# Federated Causal Inference: Multi-Source ATE Estimation beyond Meta-Analysis

Rémi Khellaf

Supervised by Aurélien Bellet & Julie Josse

November 27, 2024

PreMeDICal team, INRIA, INSERM, Université de Montpellier

# Motivation

# Federated Causal Inference beyond Meta-Analysis

- Goal of causal inference: measure the (average) impact of a treatment on an outcome

# Federated Causal Inference beyond Meta-Analysis

- Goal of causal inference: measure the (average) impact of a treatment on an outcome
  - Medical studies: impact of vaccination on Covid-19 outbreaks? (Moghadas et al., 2021)

# Federated Causal Inference beyond Meta-Analysis

- Goal of causal inference: measure the (average) impact of a treatment on an outcome
  - Medical studies: impact of vaccination on Covid-19 outbreaks? (Moghadas et al., 2021)
  - Economics and social sciences: impact of studies on future earnings in developing countries? (Duflo, Glennerster, and Kremer, 2007)

# Federated Causal Inference beyond Meta-Analysis

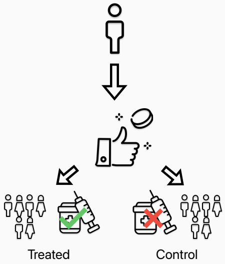
- Goal of causal inference: measure the (average) impact of a treatment on an outcome
    - Medical studies: impact of vaccination on Covid-19 outbreaks? (Moghadas et al., 2021)
    - Economics and social sciences: impact of studies on future earnings in developing countries? (Duflo, Glennerster, and Kremer, 2007)
    - Public Health & Economy: evaluating drugs efficacy. French social security reimburses drugs based on their proven efficacy. (French Health Authority, 2024)

# Federated Causal Inference beyond Meta-Analysis

- Goal of causal inference: measure the (average) impact of a treatment on an outcome
- Experimental method: randomized clinical trials (RCTs) are the gold standard

# Federated Causal Inference beyond Meta-Analysis

- Goal of causal inference: measure the (average) impact of a treatment on an outcome
- Experimental method: randomized clinical trials (RCTs) are the gold standard
  - Randomization of treatment assignment disentangles correlation from causation: differences in outcomes are attributable to the treatment only rather than confounding factors



## Federated Causal Inference beyond Meta-Analysis

- Goal of causal inference: measure the (average) impact of a treatment on an outcome
- Experimental method: randomized clinical trials (RCTs) are the gold standard
  - Randomization of treatment assignment disentangles correlation from causation: differences in outcomes are attributable to the treatment only rather than confounding factors
  - Limits: expensive, not always feasible, stringent eligibility criteria, short timeframes, small sample sizes

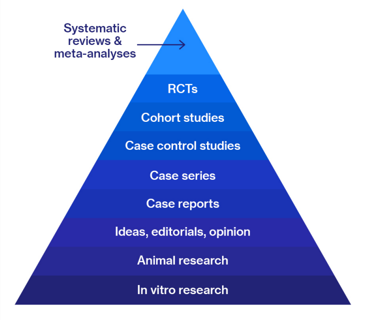
# Federated Causal Inference beyond Meta-Analysis

- Goal of causal inference: measure the (average) impact of a treatment on an outcome
- Experimental method: randomized clinical trials (RCTs) are the gold standard
  - Randomization of treatment assignment disentangles correlation from causation: differences in outcomes are attributable to the treatment only rather than confounding factors
  - Limits: expensive, not always feasible, stringent eligibility criteria, short timeframes, small sample sizes

  $\rightarrow$ Multisource approach: several RCTs are better than 1!

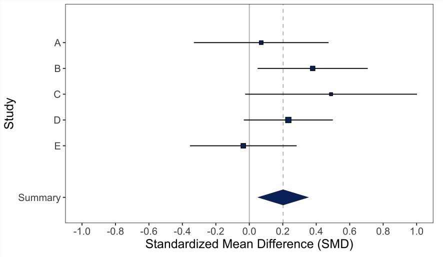
# Federated Causal Inference beyond Meta-Analysis

- Goal of causal inference: measure the (average) impact of a treatment on an outcome
- Experimental method: randomized clinical trials (RCTs) are the gold standard
- Multisource framework: meta-analysis on summary statistics as the top of the pyramid of evidence



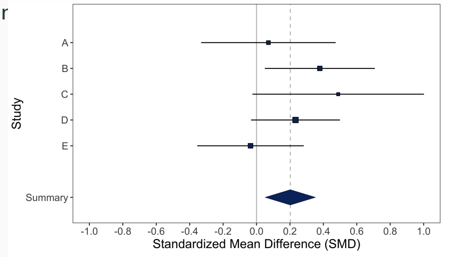
# Federated Causal Inference beyond Meta-Analysis

- Goal of causal inference: measure the (average) impact of a treatment on an outcome
- Experimental method: randomized clinical trials (RCTs) are the gold standard
- Multisource framework: meta-analysis on summary statistics as the top of the pyramid of evidence
    - Aggregates estimated causal effects from published summary statistics of multiple studies



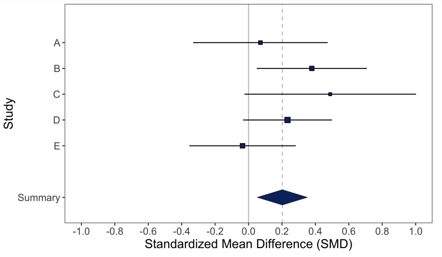
## Federated Causal Inference beyond Meta-Analysis

- Goal of causal inference: measure the (average) impact of a treatment on an outcome
- Experimental method: randomized clinical trials (RCTs) are the gold standard
- Multisource framework: meta-analysis on summary statistics as the top of the pyramid of evidence
  - Aggregates estimated causal effects from published summary statistics of multiple studies
  - Increased statistical power and more precise estir



# Federated Causal Inference beyond Meta-Analysis

- Goal of causal inference: measure the (average) impact of a treatment on an outcome
- Experimental method: randomized clinical trials (RCTs) are the gold standard
- Multisource framework: meta-analysis on summary statistics as the top of the pyramid of evidence
  - Aggregates estimated causal effects from published summary statistics of multiple studies
  - Increased statistical power and more precise estir
  - Keeps the data decentralized

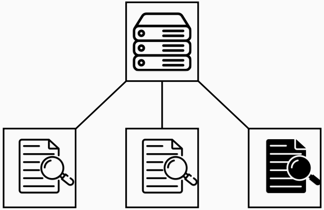


# Federated Causal Inference beyond Meta-Analysis

- Goal of causal inference: measure the (average) impact of a treatment on an outcome
- Experimental method: randomized clinical trials (RCTs) are the gold standard
- Multisource framework: meta-analysis on summary statistics as the top of the pyramid of evidence
    - Aggregates estimated causal effects from published summary statistics of multiple studies
    - Increased statistical power and more precise estimates
    - Keeps the data decentralized
    - Limits: no direct access of individual observations $\Longrightarrow$ face important challenges in presence of heterogeneity between the studies

# Federated Causal Inference beyond Meta-Analysis

- Goal of causal inference: measure the (average) impact of a treatment on an outcome
- Experimental method: randomized clinical trials (RCTs) are the gold standard
- Multisource framework: meta-analysis on summary statistics as the top of the pyramid of evidence
- Our contribution: leverage federated learning (FL) to perform causal inference on decentralized invididual data



# Federated Causal Inference beyond Meta-Analysis

- Goal of causal inference: measure the (average) impact of a treatment on an outcome
- Experimental method: randomized clinical trials (RCTs) are the gold standard
- Multisource framework: meta-analysis on summary statistics as the top of the pyramid of evidence
- Our contribution: leverage federated learning (FL) to perform causal inference on decentralized invidividual data
    - Propose the Gradient Descent Federated estimator of the Average Treatment Effect (ATE)

# Federated Causal Inference beyond Meta-Analysis

- Goal of causal inference: measure the (average) impact of a treatment on an outcome
- Experimental method: randomized clinical trials (RCTs) are the gold standard
- Multisource framework: meta-analysis on summary statistics as the top of the pyramid of evidence
- Our contribution: leverage federated learning (FL) to perform causal inference on decentralized invidividual data
  - Propose the Gradient Descent Federated estimator of the Average Treatment Effect (ATE)
  - Study and compare (bias and asymptotic variances) Meta-Analysis, One-Shot federated and Gradient Descent federated estimators of ATE

# Federated Causal Inference beyond Meta-Analysis

- Goal of causal inference: measure the (average) impact of a treatment on an outcome
- Experimental method: randomized clinical trials (RCTs) are the gold standard
- Multisource framework: meta-analysis on summary statistics as the top of the pyramid of evidence
- Our contribution: leverage federated learning (FL) to perform causal inference on decentralized invidividual data
  - Propose the Gradient Descent Federated estimator of the Average Treatment Effect (ATE)
  - Study and compare (bias and asymptotic variances) Meta-Analysis, One-Shot federated and Gradient Descent federated estimators of ATE
  - Compare their robustness to several heterogeneity scenarios

# Introduction to Causal Inference

# Usual Causal Framework - RCT

- Causal inference goal: estimate the (average) impact of **treatment** $W$ on $Y$, given $X$ describing a population.

