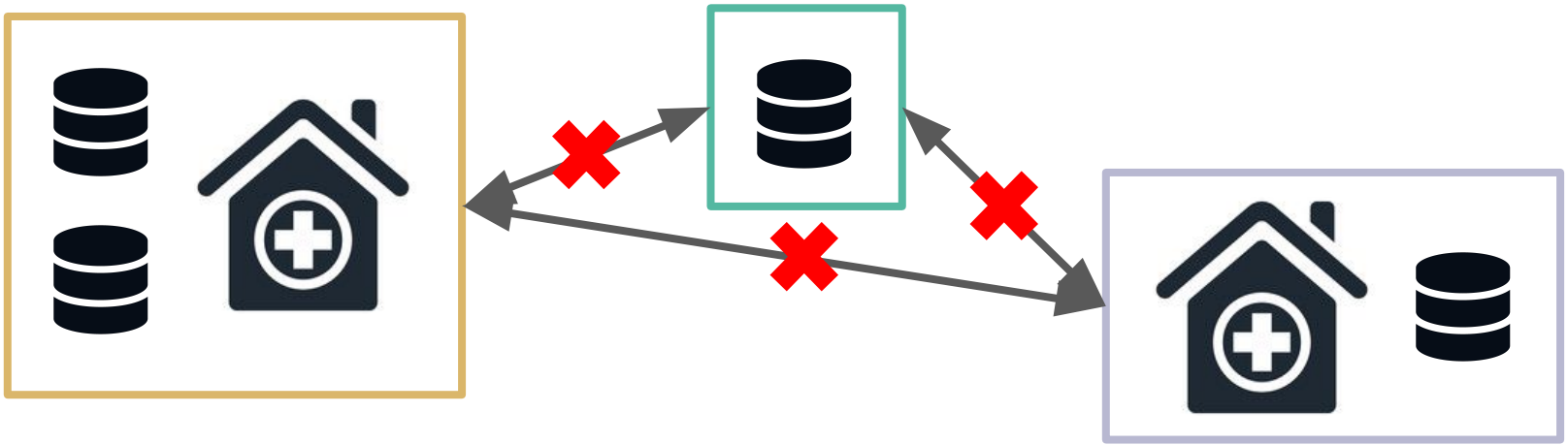
ORIGAMI
Lab

Data-sharing in studies involving clinical populations

Not all data can be openly shared

- Due to strong data privacy frameworks (e.g., the European Union's General Data Protection Regulation [GDPR]), etc.

Reality: Despite promotion of open science and data-sharing practices, **many datasets still live in so-called "silos"**



Decentralized methods for multi-centric analyses

Data **processing**

- Share software pipelines (including versions and runtime parameters)

Nipoppy framework for organization and decentralized processing of neuroimaging-clinical datasets

