

Methods Innovation Team

Team Meeting

November 15, 2022



Methods Innovation Team Meeting Participants

Name	Organization	
Fotios Kokkotos	NEWDIGS	Methods Innovation Team Lead & Meeting Facilitator
Keileen Hopps	NEWDIGS	Scribe
Paul Beninger	Tufts	
John Ferguson	Sanofi	
Smita Kothari	Merck	
Jiaching Lin	Takeda	
Peter Loupos	Princeton Healthcare Strategies, LLC	
Lynn Sanders	Takeda	
Asvin Srinivasan	Onc.AI	
Wenting Wang	Sanofi	
Chunlei Zheng	VA Boston Healthcare System	



Team Meeting & Report Out Timing

1:45-3:15 pm Team Meeting

- 10 minutes – Introductions and presenter selection
- 70 minutes – Discussion
- 10 minutes – Presentation coordination

Notes and instructions

- Each team will need to pick a presenter for the Report Out session that follows the team meeting session. **Please do this first!**
- Each team meeting will be facilitated by the Team Lead and have a pre-selected scribe(s).
- At 3:15 pm, please send the presentation to the email hyperlinked on the last slide.

3:15-3:30 pm Break

3:30-4:30 pm Team Reports and Group Discussions

Per team:

- 10 minutes – report out
- 5 minutes - Q&A



Agenda

- ❑ Discussion
 - Clinical Practice Gaps
 - Data Exclusion Cycle
 - Growing Library of Risk Bias Tools
 - Metrics under consideration
 - **Clinical Outcomes Metrics**
 - Median Overall Survival (OS)
 - Time to treatment failure
 - Median Progression Free Survival (PFS)
 - **Impact Measures**
 - Time to treatment/"effective treatment"
 - Total cost of care
- ❑ Review & Refine
 - Data Characterization and Technical Skills Framework
 - Machine Learning & Statistical Methods Outline
 - Federated Machine Learning Process and FL Network Outline
- ❑ Future Action
 - Timeline? Deliverables? Etc.