# Usual Causal Framework - RCT

- Causal inference goal: estimate the (average) impact of **treatment** $W$ on $Y$, given $X$ describing a population.
- Data is centralized:

| Obs. | Covariates | | | Treatment | Outcome | Potential Outcomes | |
|---|---|---|---|---|---|---|---|
| $i$ | $X_1$ | $X_2$ | $X_3$ | $W$ | $Y$ | $Y^{(1)}$ | $Y^{(0)}$ |
| 1 | 2.3 | 1.5 | M | 1 | 3.2 | 3.2 | ?? |
| 2 | 2.2 | 3.1 | F | 0 | 2.8 | ?? | 2.8 |
| 3 | 3.5 | 2.0 | F | 1 | 2.1 | 2.1 | ?? |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $n-1$ | 3.7 | 2.0 | F | 0 | 2.8 | ?? | 2.8 |
| $n$ | 2.5 | 1.7 | M | 1 | 3.2 | 3.2 | ?? |

# Usual Causal Framework - RCT

- Causal inference goal: estimate the (average) impact of **treatment** $W$ on $Y$, given $X$ describing a population.
- Data is centralized:

| Obs. | Covariates | | | Treatment | Outcome | Potential Outcomes | |
|---|---|---|---|---|---|---|---|
| $i$ | $X_1$ | $X_2$ | $X_3$ | $W$ | $Y$ | $Y^{(1)}$ | $Y^{(0)}$ |
| 1 | 2.3 | 1.5 | M | 1 | 3.2 | 3.2 | ?? |
| 2 | 2.2 | 3.1 | F | 0 | 2.8 | ?? | 2.8 |
| 3 | 3.5 | 2.0 | F | 1 | 2.1 | 2.1 | ?? |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $n-1$ | 3.7 | 2.0 | F | 0 | 2.8 | ?? | 2.8 |
| $n$ | 2.5 | 1.7 | M | 1 | 3.2 | 3.2 | ?? |

**Average Treatment Effect (ATE)**

measured as a risk difference:

$$\tau = \mathbb{E}\left(Y_i(1) - Y_i(0)\right)$$

# Usual Causal Framework - RCT

- Causal inference goal: estimate the (average) impact of **treatment** $W$ on $Y$, given $X$ describing a population.
- Data is centralized:

| Obs. | Covariates | | | Treatment | Outcome | Potential Outcomes | |
|---|---|---|---|---|---|---|---|
| $i$ | $X_1$ | $X_2$ | $X_3$ | $W$ | $Y$ | $Y^{(1)}$ | $Y^{(0)}$ |
| 1 | 2.3 | 1.5 | M | 1 | 3.2 | 3.2 | ?? |
| 2 | 2.2 | 3.1 | F | 0 | 2.8 | ?? | 2.8 |
| 3 | 3.5 | 2.0 | F | 1 | 2.1 | 2.1 | ?? |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $n-1$ | 3.7 | 2.0 | F | 0 | 2.8 | ?? | 2.8 |
| $n$ | 2.5 | 1.7 | M | 1 | 3.2 | 3.2 | ?? |

**Average Treatment Effect (ATE)**

measured as a risk difference:

$$\tau = \mathbb{E}(Y_i(1) - Y_i(0))$$
$$= \mathbb{E}(\mathbb{E}(Y_i(1) \mid X_i) - \mathbb{E}(Y_i(0) \mid X_i))$$

# Usual Causal Framework - RCT

- Causal inference goal: estimate the (average) impact of **treatment** $W$ on $Y$, given $X$ describing a population.
- Data is centralized:

| Obs. | Covariates | | | Treatment | Outcome | Potential Outcomes | |
|---|---|---|---|---|---|---|---|
| $i$ | $X_1$ | $X_2$ | $X_3$ | $W$ | $Y$ | $Y^{(1)}$ | $Y^{(0)}$ |
| 1 | 2.3 | 1.5 | M | 1 | 3.2 | 3.2 | ?? |
| 2 | 2.2 | 3.1 | F | 0 | 2.8 | ?? | 2.8 |
| 3 | 3.5 | 2.0 | F | 1 | 2.1 | 2.1 | ?? |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $n-1$ | 3.7 | 2.0 | F | 0 | 2.8 | ?? | 2.8 |
| $n$ | 2.5 | 1.7 | M | 1 | 3.2 | 3.2 | ?? |

**Average Treatment Effect (ATE)**

measured as a risk difference:

$$\begin{aligned}\tau &= \mathbb{E}(Y_i(1) - Y_i(0)) \\ &= \mathbb{E}(\mathbb{E}(Y_i(1) \mid X_i) - \mathbb{E}(Y_i(0) \mid X_i)) \\ &= \mathbb{E}\big(\mathbb{E}(Y_i(1) \mid W_i = 1, X_i) \\ &\qquad - \mathbb{E}(Y_i(0) \mid W_i = 0, X_i)\big) \qquad \text{(a)}\end{aligned}$$

(a) Unconfoundedness: $W_i \perp\!\!\!\perp \{Y_i(1), Y_i(0)\}$

RCTs: $W_i \sim \mathcal{B}(p_i), \implies W_i \perp\!\!\!\perp X_i$

# Usual Causal Framework - RCT

- Causal inference goal: estimate the (average) impact of **treatment** $W$ on $Y$, given $X$ describing a population.
- Data is centralized:

| Obs. | Covariates | | | Treatment | Outcome | Potential Outcomes | |
|---|---|---|---|---|---|---|---|
| $i$ | $X_1$ | $X_2$ | $X_3$ | $W$ | $Y$ | $Y^{(1)}$ | $Y^{(0)}$ |
| 1 | 2.3 | 1.5 | M | 1 | 3.2 | 3.2 | ?? |
| 2 | 2.2 | 3.1 | F | 0 | 2.8 | ?? | 2.8 |
| 3 | 3.5 | 2.0 | F | 1 | 2.1 | 2.1 | ?? |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| $n-1$ | 3.7 | 2.0 | F | 0 | 2.8 | ?? | 2.8 |
| $n$ | 2.5 | 1.7 | M | 1 | 3.2 | 3.2 | ?? |

**Average Treatment Effect (ATE)**

measured as a risk difference:

$$\begin{aligned}\tau &= \mathbb{E}(Y_i(1) - Y_i(0)) \\ &= \mathbb{E}(\mathbb{E}(Y_i(1) \mid X_i) - \mathbb{E}(Y_i(0) \mid X_i)) \\ &= \mathbb{E}\big(\mathbb{E}(Y_i(1) \mid W_i = 1, X_i) \\ &\qquad - \mathbb{E}(Y_i(0) \mid W_i = 0, X_i)\big) \qquad \text{(a)} \\ &= \mathbb{E}\big(\mathbb{E}(Y_i \mid W_i = 1, X_i) \\ &\qquad - \mathbb{E}(Y_i \mid W_i = 0, X_i)\big) \qquad \text{(b)}\end{aligned}$$

(b) Consistency: $Y_i = W_i Y_i(1) + (1 - W_i) Y_i(0)$

# Usual Causal Framework - RCT

- Causal inference goal: estimate the (average) impact of **treatment** $W$ on $Y$, given $X$ describing a population.
- Data is centralized:

| Obs. | Covariates | | | Treatment | Outcome | Potential Outcomes | |
|---|---|---|---|---|---|---|---|
| $i$ | $X_1$ | $X_2$ | $X_3$ | $W$ | $Y$ | $Y^{(1)}$ | $Y^{(0)}$ |
| 1 | 2.3 | 1.5 | M | 1 | 3.2 | 3.2 | ?? |
| 2 | 2.2 | 3.1 | F | 0 | 2.8 | ?? | 2.8 |
| 3 | 3.5 | 2.0 | F | 1 | 2.1 | 2.1 | ?? |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| $n-1$ | 3.7 | 2.0 | F | 0 | 2.8 | ?? | 2.8 |
| $n$ | 2.5 | 1.7 | M | 1 | 3.2 | 3.2 | ?? |

**Unbiased ATE estimation:**

- Difference-in-Means:
$\hat{\tau}_{\mathrm{DM}} = \overline{Y_{|W=1}} - \overline{Y_{|W=0}}$

