Predictive Modelling Under Clinical Presence

Vincent Jeanselme

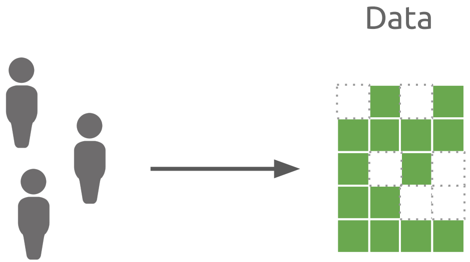
MRC Biostatistics Unit
University of Cambridge

Nokia Bell Labs- Responsible AI seminar series

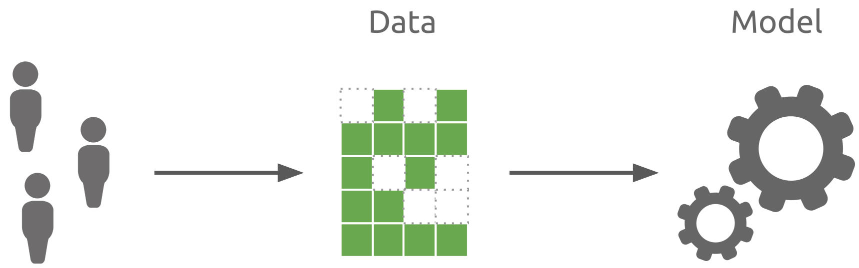
19.10.2023

Clinical Modelling

