Bayesian Federated Inference

A statistical approach

Marianne Jonker, Radboudumc, Nijmegen
In collaboration with: Hassan Pazira, Emanuele Massa, **Zoë van den Heuvel**, Ton Coolen

Radboudumc
university medical center

Challenge

Aim:

Datasets at multiple locations are available for analysis.

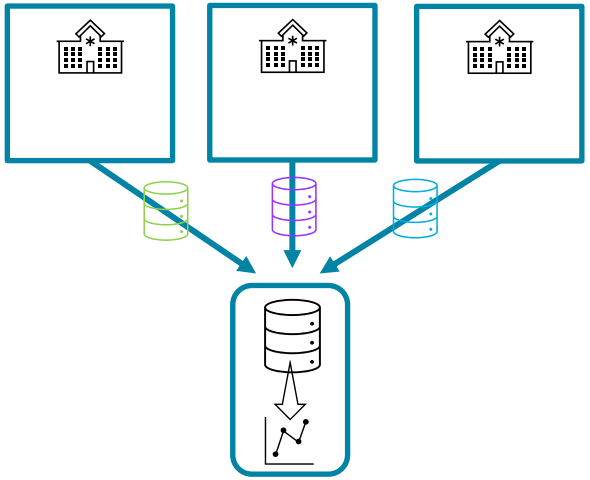
The aim is to do a statistical analysis using all data.

Solution:

- Data sharing
- Use one dataset only
- Federated Learning (machine learning community)
- Federated Inference (statistical community)

Data sharing

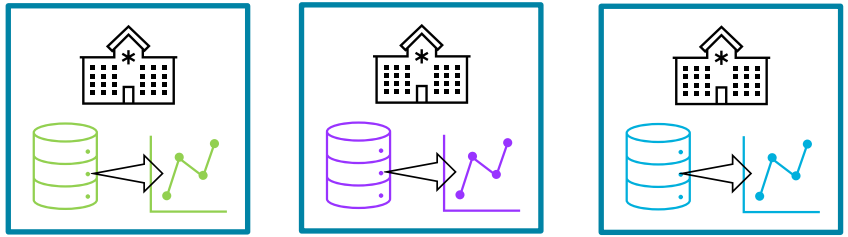
Each location sends their local data to a central server where the datasets are combined for statistical analysis.



Problems:

- Privacy of the patients
- Data sharing agreements

Separate Analyses



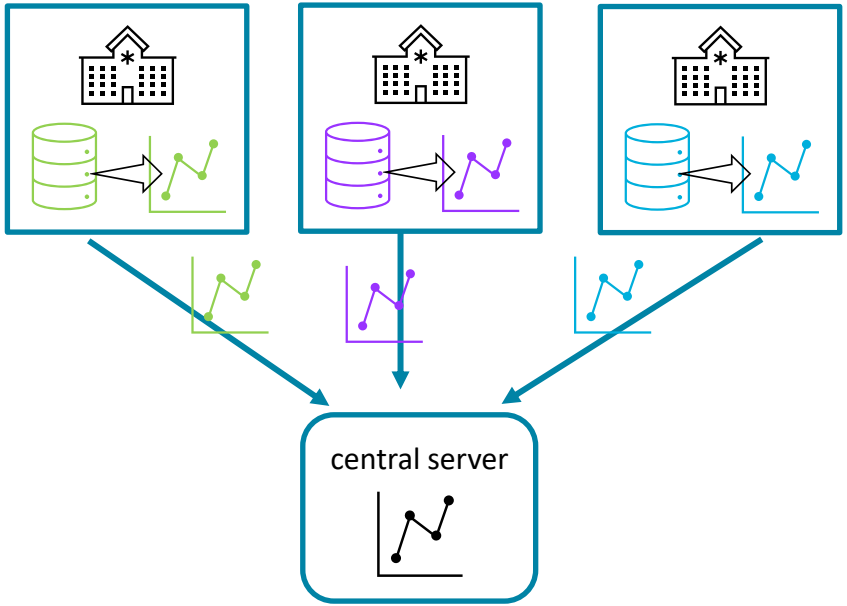
Loss of information, power, generability

Federated Learning

Federated Learning strategies have been around for some time, but it was not developed with medical applications in mind.