Clinical Practice Gaps for aNSCLC

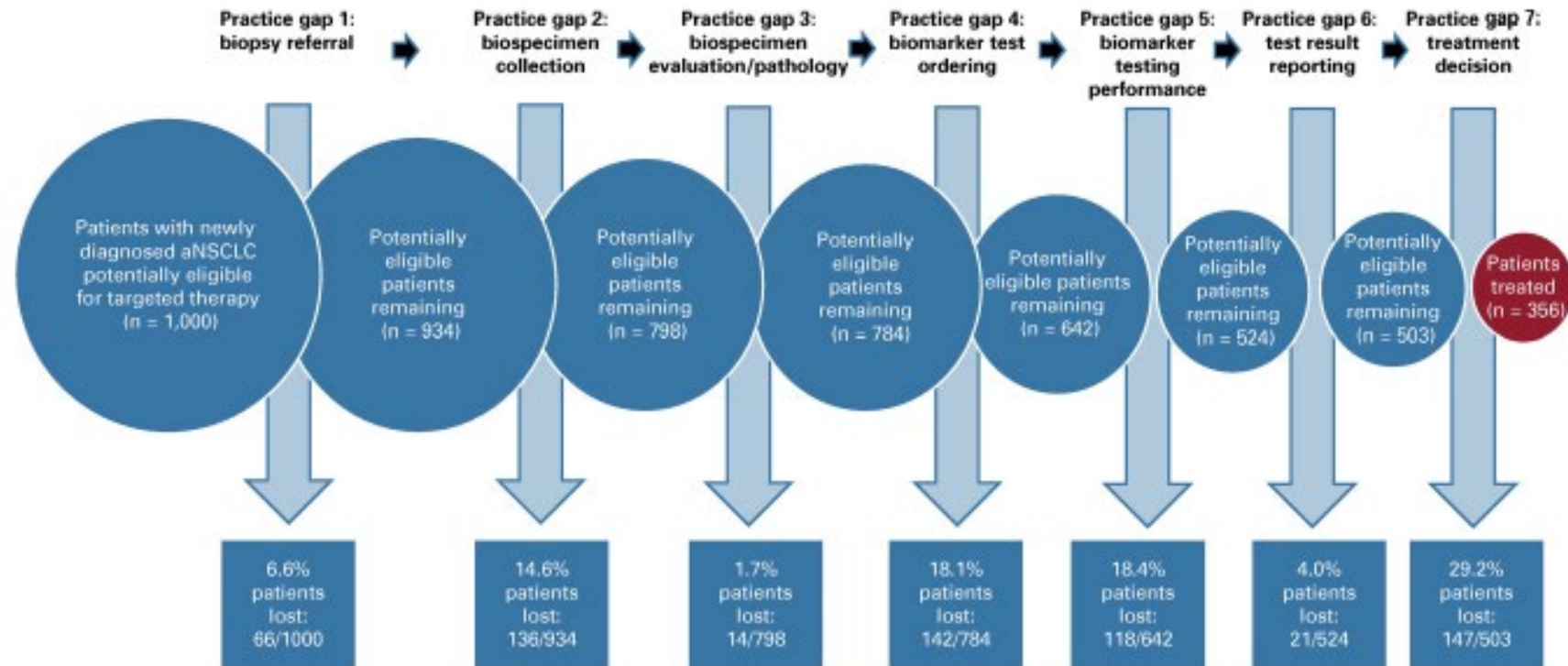


Illustration from Sadik, H., et al. Impact of Clinical Practice Gaps on the Implementation of Personalized Medicine in Advanced Non-Small-Cell Lung Cancer. *JCO Precision Oncology* 2022:6



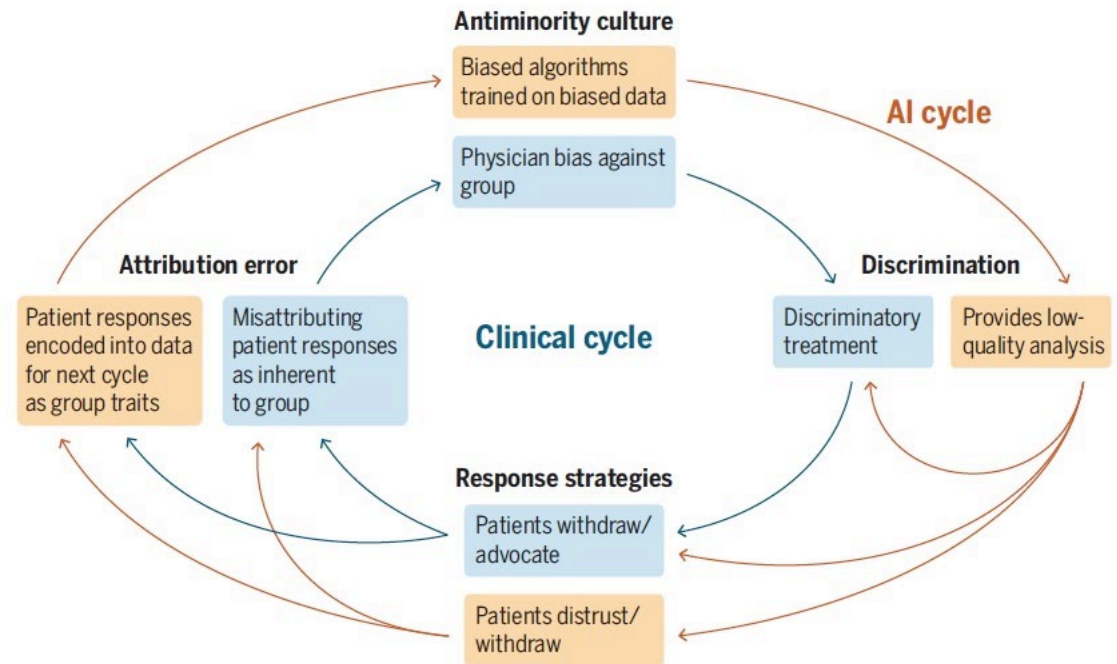
Data Exclusion Cycles & Health Disparities Biases

How can we expand what the “real world” is viewed as being?