Data **discovery**

- Share information about availability of different datatypes/measures

Neurobagel ecosystem for decentralized metadata harmonization and search

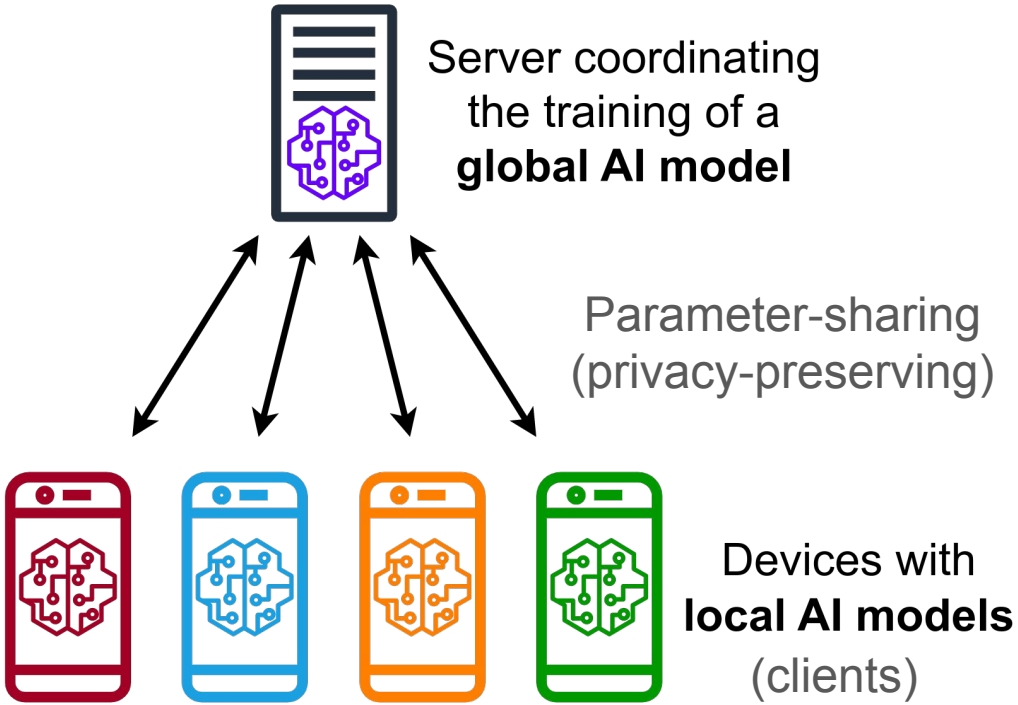
Data **analysis**

- Share fitted model parameters

Federated learning (McMahan *et al.*, 2016)

Federated learning

- Decentralized data
- Sharing of fitted model parameters to create a **global model**



We compared a simple **federated learning** setup with more traditional setups in clinical neuroimaging analyses

Data

Parkinson's/Alzheimer's disease (PD/AD) datasets:

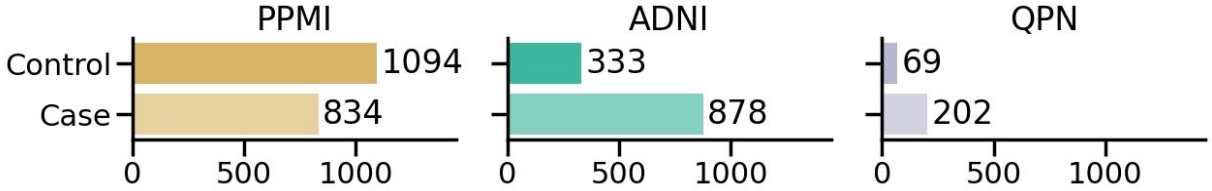
- Parkinson's Progression Markers Initative (**PPMI**) (Marek et al., 2018)
- Alzheimer's Disease Neuroimaging Initative (**ADNI**) (Jack et al., 2008)
- Quebec Parkinson Network (**QPN**) (Gan-Or et al., 2020)

Measures

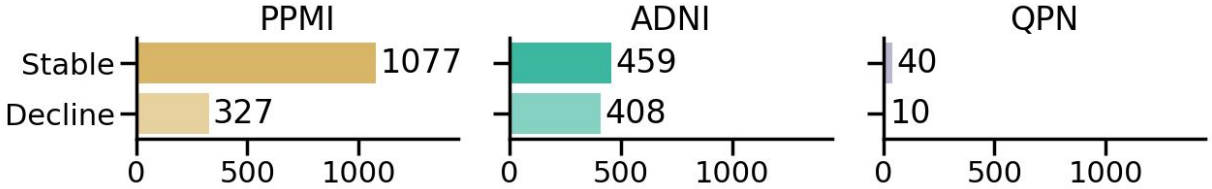
- Age
- Sex
- FreeSurfer cortical thicknesses (**CT**) and subcortical volumes (**SV**)
- Cognitive assessment score
  - PPMI and QPN: Montreal Cognitive Assessment (MoCA)
  - ADNI: Mini-Mental State Examination (MMSE)

Prediction tasks

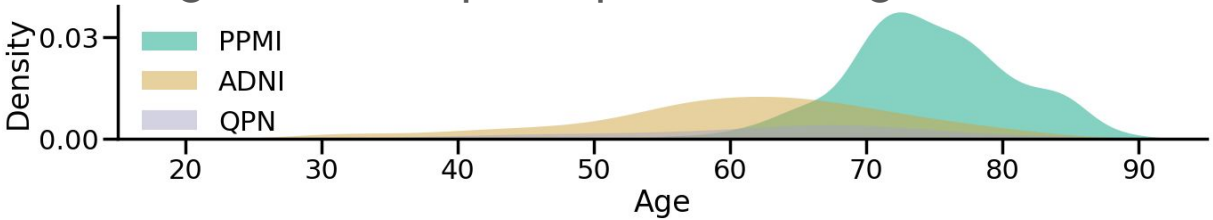
Predict **diagnosis** (PD/AD vs control) from **age + sex + CT + SV**