# Usual Causal Framework - RCT

- Causal inference goal: estimate the (average) impact of **treatment** $W$ on $Y$, given $X$ describing a population.
- Data is centralized:

| Obs. | Covariates | | | Treatment | Outcome | Potential Outcomes | |
|---|---|---|---|---|---|---|---|
| $i$ | $X_1$ | $X_2$ | $X_3$ | $W$ | $Y$ | $Y^{(1)}$ | $Y^{(0)}$ |
| 1 | 2.3 | 1.5 | M | 1 | 3.2 | 3.2 | $\widehat{2.4}$ |
| 2 | 2.2 | 3.1 | F | 0 | 2.8 | $\widehat{3.5}$ | 2.8 |
| 3 | 3.5 | 2.0 | F | 1 | 2.1 | 2.1 | $\widehat{2.1}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $n-1$ | 3.7 | 2.0 | F | 0 | 2.8 | $\widehat{3.7}$ | 2.8 |
| $n$ | 2.5 | 1.7 | M | 1 | 3.2 | 3.2 | $\widehat{2.9}$ |

**Unbiased ATE estimation:**

- Difference-in-Means:
  $\hat{\tau}_{\mathrm{DM}} = \overline{Y_{|W=1}} - \overline{Y_{|W=0}}$
- Covariate-adjusted G-Formula:
  $\hat{\tau}_{\mathrm{G}} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\mu}_1(X_i) - \hat{\mu}_0(X_i)\right)$

# Usual Causal Framework - RCT

- Causal inference goal: estimate the (average) impact of **treatment** $W$ on $Y$, given $X$ describing a population.
- Data is centralized:

| Obs. | Covariates | | | Treatment | Outcome | Potential Outcomes | |
|---|---|---|---|---|---|---|---|
| $i$ | $X_1$ | $X_2$ | $X_3$ | $W$ | $Y$ | $Y^{(1)}$ | $Y^{(0)}$ |
| 1 | 2.3 | 1.5 | M | 1 | 3.2 | 3.2 | $\widehat{2.4}$ |
| 2 | 2.2 | 3.1 | F | 0 | 2.8 | $\widehat{3.5}$ | 2.8 |
| 3 | 3.5 | 2.0 | F | 1 | 2.1 | 2.1 | $\widehat{2.1}$ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| $n-1$ | 3.7 | 2.0 | F | 0 | 2.8 | $\widehat{3.7}$ | 2.8 |
| $n$ | 2.5 | 1.7 | M | 1 | 3.2 | 3.2 | $\widehat{2.9}$ |

**Unbiased ATE estimation:**

- Difference-in-Means:
  $\hat{\tau}_{\mathrm{DM}} = \overline{Y_{|W=1}} - \overline{Y_{|W=0}}$
- Linearly-adjusted G-Formula:
  $\hat{\tau}_{\mathrm{OLS}} = \frac{1}{n}\sum_{i=1}^{n}\left(X_i\hat{\beta}^{(1)} - X_i\hat{\beta}^{(0)}\right)$

# Usual Causal Framework - RCT

- Causal inference goal: estimate the (average) impact of **treatment** $W$ on $Y$, given $X$ describing a population.
- Data is centralized:

| Obs. | Covariates | | | Treatment | Outcome | Potential Outcomes | |
|---|---|---|---|---|---|---|---|
| $i$ | $X_1$ | $X_2$ | $X_3$ | $W$ | $Y$ | $Y^{(1)}$ | $Y^{(0)}$ |
| 1 | 2.3 | 1.5 | M | 1 | 3.2 | 3.2 | ?? |
| 2 | 2.2 | 3.1 | F | 0 | 2.8 | ?? | 2.8 |
| 3 | 3.5 | 2.0 | F | 1 | 2.1 | 2.1 | ?? |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| $n-1$ | 3.7 | 2.0 | F | 0 | 2.8 | ?? | 2.8 |
| $n$ | 2.5 | 1.7 | M | 1 | 3.2 | 3.2 | ?? |

**Unbiased ATE estimation:**

- Difference-in-Means:
  $\hat{\tau}_{\mathrm{DM}} = \overline{Y_{|W=1}} - \overline{Y_{|W=0}}$
- Linearly-adjusted G-Formula:
  $\hat{\tau}_{\mathrm{OLS}} = \frac{1}{n}\sum_{i=1}^{n}\left(X_i\hat{\beta}^{(1)} - X_i\hat{\beta}^{(0)}\right)$
  with $\hat{\beta}^{(w)}$ the OLS regressor learned on individuals with $W = w$.

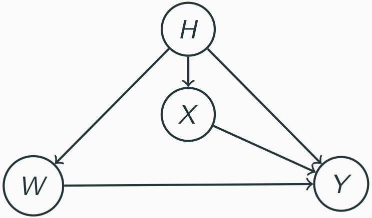
Refs.: U.S. Food and Drug Administration, 2023, European Medicines Agency, 2024, Tsiatis et al., 2008, Benkeser et al., 2021, Lin, 2013, Wager, 2020, Lei and Ding, 2021, Van Lancker, Bretz, and Dukes, 2024

Main Motivation: $\mathbb{V}(\hat{\tau}_{\mathrm{OLS}}) \leq \mathbb{V}(\hat{\tau}_{\mathrm{DM}})$ even when $\mu_1$ and $\mu_0$ are not linear functions!

# Multisource Causal Framework

- **Multisource** inference goal: estimate the impact of $W$ on $Y$ given $X$ describing a population, split across $K$ studies.
- Data is decentralized:

| Source | Obs. | Covariates | | | Treatment | Outcomes |
|---|---|---|---|---|---|---|
| $H$ | $i$ | $X_1$ | $X_2$ | $X_3$ | $W$ | $Y$ |
| 1 | 1 | 2.3 | 1.5 | M | 1 | 3.2 |
| 1 | 2 | 2.2 | 3.1 | F | 0 | 2.8 |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| 2 | 1 | 4.5 | 5.0 | F | 1 | 4.1 |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $K$ | 1 | 3.7 | 2.0 | F | 0 | 2.8 |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $K$ | $n_K$ | 2.5 | 1.7 | M | 0 | 3.2 |




**ATE**:

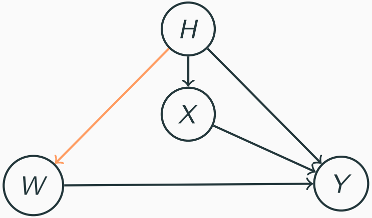
$$\tau = \mathbb{E}\left(\mathbb{E}\left(Y^{(1)} - Y^{(0)} \mid H\right)\right)$$

# Multisource Causal Framework

- **Multisource** inference goal: estimate the impact of $W$ on $Y$ given $X$ describing a population, split across $K$ studies.
- Data is decentralized:

| Source | Obs. | Covariates | | | Treatment | Outcomes | |
|---|---|---|---|---|---|---|---|
| $H$ | $i$ | $X_1$ | $X_2$ | $X_3$ | $W$ | $Y$ | $\mathbb{P}(W_i)$ |
| 1 | 1 | 2.3 | 1.5 | M | 1 | 3.2 | $p_1$ |
| 1 | 2 | 2.2 | 3.1 | F | 0 | 2.8 | $p_1$ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 2 | 1 | 4.5 | 5.0 | F | 1 | 4.1 | $p_2$ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| $K$ | 1 | 3.7 | 2.0 | F | 0 | 2.8 | $p_3$ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| $K$ | $n_K$ | 2.5 | 1.7 | M | 0 | 3.2 | $p_3$ |

Heterogeneity in **treatment allocation**




**ATE**:

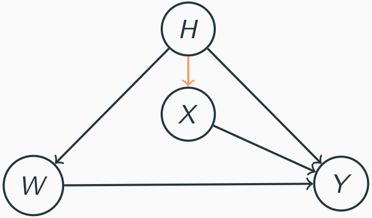
$$\tau = \mathbb{E}\left(\mathbb{E}\left(Y^{(1)} - Y^{(0)} \mid H\right)\right)$$

# Multisource Causal Framework

- **Multisource** inference goal: estimate the impact of $W$ on $Y$ given $X$ describing a population, split across $K$ studies.
- Data is decentralized:

| Source | Obs. | Covariates | | | Treatment | Outcomes | |
|---|---|---|---|---|---|---|---|
| $H$ | $i$ | $X_1$ | $X_2$ | $X_3$ | $W$ | $Y$ | $\mathcal{D}$ |
| 1 | 1 | 2.3 | 1.5 | M | 1 | 3.2 | $\mathcal{D}_1$ |
| 1 | 2 | 2.2 | 3.1 | F | 0 | 2.8 | $\mathcal{D}_1$ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 2 | 1 | 4.5 | 5.0 | F | 1 | 4.1 | $\mathcal{D}_2$ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| $K$ | 1 | 3.7 | 2.0 | F | 0 | 2.8 | $\mathcal{D}_3$ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| $K$ | $n_K$ | 2.5 | 1.7 | M | 0 | 3.2 | $\mathcal{D}_3$ |