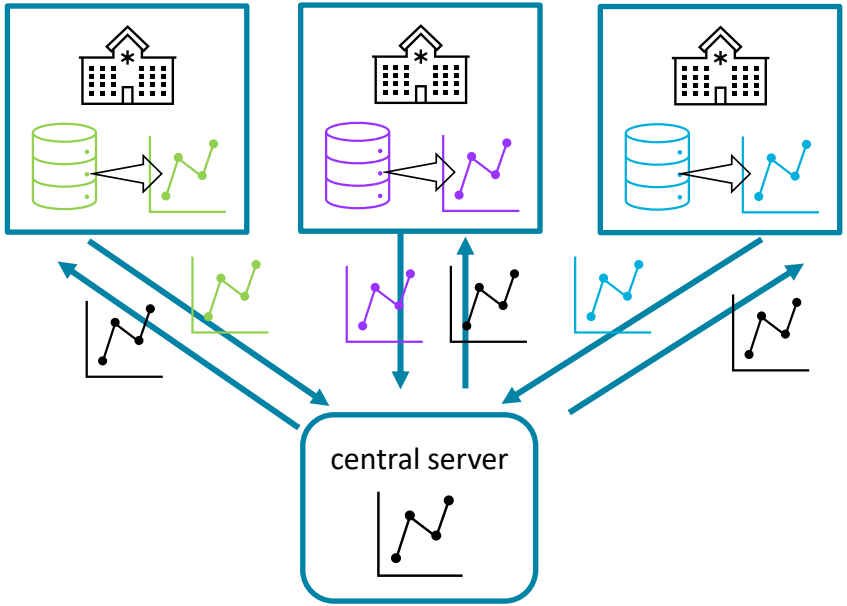
As more medical data becomes available and the need for privacy-preserving methods grows, federated learning methods are increasingly finding its way into healthcare.

Federated Learning



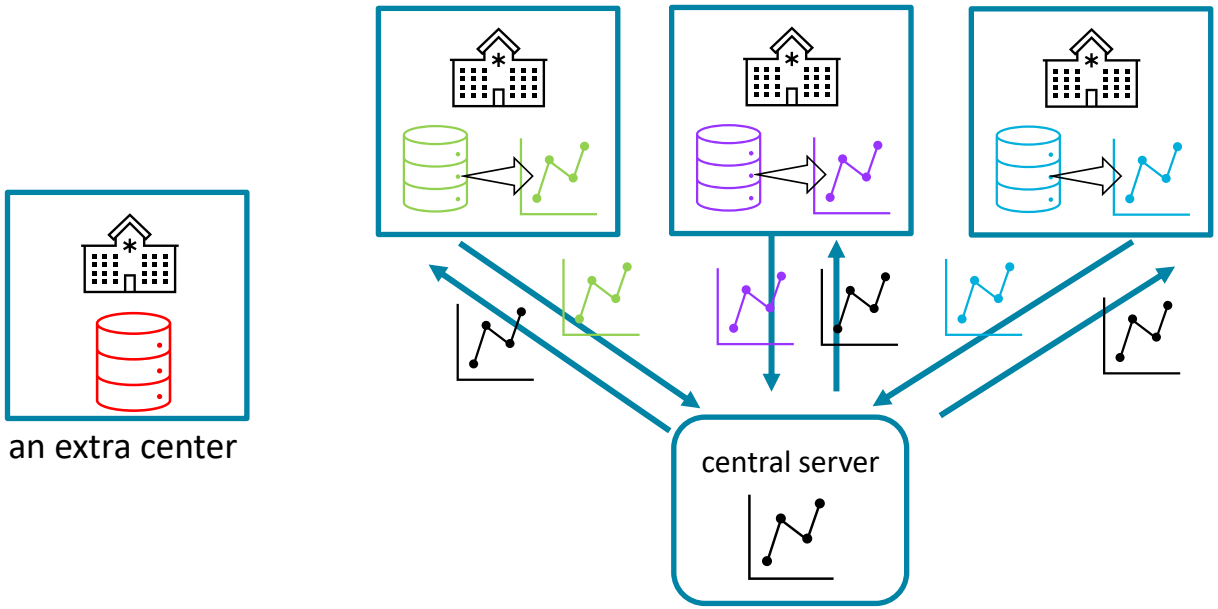
Iterative algorithm

Federated Learning



Iterative algorithm

Federated Learning



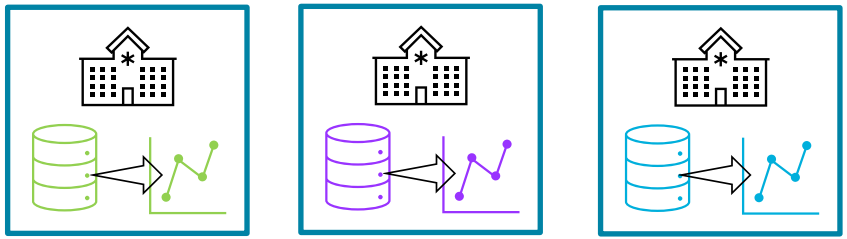
Federated Learning

Especially the fact that the algorithm works cyclic, can be a disadvantage:

- Has to be done in a semi-automatic way (cross firewalls)
- All centers need to have the data + software available at the same time
- The estimates cannot be easily updated if one center joins later.