Predict **cognitive decline** in patients from **age + sex + CT**



Predict **age** of control participants from **age + sex + SV**



Model training setups

Setup	Train data	Model
Siloed (PPMI)	PPMI →	$M_{siloed,\,ppmi}$
Siloed (ADNI)	ADNI →	$M_{siloed,\,adni}$
Siloed (QPN)	QPN →	$M_{siloed,\,qpn}$
Federated	PPMI → M_{ppmi} ADNI → M_{adni} QPN → M_{qpn}	$M_{federated}$ (weighted avg. of params.)
Mega-analysis	PPMI ADNI QPN	$M_{mega-analysis}$

Base model pipeline:

- Z-score
- Logistic/linear regression

10-fold cross-validation

Test datasets:

- PPMI
- ADNI
- QPN
- PPMI+ADNI+QPN

Model training setups

Setup	Train data	Model
Siloed (PPMI)	PPMI →	$M_{siloed,\,ppmi}$
Siloed (ADNI)	ADNI →	$M_{siloed,\,adni}$
Siloed (QPN)		
Federated		
Mega-analysis	PPMI ADNI QPN →	$M_{mega-analysis}$

Hypothesis

Siloed < Federated ≤ Mega-analysis

Base model pipeline:

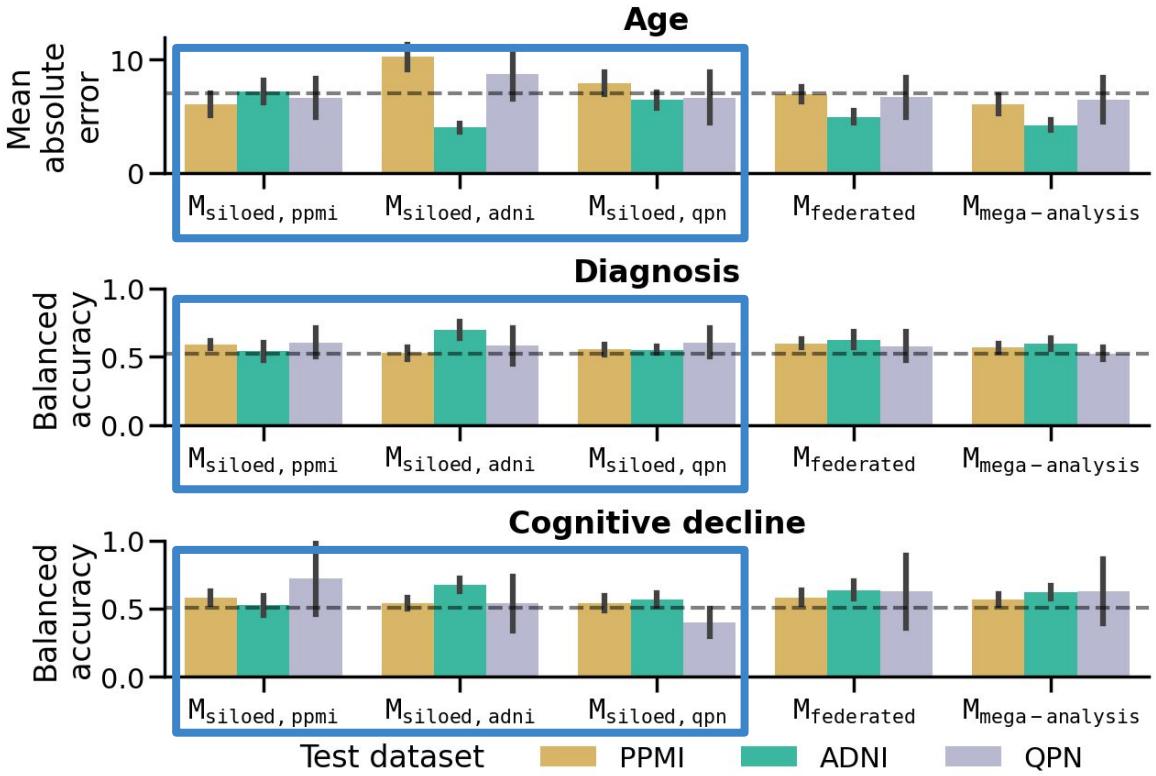
- Z-score
- Logistic/linear regression

-fold cross-validation

st datasets:

- PPMI
- ADNI
- QPN
- PPMI+ADNI+QPN

Preliminary results

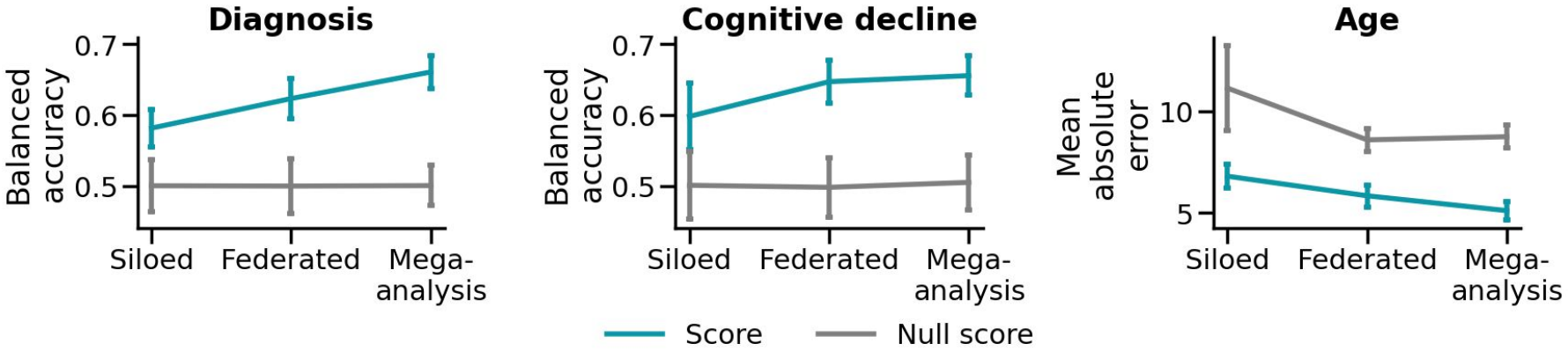


Effects of **dataset shift**:

- Poor generalizability of models trained with **Siloed** setup

Preliminary results

Test performance scores on **PPMI+ADNI+QPN** show **improvement from Siloed to Federated setup** in all three prediction tasks



Discussion

- Preliminary results suggest that, for the datasets and use-cases investigated, **federated setup shows generalizable and comparable performance to the mega-analysis approach**
  - But does not require sharing of sensitive clinical data
- Neuroinformatics infrastructure is needed for handling the practical aspects of this type of analysis

Nipoppy framework for organization and decentralized processing of neuroimaging-clinical datasets

Neurobagel ecosystem for decentralized metadata harmonization and search

Fed-BioMed framework for federated learning

Future!

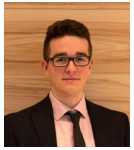
Thank you!

- Contact: michelle.wang6@mail.mcgill.ca
- OHBM poster **#1146** (June 27-28, 13:45-15:45)


Nikhil Bhagwat


Alyssa Dai


Sebastian Urchs


Mathieu Dugré


Arman Jahanpour


Jean-Baptiste Poline

NeuroDataScience-ORIGAMI Lab

ORIGAMI Lab · neuro Montreal Neurological Institute-Hospital · CIHR IRSC Canadian Institutes of Health Research / Instituts de recherche en santé du Canada · Québec Fonds de recherche – Nature et technologies / Fonds de recherche – Santé / Fonds de recherche – Société et culture · PARKINSON QUÉBEC · UNIQUE CENTRE · NIH National Institutes of Health · THE MICHAEL J. FOX FOUNDATION FOR PARKINSON'S RESEARCH · Parkinson's Progression Markers Initiative

McGill UNIVERSITY · Chan Zuckerberg Initiative · HBHL NEUROHUB · QUEBEC PARKINSON NETWORK · Fondation Brain Canada Foundation · ADNI Alzheimer's Disease Neuroimaging Initiative