Heterogeneity in **covariates distribution**




**ATE**:

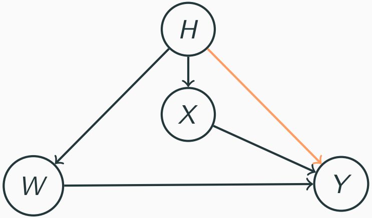
$$\tau = \mathbb{E}\left(\mathbb{E}\left(Y^{(1)} - Y^{(0)} \mid H\right)\right)$$

# Multisource Causal Framework

- **Multisource** inference goal: estimate the impact of $W$ on $Y$ given $X$ describing a population, split across $K$ studies.
- Data is decentralized:

| Source | Obs. | Covariates | | | Treatment | Outcomes | |
|---|---|---|---|---|---|---|---|
| $H$ | $i$ | $X_1$ | $X_2$ | $X_3$ | $W$ | $Y$ | + cste |
| 1 | 1 | 2.3 | 1.5 | M | 1 | 3.2 | $h_1$ |
| 1 | 2 | 2.2 | 3.1 | F | 0 | 2.8 | $h_1$ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 2 | 1 | 4.5 | 5.0 | F | 1 | 4.1 | $h_2$ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| $K$ | 1 | 3.7 | 2.0 | F | 0 | 2.8 | $h_3$ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| $K$ | $n_K$ | 2.5 | 1.7 | M | 0 | 3.2 | $h_3$ |

Heterogeneity in **center effects**




**ATE**:

$$\tau = \mathbb{E}\left(\mathbb{E}\left(Y^{(1)} - Y^{(0)} \mid H\right)\right)$$

$\rightarrow$ How to estimate $\tau$?

# Federated Causal Inferences: estimation strategies

$k = 1$  $\hat{\tau}_1$

$k = 2$  $\hat{\tau}_2$

$\vdots$

$k = K$  $\hat{\tau}_K$

# Federated Causal Inferences: estimation strategies

$k = 1$  $\hat{\tau}_1$

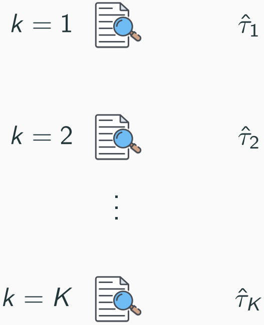
$k = 2$  $\hat{\tau}_2$

$\vdots$

$k = K$  $\hat{\tau}_K$

**Global estimates of the treatment effect**:

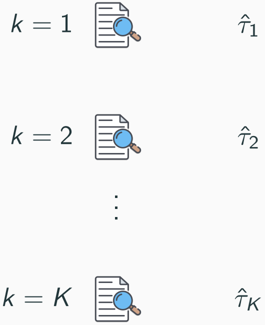
# Federated Causal Inferences: estimation strategies




**Global estimates of the treatment effect**:

- **Meta-analysis**: average-weighted aggregation of local estimates $\rightarrow \hat{\tau} = \sum_{k=1}^{K} w_k \hat{\tau}_k$

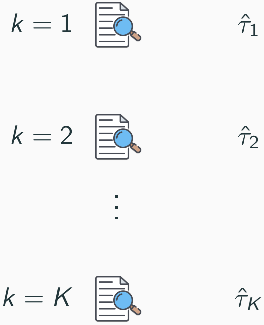
# Federated Causal Inferences: estimation strategies




**Global estimates of the treatment effect**:

- **Meta-analysis**: average-weighted aggregation of local estimates $\rightarrow \hat{\tau} = \sum_{k=1}^{K} w_k \hat{\tau}_k$
- Federated estimation:

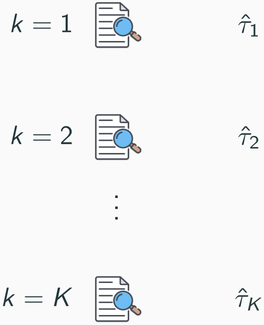
# Federated Causal Inferences: estimation strategies




**Global estimates of the treatment effect**:

- **Meta-analysis**: average-weighted aggregation of local estimates $\rightarrow \hat{\tau} = \sum_{k=1}^{K} w_k \hat{\tau}_k$
- Federated estimation:
  a. Learn the parameters for the outcome and/or propensity score models

# Federated Causal Inferences: estimation strategies




**Global estimates of the treatment effect**:

- **Meta-analysis**: average-weighted aggregation of local estimates $\rightarrow \hat{\tau} = \sum_{k=1}^{K} w_k \hat{\tau}_k$
- Federated estimation:
  a. Learn the parameters for the outcome and/or propensity score models
  b. Build a global estimate $\hat{\tau}$ from these models (e.g., G-Formula, IPW, AIPW).

# Multi-sources ATE Estimation

# Federated ATE estimation in linear outcome modelling

- Denote $X_k^{(w)}$ (resp. $Y_k^{(w)}$) the covariates matrix (resp. outcome vector) of study $k$ under treatment arm $w$

# Federated ATE estimation in linear outcome modelling

- Denote $X_k^{(w)}$ (resp. $Y_k^{(w)}$) the covariates matrix (resp. outcome vector) of study $k$ under treatment arm $w$
- Linear outcome model: $\forall k, Y_{k,i}^{(w)} = c^{(w)} + X_{k,i}\beta^{(w)} + \varepsilon_{k,i}^{(w)}$ with $w \in \{0, 1\}$.

# Federated ATE estimation in linear outcome modelling

- Denote $X_k^{(w)}$ (resp. $Y_k^{(w)}$) the covariates matrix (resp. outcome vector) of study $k$ under treatment arm $w$
- Linear outcome model: $\forall k$, $Y_{k,i}^{(w)} = c^{(w)} + X_{k,i}\beta^{(w)} + \varepsilon_{k,i}^{(w)}$ with $w \in \{0,1\}$.

$\rightarrow$ Goal: estimate $\tau := \mathbb{E}\left(\mathbb{E}(Y_i^{(1)} - Y_i^{(0)} \mid H_i)\right) = c^{(1)} - c^{(0)} + \mathbb{E}(\mathbb{E}(X_i \mid H_i))(\beta^{(1)} - \beta^{(0)})$.

# Federated ATE estimation in linear outcome modelling

- Denote $X_k^{(w)}$ (resp. $Y_k^{(w)}$) the covariates matrix (resp. outcome vector) of study $k$ under treatment arm $w$
- Linear outcome model: $\forall k$, $Y_{k,i}^{(w)} = c^{(w)} + X_{k,i}\beta^{(w)} + \varepsilon_{k,i}^{(w)}$ with $w \in \{0,1\}$.

**Local causality assumptions**:

- Consistency: $\forall i, Y_i = W_i Y_i^{(1)} + (1 - W_i) Y_i^{(0)}$
- Positivity: $\forall x \in \mathcal{X}, \exists \eta > 0$ s.t. $\eta \leq \mathbb{P}(W_i = 1 \,|\, X_i = x) \leq 1 - \eta$
- Unconfoundedness: $W_i \perp\!\!\!\perp \{Y_i^{(1)}, Y_i^{(0)}\} \,|\, X_i, H_i$

# Federated ATE estimation in linear outcome modelling

- Denote $X_k^{(w)}$ (resp. $Y_k^{(w)}$) the covariates matrix (resp. outcome vector) of study $k$ under treatment arm $w$
- Linear outcome model: $\forall k, Y_{k,i}^{(w)} = c^{(w)} + X_{k,i}\beta^{(w)} + \varepsilon_{k,i}^{(w)}$ with $w \in \{0, 1\}$.

**Regression assumptions**:

- $\forall (k, w), \mathbb{E}(X_k^\top \varepsilon(w)) = 0, \mathbb{V}\left(\varepsilon(w) \mid X_k\right) = \sigma^2,$
- Local Full Rank: $\text{rank}(X_k^\top X_k) = d$

# Federated ATE estimation in linear outcome modelling

- Denote $X_k^{(w)}$ (resp. $Y_k^{(w)}$) the covariates matrix (resp. outcome vector) of study $k$ under treatment arm $w$
- Linear outcome model: $\forall k$, $Y_{k,i}^{(w)} = c^{(w)} + X_{k,i}\beta^{(w)} + \varepsilon_{k,i}^{(w)}$ with $w \in \{0, 1\}$.

**Pool G-Formula estimator:**

$$\hat{\tau} = \frac{1}{n}\sum_{i=1}^{n}(\hat{\mu}_1(X) - \hat{\mu}_0(X)) = \frac{1}{n}\sum_{i=1}^{n} X_i(\hat{\theta}^{(1)}_{\text{pool}} - \hat{\theta}^{(0)}_{\text{pool}})$$

with $\hat{\theta}^{(w)}_{\text{pool}} = \{\hat{c}^{(w)}_{\text{pool}}, \hat{\beta}^{(w)}_{\text{pool}}\}$ the OLS regressor over the pooled dataset.