- Ex.: Payers using their data to develop population-level statistics and analysis that frame the overall population within their own narrow confines (e.g., higher-income level individuals with private healthcare coverage, etc.)
- We’re focusing on minimizing the health disparities biases inherent in many current large data sets, e.g., Medicare

Medical practice and AI create overlapping exclusion cycles

Around a generalized structure of the four-step cycle, we depict two interacting exclusion cycles: clinical encounters (blue) and artificial intelligence (AI) products that result from, and influence, data surrounding those clinical encounters (orange). Each cycle self-perpetuates, but the cycles also interact at various points.





Growing Library of *Risk of Bias* Tools

Riskofbias.info

RoB 2 tool (revised tool for Risk of Bias in randomized trials)

ROBINS-E tool (Risk Of Bias in non-randomized Studies - of Exposures)

ROB ME (Risk Of Bias due to Missing Evidence in a synthesis)

ROBINS-I tool (Risk Of Bias in Non-randomized Studies - of Interventions)

robvis (visualization tool for risk of bias assessments in a systematic review)

PROBAST: Prediction model Risk of Bias Assessment Tool

Potential LEAPS Opportunity

Data Assessment and Risk Engineering (DARE Framework)

- Question: Can the diversity of data sets strategically selected for a Federated Learning study (and platform) reduce the risk of bias related to under-represented patient sub-populations?



Metrics under consideration

Clinical Outcomes Metrics

- Median Overall Survival (OS)
- Time to treatment failure (TTF)
- Median Progression Free Survival (PFS)

Impact Measures

- Time to treatment/"effective treatment"
- Total cost of care



Review & Refine

- Data Characterization and Technical Skills Framework
- Machine Learning & Statistical Methods Outline
- Federated Machine Learning Process and FL Network Outline

TuftsMedicine
Tufts Medical Center
NEW DIGS

Data & Technical Skills Characterization Framework

LEAPS Methods Innovation Team

Case Study Background
For the initial advanced NSCLC case study, develop a predictive model(s) that can improve decision-making for all stakeholders related to immune checkpoint inhibitor use in patients with advanced NSCLC.

Objectives
Assess feasibility of using federated (machine) learning methods, leveraging diverse data types (e.g., EHR, administrative claims, social determinants of health, biologic, clinical trials, patient-generated, etc.) to:

- Identify signals, generate hypotheses about clinically meaningful sub-populations
- Define next step in corroborating/validating promising hypotheses
- Reduce bias in algorithm development through the use of diverse data sets
- Establish federated learning environment (technology enablers, cross-functional expertise, governance) that is scalable

Purpose of Data & Technical Skills Framework

- Develop list of organizations with data & their needed for the successful completion of a proof of concept
- Define the data privacy and security framework concerns of the data providers
- Develop & apply a framework for characterizing data sources in order to identify the initial set of data sources

Framework Application and Approach
The Data & Technical Skills Characterization Framework begins with the application of the NEWDIGS 5-L in the search of the data sources under consideration. A Predictive Outcomes Platform (POP). A primary strategic inclusivity of data sources.

The Framework seeks to identify and determine data sources and needed technical skills to the POP specifically. In addition to capturing general data

TuftsMedicine
Tufts Medical Center
NEW DIGS

Outline of Machine Learning & Statistical Methods

LEAPS Methods Innovation Team

Case Study Background
For the initial case study, develop a predictive model(s) that can improve decision-making for all stakeholders related to immune checkpoint inhibitor use in patients with advanced NSCLC.