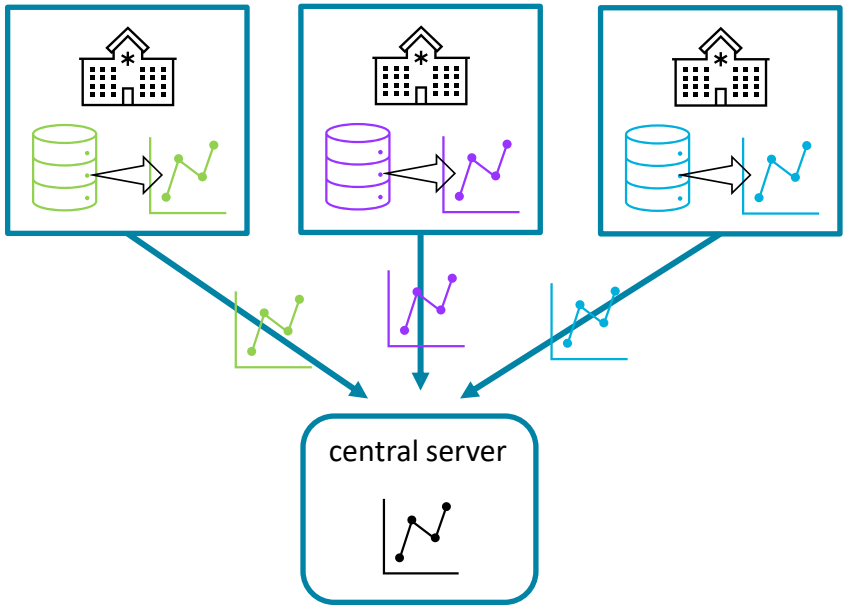
Federated Inference

our (statistical) proposal



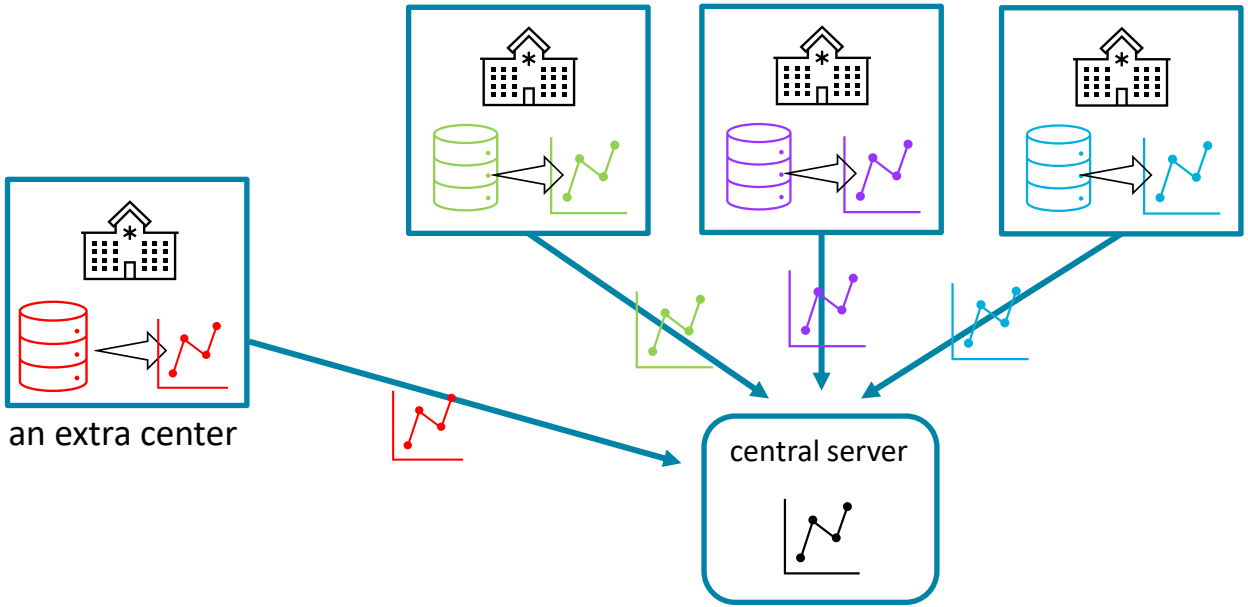
Federated Inference

our (statistical) proposal



Federated Inference

our (statistical) proposal



Federated Inference in steps

1. Decide which statistical model should be estimated to answer the research question.
2. In each medical center the same statistical analysis is performed.
3. The inference results are sent to the central server.
4. In the central server the results are aggregated to estimate the global statistical model: A weighted average of the local estimates is computed. Weights depend on the local sample sizes and accuracy of the estimates

With this method the final results are **approximately** equal to that what would have been found if the analysis was performed on the combined data

Example, a pilot study

1. Swedish MS registry split in two parts: (1) Stockholm, (2) rest of Sweden.
2. Outcome: SPMS (Secondary progressive MS) vs RRMS (Relapsing Remitting MS).
3. Variables: age, gender, location (Stockholm, yes/no)
4. Estimate a logistic regression model that includes these three variables. Estimation should be done in a federated way

Example, a pilot study

	Estimate			SE		
	comb.	FedL	BFI	comb.	FedL	BFI