# Federated ATE estimation in linear outcome modelling

- Denote $X_k^{(w)}$ (resp. $Y_k^{(w)}$) the covariates matrix (resp. outcome vector) of study $k$ under treatment arm $w$
- Linear outcome model: $\forall k$, $Y_{k,i}^{(w)} = c^{(w)} + X_{k,i}\beta^{(w)} + \varepsilon_{k,i}^{(w)}$ with $w \in \{0, 1\}$.

**Pool G-Formula estimator:**

$$\hat{\tau} = \frac{1}{n}\sum_{i=1}^{n}(\hat{\mu}_1(X) - \hat{\mu}_0(X)) = \frac{1}{n}\sum_{i=1}^{n} X_i(\hat{\theta}_{\text{pool}}^{(1)} - \hat{\theta}_{\text{pool}}^{(0)})$$

with $\hat{\theta}_{\text{pool}}^{(w)} = \{\hat{c}_{\text{pool}}^{(w)}, \hat{\beta}_{\text{pool}}^{(w)}\}$ the OLS regressor over the pooled dataset.

→ Problem: we do not have access to the pooled dataset !

# Federated ATE estimation in linear outcome modelling

- Denote $X_k^{(w)}$ (resp. $Y_k^{(w)}$) the covariates matrix (resp. outcome vector) of study $k$ under treatment arm $w$
- Linear outcome model: $\forall k$, $Y_{k,i}^{(w)} = c^{(w)} + X_{k,i}\beta^{(w)} + \varepsilon_{k,i}^{(w)}$ with $w \in \{0,1\}$.

**Pool G-Formula estimator:**

$$\hat{\tau} = \frac{1}{n}\sum_{i=1}^{n}(\hat{\mu}_1(X) - \hat{\mu}_0(X)) = \frac{1}{n}\sum_{i=1}^{n} X_i(\hat{\theta}_{\mathrm{pool}}^{(1)} - \hat{\theta}_{\mathrm{pool}}^{(0)})$$

with $\hat{\theta}_{\mathrm{pool}}^{(w)} = \{\hat{c}_{\mathrm{pool}}^{(w)}, \hat{\beta}_{\mathrm{pool}}^{(w)}\}$ the OLS regressor over the pooled dataset.

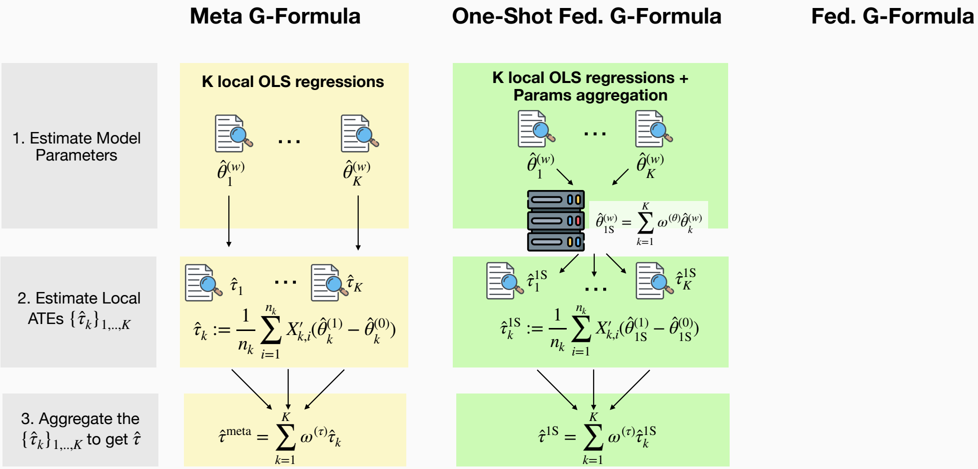
$\rightarrow$ Problem: we do not have access to the pooled dataset !

$\rightarrow$ Need to consider other estimation strategies for $\hat{\tau}$.

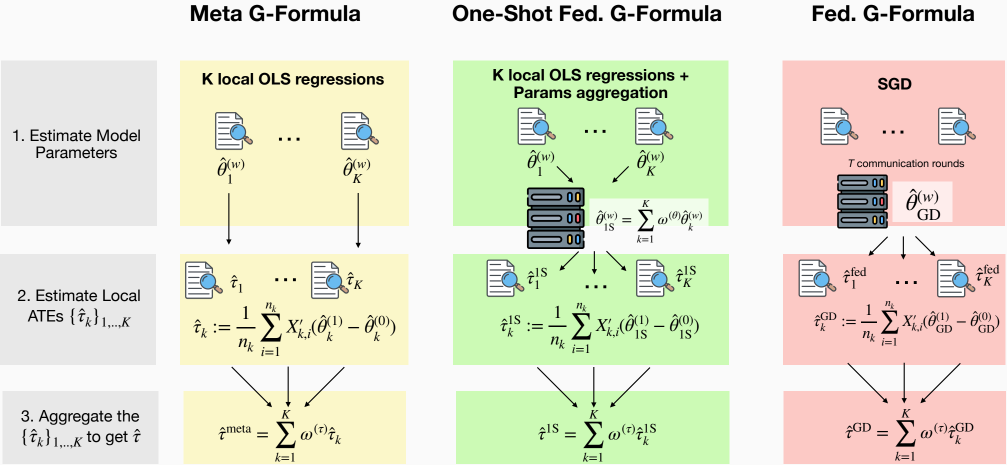
# Meta vs. (One Shot) Federated G-Formula

| | Meta G-Formula | One-Shot Fed. G-Formula | Fed. G-Formula |
|---|---|---|---|
| 1. Estimate Model Parameters | **K local OLS regressions**<br>$\hat{\theta}_1^{(w)}$ ... $\hat{\theta}_K^{(w)}$ | | |
| 2. Estimate Local ATEs $\{\hat{\tau}_k\}_{1,..,K}$ | $\hat{\tau}_1$ ... $\hat{\tau}_K$<br>$\hat{\tau}_k := \frac{1}{n_k} \sum_{i=1}^{n_k} X'_{k,i}(\hat{\theta}_k^{(1)} - \hat{\theta}_k^{(0)})$ | | |
| 3. Aggregate the $\{\hat{\tau}_k\}_{1,..,K}$ to get $\hat{\tau}$ | $\hat{\tau}^{\text{meta}} = \sum_{k=1}^{K} \omega^{(\tau)} \hat{\tau}_k$ | | |

# Meta vs. (One Shot) Federated G-Formula



# Meta vs. (One Shot) Federated G-Formula



# Comparison of the Estimators

## Homogeneous setting

# Asymptotic Variances

Homogeneous population setting:

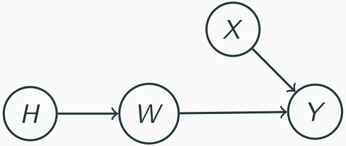



**Figure 1:** Graphical model: $K$ RCTs

$\implies \forall\{k, i\}, W_{k,i} \sim \mathcal{B}(p_k)$, different treatment allocation schemes.

# Asymptotic Variances

Under an Homogeneous setting, all estimators are unbiased and:

| Estimator | Notation | $\mathbb{V}^\infty$ | Com. rounds | Com. cost |
|---|---|---|---|---|
| Meta-SW | $\hat{\tau}_{\text{Meta-SW}}$ | $\frac{\sigma^2}{n}\sum_{k=1}^{K}\frac{\rho_k}{p_k(1-p_k)}+\frac{1}{n}\lVert\beta^{(1)}-\beta^{(0)}\rVert_\Sigma^2$ | 1 | $O(1)$ |
| Meta-IVW | $\hat{\tau}_{\text{Meta-IVW}}$ | $\left(\sum_{k=1}^{K}(\sigma^2\frac{n\rho_k}{p_k(1-p_k)}+\frac{1}{n_k}\lVert\beta^{(1)}-\beta^{(0)}\rVert_\Sigma^2)^{-1}\right)^{-1}$ | 1 | $O(1)$ |
| 1S-SW | $\hat{\tau}_{\text{1S-SW}}$ | $V_{\text{pool}}$ | 2 | $O(d)$ |
| 1S-IVW | $\hat{\tau}_{\text{1S-IVW}}$ | $V_{\text{pool}}$ | 2 | $O(d^2)$ |
| GD | $\hat{\tau}_{\text{GD}}$ | $V_{\text{pool}}$ | $T+1$ | $O(Td)$ |
| Pool | $\hat{\tau}_{\text{pool}}$ | $V_{\text{pool}}=\frac{\sigma^2}{n}\frac{1}{p(1-p)}+\frac{1}{n}\lVert\beta^{(1)}-\beta^{(0)}\rVert_\Sigma^2$ | — | — |

with $\rho_k := \mathbb{P}(H_i = k) = \mathbb{E}(n_k)/n$ and $p = \sum_{k=1}^K \frac{n_k}{n} p_k$.