Objectives
Assess feasibility of using federated (machine) learning methods, leveraging diverse data types (e.g., EHR, claims, social determinants of health, biologic, clinical trials, patient-generated, etc.) to:

- Identify signals, generate hypotheses about clinically meaningful sub-populations
- Define next step in corroborating/validating promising hypotheses
- Reduce bias in algorithm development through the use of diverse data sets
- Establish federated learning environment (technology enablers, cross-functional expertise, governance) that is scalable

Purpose of the Machine Learning & Statistical Methods Outline

- Discuss a list of machine learning and statistical methods to be applied to a series of case studies within LEAPS
- Identify the right machine learning/statistical models to fit the response variables with emphasis on full transparency
- Develop a framework for assessing and validating the strengths and limitations of available machine learning and statistical methods

Application and Approach
The Machine Learning (ML) & Statistical Methods Outline (Outline) and characterize the strengths and limitations of available machine learning and statistical methods to be applied to the Advanced NSCLC Use Case more generally to other use cases as identified by the LEAPS team. Capturing general details, the strengths and limitations of the ML methods are characterized and assessed across multiple objectives.

TuftsMedicine
Tufts Medical Center
NEW DIGS

Federated Learning Network Outline

LEAPS Methods Innovation Team

Case Study Background
For the initial case study, develop a predictive model(s) that can improve decision-making for all stakeholders related to immune checkpoint inhibitor use in patients with advanced NSCLC.

Objectives
Validate the potential of using federated (machine) learning methods, leveraging diverse data types (e.g., EHR, claims, social determinants of health, biologic, clinical trials, patient-generated, etc.) to:

- Identify signals, generate hypotheses about clinically meaningful sub-populations
- Define next step in corroborating/validating promising hypotheses
- Reduce bias in algorithm development through the use of diverse data sets
- Establish federated learning environment (technology enablers, cross-functional expertise, governance) that is scalable

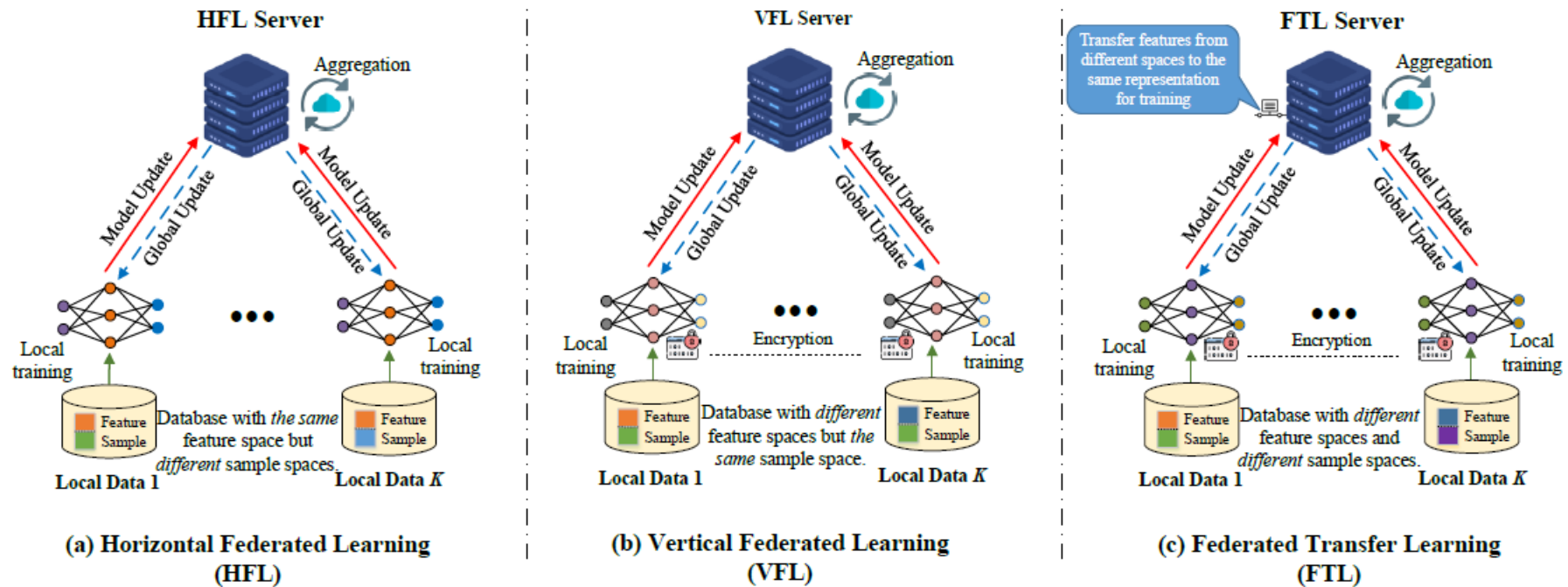
Purpose of the Federated Learning Network Outline