intercept	-7.620	-7.620	-7.571	0.124	0.124	0.124
age	0.114	0.114	0.113	0.002	0.002	0.002
sex	-0.125	-0.125	-0.127	0.047	0.047	0.047
Stockholm	0.225	0.225	0.228	0.049	0.049	0.050

FedL: Federated Learning strategy

SPMS: outcome = 1. RRMS: outcome = 0.
men: sex=0, women: sex=1.
patient in Stockholm region: Stockholm = 1
patient in the rest of the country: Stockholm = 0

Extensions

The methodology has been extended to deal with:

- Heterogeneous populations/models
- Clustering among locations
- Unobserved Features
- Many features
- Federated Causal inference (ongoing)
- Small Centers (ongoing)

Next Project

Study the effectiveness of rituximab in MS patients.

Observational Data:

multiple MS registries

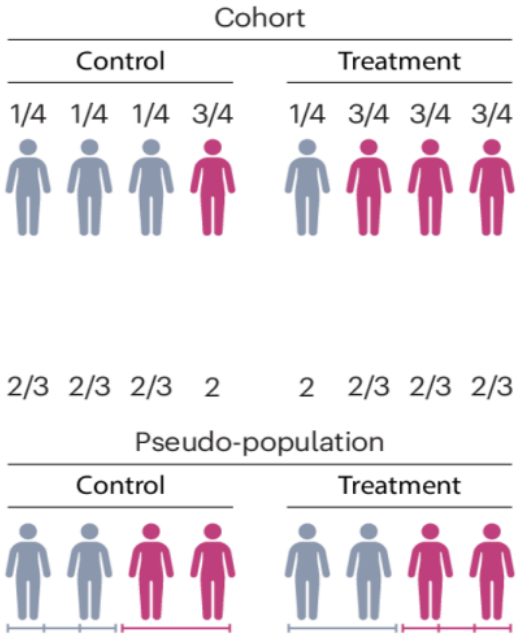
Statistical Methods for Federated Causal Inference:

1. propensity score model (for the probability to be in the treatment group)
2. weighted analysis

Statistical Analysis

Statistical Methods for Causal Inference:

1. propensity score model (for the probability to be in the treatment group)
2. weighted analysis

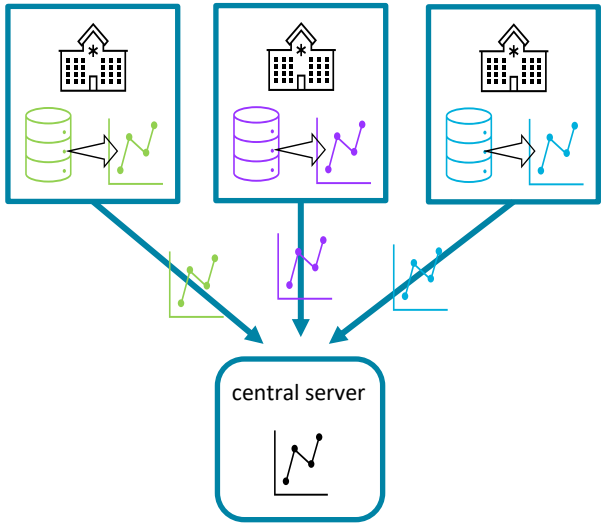


modified from: https://doi.org/10.1038/s41592-025-02629-y

Propensity Score Model

In every medical center a model for the propensity scores is estimated (logistic regression model, accounting for heterogeneity)