# Comparison of Variances - Homogeneous Setting

$$\begin{aligned}\mathbb{V}^{\infty}(\hat{\tau}_{\mathsf{pool}}) &= \mathbb{V}^{\infty}(\hat{\tau}_{\mathrm{GD}}) \\ &= \mathbb{V}^{\infty}(\hat{\tau}_{\mathrm{1S-SW}}) \\ &= \mathbb{V}^{\infty}(\hat{\tau}_{\mathrm{1S-IVW}})\end{aligned}$$

# Comparison of Variances - Homogeneous Setting

$$\begin{aligned}\mathbb{V}^{\infty}(\hat{\tau}_{\text{pool}}) &= \mathbb{V}^{\infty}(\hat{\tau}_{\mathrm{GD}}) \\ &= \mathbb{V}^{\infty}(\hat{\tau}_{\mathrm{1S-SW}}) \\ &= \mathbb{V}^{\infty}(\hat{\tau}_{\mathrm{1S-IVW}}) \\ &\qquad \leq \mathbb{V}^{\infty}(\hat{\tau}_{\mathrm{Meta-IVW}}) \left\{ \right.\end{aligned}$$

# Comparison of Variances - Homogeneous Setting

$$
\begin{aligned}
\mathbb{V}^{\infty}(\hat{\tau}_{\text{pool}}) &= \mathbb{V}^{\infty}(\hat{\tau}_{\mathrm{GD}}) \\
&= \mathbb{V}^{\infty}(\hat{\tau}_{\mathrm{1S-SW}}) \\
&= \mathbb{V}^{\infty}(\hat{\tau}_{\mathrm{1S-IVW}}) \\
&\qquad\qquad \leq \mathbb{V}^{\infty}(\hat{\tau}_{\mathrm{Meta-IVW}}) \Bigg\{ \\
&\qquad\qquad\qquad\qquad \leq \mathbb{V}^{\infty}(\hat{\tau}_{\mathrm{Meta-SW}}) \Bigg\{
\end{aligned}
$$

# Comparison of Variances - Homogeneous Setting

$$
\begin{aligned}
\mathbb{V}^{\infty}(\hat{\tau}_{\text{pool}}) &= \mathbb{V}^{\infty}(\hat{\tau}_{\mathrm{GD}}) \\
&= \mathbb{V}^{\infty}(\hat{\tau}_{\mathrm{1S-SW}}) \\
&= \mathbb{V}^{\infty}(\hat{\tau}_{\mathrm{1S-IVW}}) \\
&\qquad \leq \mathbb{V}^{\infty}(\hat{\tau}_{\mathrm{Meta-IVW}}) \begin{cases} = & \text{if same } \{p_k\}_k, \\ & \end{cases} \\
&\qquad\qquad \leq \mathbb{V}^{\infty}(\hat{\tau}_{\mathrm{Meta-SW}}) \begin{cases} = & \text{if same } \{p_k(1-p_k)\}_k, \\ & \end{cases}
\end{aligned}
$$

# Comparison of Variances - Homogeneous Setting

$$
\begin{aligned}
\mathbb{V}^\infty(\hat{\tau}_{\text{pool}}) &= \mathbb{V}^\infty(\hat{\tau}_{\mathrm{GD}}) \\
&= \mathbb{V}^\infty(\hat{\tau}_{\mathrm{1S-SW}}) \\
&= \mathbb{V}^\infty(\hat{\tau}_{\mathrm{1S-IVW}}) \\
&\qquad \leq \mathbb{V}^\infty(\hat{\tau}_{\mathrm{Meta-IVW}}) \begin{cases} = & \text{if same } \{p_k\}_k, \\ < & \text{if different } \{p_k\}_k \end{cases} \\
&\qquad\qquad \leq \mathbb{V}^\infty(\hat{\tau}_{\mathrm{Meta-SW}}) \begin{cases} = & \text{if same } \{p_k(1-p_k)\}_k, \\ < & \text{if different } \{p_k(1-p_k)\}_k \end{cases}
\end{aligned}
$$

# Comparison of Variances - Homogeneous Setting

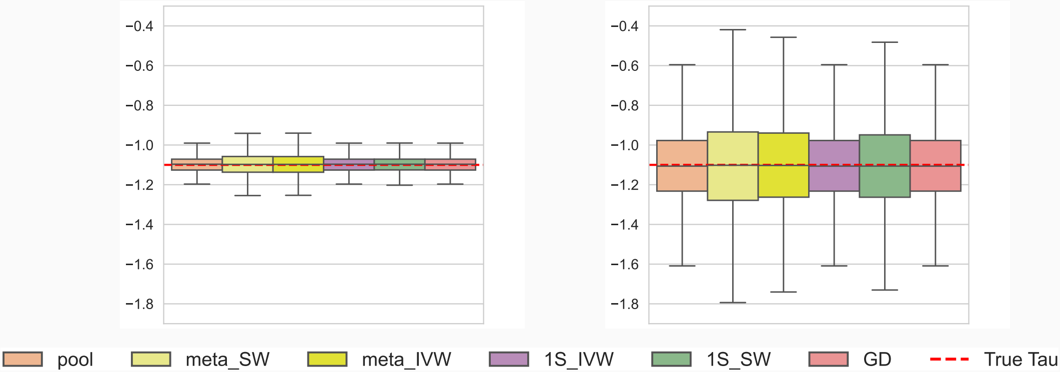
Parameters: $d = 10, K = 5$

**Large**

$\{n_k\}_k = 100 * d$, 3 studies have $p_k = 0.9$, 2 have $p_k = 0.1$

**Small**

$\{n_k\}_k = 5 * d$, 3 studies have $p_k = 0.65$, 2 have $p_k = 0.35$



# Comparison of the Estimators

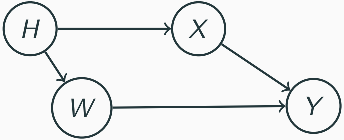
## Heterogeneous Distributions

# Comparison of Variances - Heterogeneity in $X$

Distributional Shift:

$$H \not\perp X \implies \mathcal{D}_k \neq \mathcal{D}_l \implies \tau_k \neq \tau_l$$

**Figure 1:** Graphical model for the heterogeneous distributions setting.



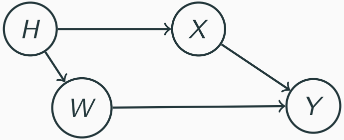

$\tau = \sum_{k=1}^{K} \rho_k \tau_k$ with $\rho_k = \mathbb{E}\left(\frac{n_k}{n}\right)$

# Comparison of Variances - Heterogeneity in $X$

Distributional Shift:

$$H \not\perp\!\!\!\perp X \implies \mathcal{D}_k \neq \mathcal{D}_l \implies \tau_k \neq \tau_l$$

**Figure 1:** Graphical model for the heterogeneous distributions setting.




$\tau = \sum_{k=1}^{K} \rho_k \tau_k$ with $\rho_k = \mathbb{E}\left(\frac{n_k}{n}\right)$

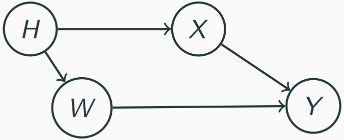
- $\hat{\tau}_{\mathrm{meta-IVW}} = \frac{\sum_{k=1}^{K} \mathbb{V}(\hat{\tau}_k)^{-1} \hat{\tau}_k}{\sum_{l=1}^{K} \mathbb{V}(\hat{\tau}_l)^{-1}}$ is biased: inverse variance weights do not aim at the $\{\rho_k\}s$.

# Comparison of Variances - Heterogeneity in $X$

Distributional Shift:

$$H \not\perp X \implies \mathcal{D}_k \neq \mathcal{D}_l \implies \tau_k \neq \tau_l$$

**Figure 1:** Graphical model for the heterogeneous distributions setting.




$\tau = \sum_{k=1}^{K} \rho_k \tau_k$ with $\rho_k = \mathbb{E}\left(\frac{n_k}{n}\right)$

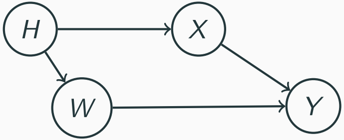
- $\hat{\tau}_{\text{meta-IVW}} = \frac{\sum_{k=1}^{K} \mathbb{V}(\hat{\tau}_k)^{-1} \hat{\tau}_k}{\sum_{l=1}^{K} \mathbb{V}(\hat{\tau}_l)^{-1}}$ is biased: inverse variance weights do not aim at the $\{\rho_k\}s$.
- $\mathbb{V}^{\infty}(\hat{\tau}_{\text{1S-SW}})$ is impacted by the different means $\{\mu_k\}_k$.