- Discuss the different types & platforms of federated (machine) learning (FL) to assist in the selection of the most appropriate types for initial prototyping
- Explore & define the right structure of the centralized server for successful implementation of a federated learning network prototype for Case Study #1 (immune-checkpoint inhibitors in advanced NSCLC)
- Demonstrate how Federated Learning expands the understanding of the impact of diseases in the broader population beyond structured information from EHR and insurance databases and addresses healthcare disparities by utilizing non-traditional data sources including patient generated information and community initiatives supporting the underserved.

Application and Approach
The Federated Learning Network Outline seeks to identify and characterize the strengths and limitations of available FL types and platforms that could be selected for application to the Advanced NSCLC Use Case specifically and more generally to other use cases as identified by the LEAPS team. In addition to providing a general



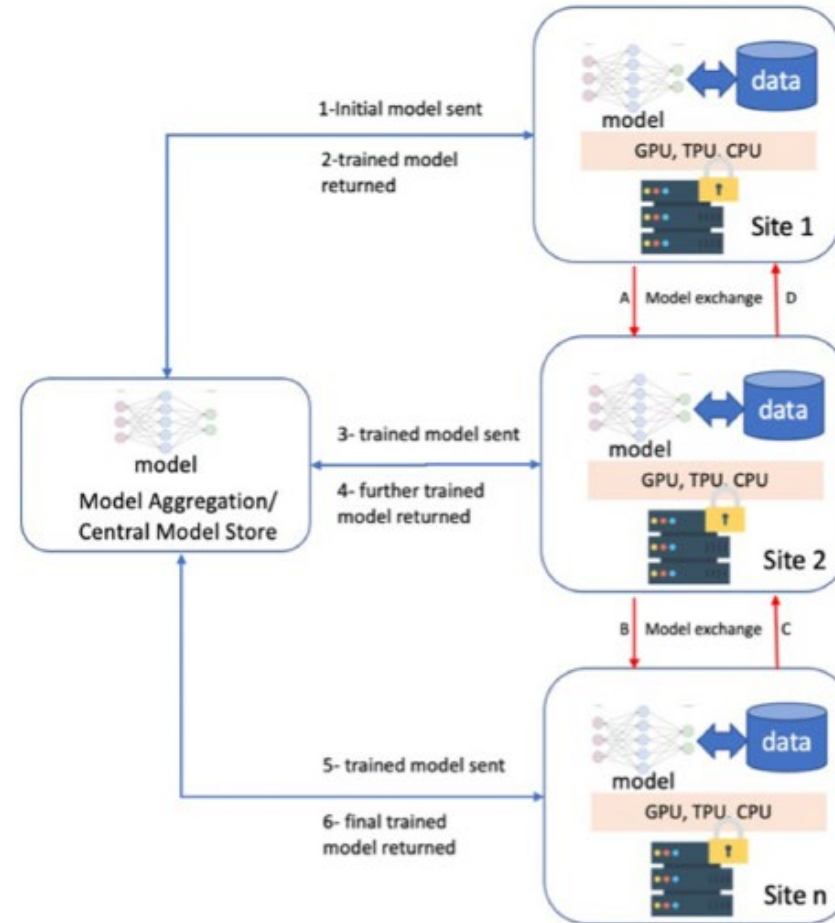
Categories of Federated Learning



Categories of FL used in smart healthcare



Federated Learning Overview



Federated Learning Overview and two different approaches; 1-6 model exchange via a centralized model store, A-D direct model exchange among the participating sites.



Future Action for Methods Team

November 2022-June 2023

TuftsMedicine
Tufts Medical Center



tuftsmedicine.org