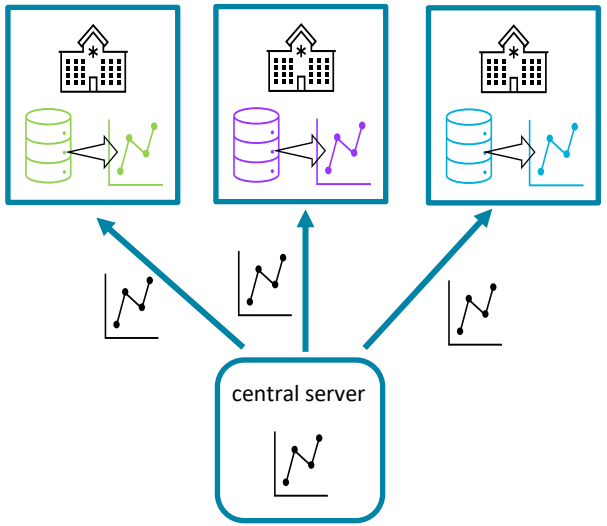
These models are combined in the central server with the BFI methodology



Send back

In every medical center a model for the propensity scores is estimated (logistic regression model, accounting for heterogeneity)

These models are combined in the central server with the BFI methodology

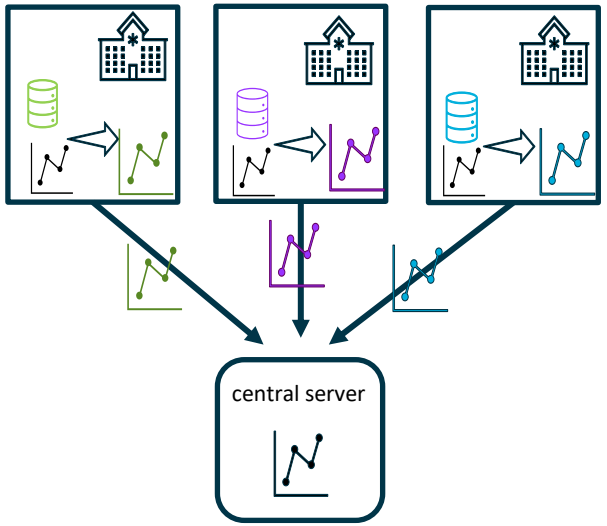


Weighted Analysis

In every medical center the weights are computed and a weighted analysis is performed.

The results are sent to the central server.

These models are combined in the central server with the BFI methodology



Note: extra pass needed

Comparison FL and FI

	Federated Learning	Federated Inference
Accuracy	exact	approximation
One pass / iterations	iterative algorithm	one pass
Updating	not possible	possible
Flexibility	restricted	many analyses possible
Automatic / hands-on	semi-automatic	hands-on
ICT + security	dealing with firewalls	E-mail with numbers
Software	multiple possible	R-software
Many researchers	possible	more difficult

Conclusion

Which method (FL or FI) should we use?

Depends on the aim, setting and the research questions.

References

- Jonker MA, Pazira H, Coolen ACC. Bayesian Federated Inference for estimating StatisticalModels based on non-shared multicenter data. Statistics in Medicine 43 (12), 2421-2438.2024. https://doi.org/10.1002/sim.10072

- Jonker MA, Pazira H, Coolen ACC. Bayesian Federated Inference for regression modelsbased on non-shared medical center data. Research Synthesis Methods, Volume 16 (2),2025, 383 – 423 DOI: https://doi.org/10.1017/rsm.2025.6

- Pazira H, Massa E, Weijers JAM, Coolen ACC, Jonker MA. Bayesian Federated Inference for Survival Models. Journal of Applied Statistics. 2025 DOI:10.1080/02664763.2025.2511932

- Pazira H, Massa E, Jonker MA. R-package BFI and users manual, 2024.
  Number of downloads: > 9400 since its release June 2024.

- Massa E, Jonker M.A. Federated Inference with missing features. Conference paper:
  The International Conference on Federated Learning Technologies and Applications