# Comparison of Variances - Heterogeneity in $X$

Distributional Shift:

$$H \not\perp\!\!\!\perp X \implies \mathcal{D}_k \neq \mathcal{D}_l \implies \tau_k \neq \tau_l$$

**Figure 1:** Graphical model for the heterogeneous distributions setting.



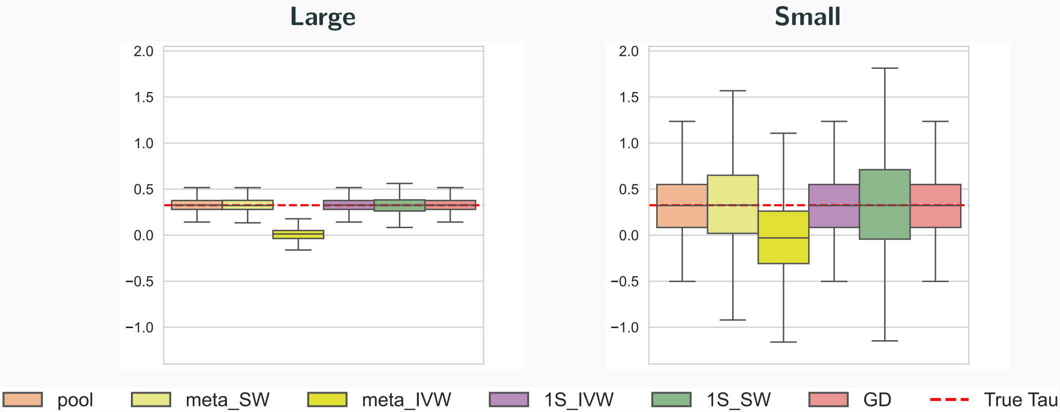

$\tau = \sum_{k=1}^{K} \rho_k \tau_k$ with $\rho_k = \mathbb{E}\left(\frac{n_k}{n}\right)$

- $\hat{\tau}_{\mathrm{meta-IVW}} = \frac{\sum_{k=1}^{K} \mathbb{V}(\hat{\tau}_k)^{-1} \hat{\tau}_k}{\sum_{l=1}^{K} \mathbb{V}(\hat{\tau}_l)^{-1}}$ is biased: inverse variance weights do not aim at the $\{\rho_k\}$s.
- $\mathbb{V}^{\infty}(\hat{\tau}_{\mathrm{1S-SW}})$ is impacted by the different means $\{\mu_k\}_k$.

$$\mathbb{V}^{\infty}(\hat{\tau}_{\mathrm{pool}}) = \mathbb{V}^{\infty}(\hat{\tau}_{\mathrm{GD}}) = \mathbb{V}^{\infty}(\hat{\tau}_{\mathrm{1S-IVW}}) \leq \mathbb{V}^{\infty}(\hat{\tau}_{\mathrm{meta-SW}})$$

# Comparison of Variances - Heterogeneity in $X$

$$\forall k, X_k \sim \mathcal{N}(\mu_k, \Sigma_k)$$



# Comparison of the Estimators

## Presence of Center Effects

# Comparison of Variances - Heterogeneity through Center Effects

Presence of a constant (real-valued) effect of the center $k$ onto the outcome $Y$:

$$Y_{k,i}^{(w)} = c^{(w)} + \underline{h_k} + X_{k,i}\beta^{(w)} + \varepsilon_i(w)$$

# Comparison of Variances - Heterogeneity through Center Effects

Presence of a constant (real-valued) effect of the center $k$ onto the outcome $Y$:

$$Y_{k,i}^{(w)} = c^{(w)} + \underline{h_k} + X_{k,i}\beta^{(w)} + \varepsilon_i(w)$$

Studies may have different baselines in individual outcomes, from varying practices or organizational contexts (e.g. hospital specialized in oncology).

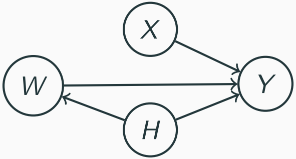



**Figure 1:** Graphical model for the center effects setting.

# Comparison of Variances - Heterogeneity through Center Effects

Presence of a constant (real-valued) effect of the center $k$ onto the outcome $Y$:

$$Y_{k,i}^{(w)} = c^{(w)} + \underline{h_k} + X_{k,i}\beta^{(w)} + \varepsilon_i(w)$$

Studies may have different baselines in individual outcomes, from varying practices or organizational contexts (e.g. hospital specialized in oncology).

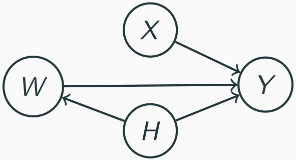



**Figure 1:** Graphical model for the center effects setting.

Caution: $H$ is now a confounder!

# Comparison of Variances - Heterogeneity through Center Effects

Presence of a constant (real-valued) effect of the center $k$ onto the outcome $Y$:

$$Y_{k,i}^{(w)} = c^{(w)} + \underline{h_k} + X_{k,i}\beta^{(w)} + \varepsilon_i(w)$$

Properties:

- All federated estimators are **biased** and need to be adjusted **except the metas** which naturally account for the center effects.

# Comparison of Variances - Heterogeneity through Center Effects

Presence of a constant (real-valued) effect of the center $k$ onto the outcome $Y$:

$$Y_{k,i}^{(w)} = c^{(w)} + \underline{h_k} + X_{k,i}\beta^{(w)} + \varepsilon_i(w)$$

Properties:

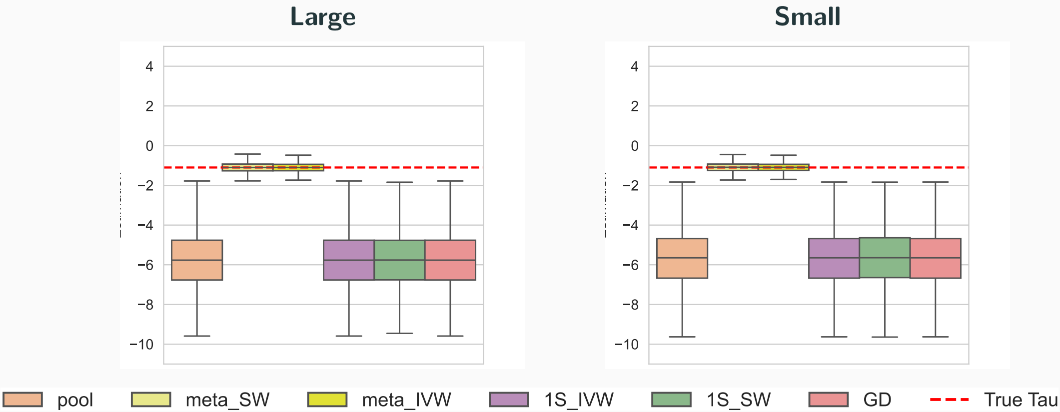
- All federated estimators are **biased** and need to be adjusted **except the metas** which naturally account for the center effects.

Adjustment:

- Adjusted One-Shot estimators: share and aggregate only the covariates coefficients $\hat{\beta}_k$, while keeping the intercepts local
- Adjusted Gradient Descent: add $H$ variable into the datasets.

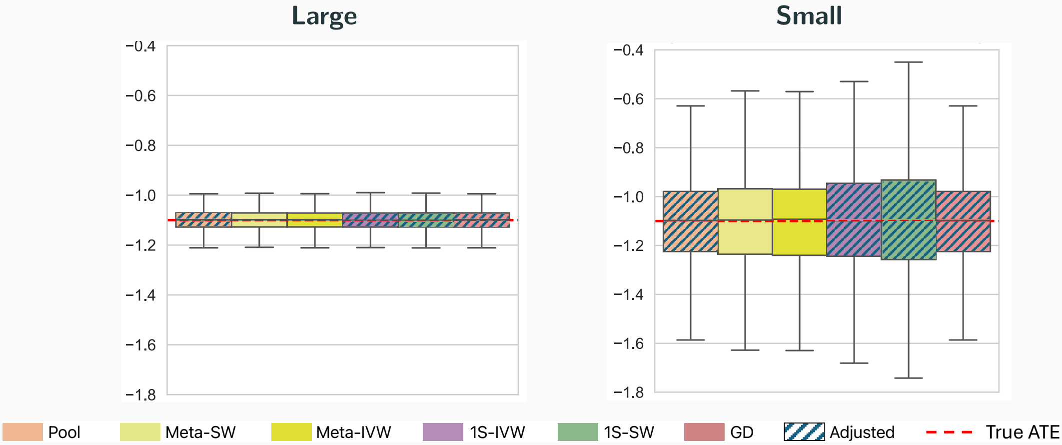
# Comparison of Variances - Heterogeneity through Center Effects

$$(h_1, h_2, h_3, h_4, h_5) = (1, .2, -1, 30, 2) \text{ and different } p_k$$



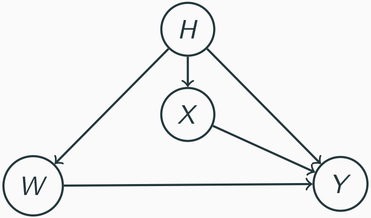
# Comparison of Variances - Heterogeneity through Center Effects

$(h_1, h_2, h_3, h_4, h_5) = (1, .2, -1, 30, 2)$ and different $p_k$



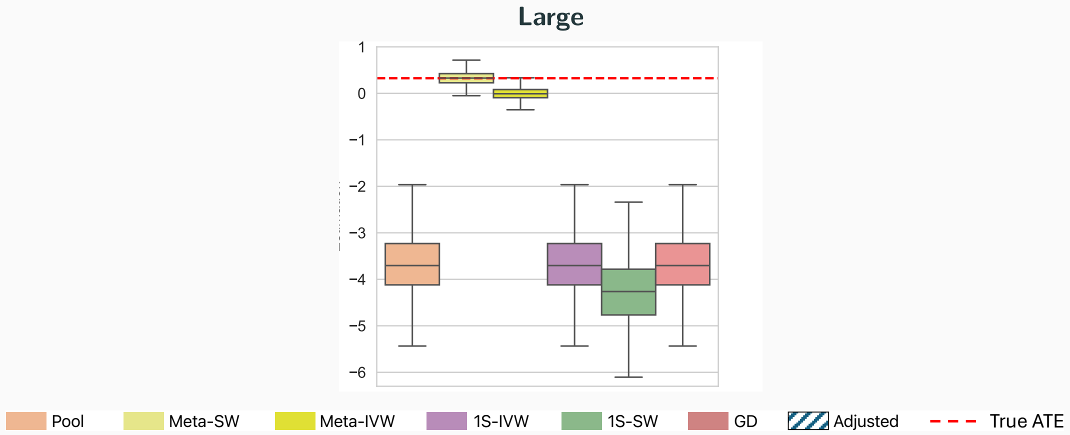
# Full Heterogeneity

Fully heterogeneous setting, realistic:



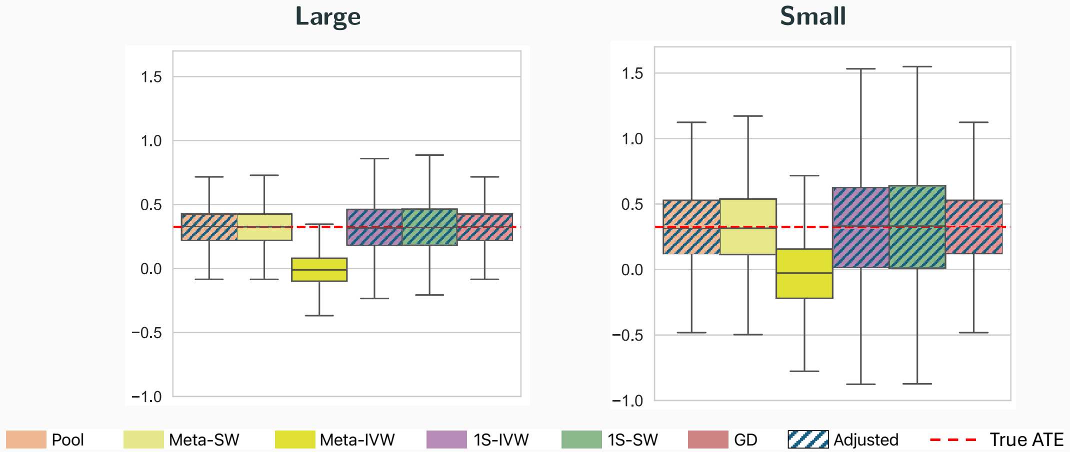
# Full Heterogeneity

Different $(h_k, p_k, \mu_k, \Sigma_k)$



# Full Heterogeneity

Different $h_k, p_k, \mu_k, \Sigma_k$



# Federated Causal Inferences: challenges

| | Meta-Analysis | One-Shot FL Learning | GD FL |
|---|---|---|---|
| + | • Easy to implement<br>• Private and low communications: 1 round<br>• Shares only summary statistics:<br>  ◦ Locally estimated ATEs $\{\hat{\tau}_k\}$<br>  ◦ Sample sizes $\{n_k\}$ or estimated variances $\{\hat{V}(\hat{\tau}_k)\}$ | • Easy to implement<br>• Private and low communications: 2 rounds<br>• Shares summary statistics:<br>  ◦ Sample sizes $\{n_k\}$ or empirical variance-covariance matrices $\{\hat{\Sigma}_k\}$ (can be costly when $d$ is large)<br>  ◦ Locally estimated ATEs with One-Shot federated outcome models $\{\hat{\tau}_k^{1\mathrm{S}}\}$ | • Flexible: (non-)parametric models, estimate <u>function</u> $\tau(X)$<br>• **Robust to locally small sample sizes** $(n_k^{(w)} < d)$<br>• **Robust to different treatment schemes**<br>• Private: using secure aggregation or differential privacy<br>• Accurate: learn from the pool dataset as if it was centralized |
| - | • Sensitive to imbalance in sample sizes<br>• No access to individual data: cannot detect and qualify heterogeneity<br>• The aggregation with lowest variance (IVW) yields a biased estimate under heterogeneity in distributions | • Not designed for heterogeneous settings | • Harder to implement in practice<br>• Heavy computations: compute $\nabla f(\hat{\theta})$ at each round |

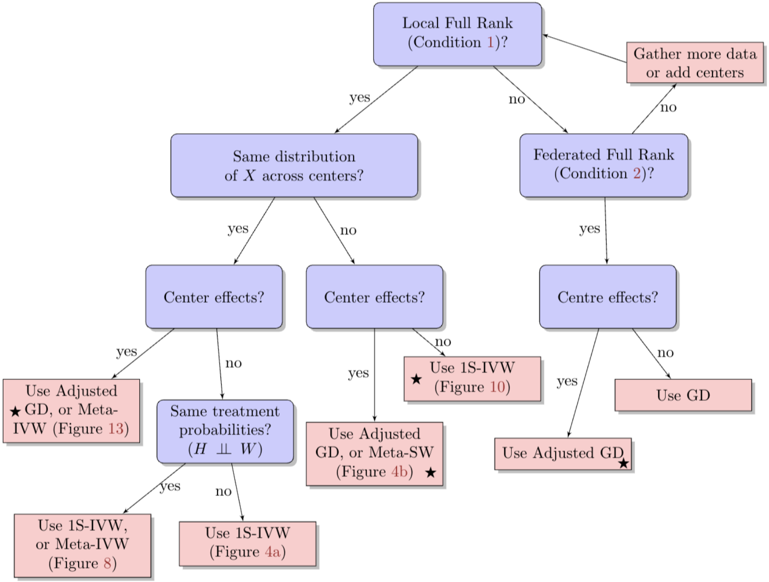



Figure 6: Decision Diagram for Practitionners. The sign ★ denotes scenarios where the DM estimator is biased.

# Perspectives

- Extend this work to **observational studies**.
- Quantify the bias and variances of the estimators in **finite sample sizes** settings.
- **Non-parametric** and non-linear approaches: federated random forests, neural networks, etc.
- Apply the **Differential-Privacy** framework to federated causal inferences.

# Scan my paper!

Thank you!

# References


Benkeser, David et al. (2021). **"Improving precision and power in randomized trials for COVID-19 treatments using covariate adjustment, for binary, ordinal, and time-to-event outcomes".** In: *Biometrics* 77.4, pp. 1467–1481.

Duflo, Esther, Rachel Glennerster, and Michael Kremer (2007). **"Using randomization in development economics research: A toolkit".** In: *Handbook of development economics* 4, pp. 3895–3962.

European Medicines Agency (2024). **ICH E9 Statistical Principles for Clinical Trials: Scientific Guideline.**

French Health Authority (2024). **Pricing & Reimbursement of drugs and HTA policies in France.** URL: https://www.has-sante.fr/upload/docs/application/pdf/2014-03/pricing_reimbursement_of_drugs_and_hta_policies_in_france.pdf.

Lei, Lihua and Peng Ding (2021). **"Regression adjustment in completely randomized experiments with a diverging number of covariates".** In: *Biometrika* 108.4, pp. 815–828.

Lin, Winston (2013). **"Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique".** In.

Moghadas, Seyed M et al. (2021). **"The impact of vaccination on coronavirus disease**