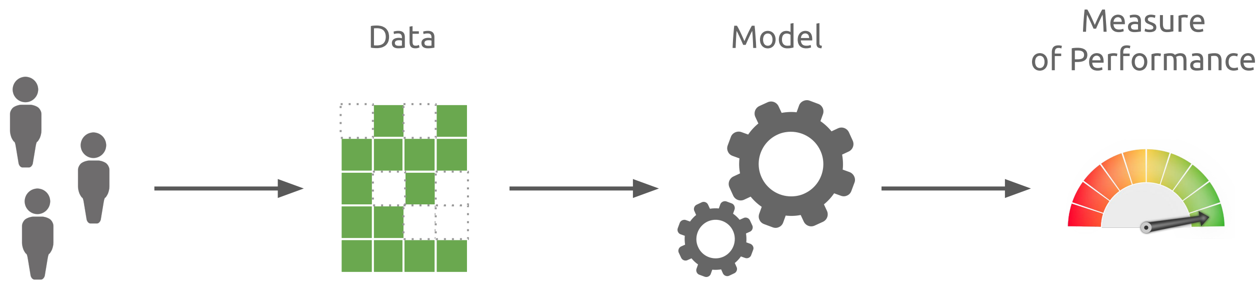
Clinical Modelling



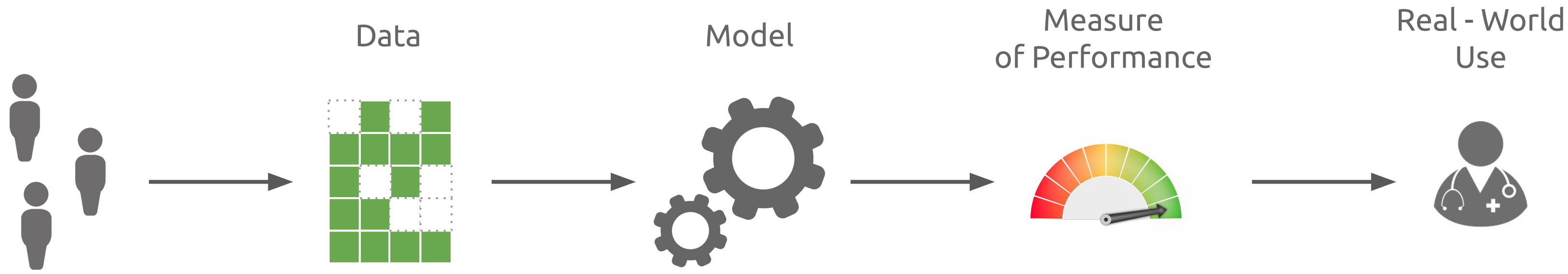
Clinical Modelling



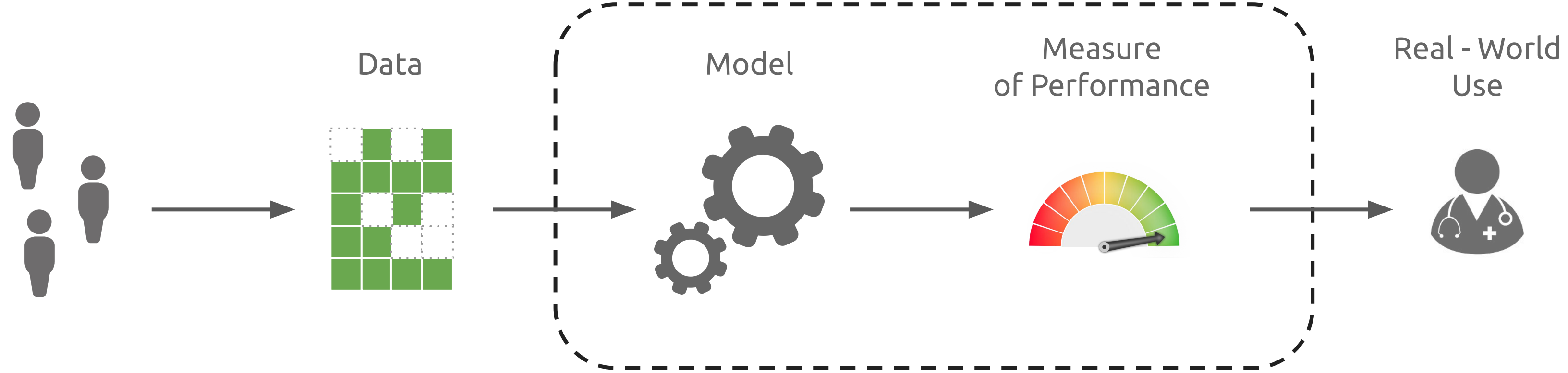
Clinical Modelling



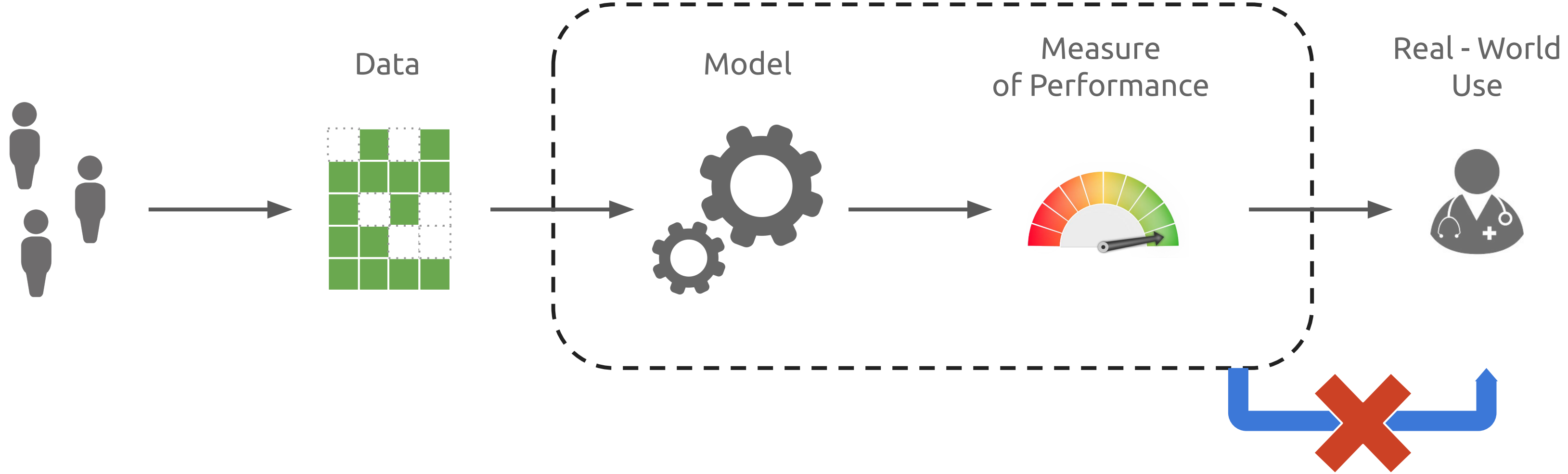
Clinical Modelling



Clinical Modelling

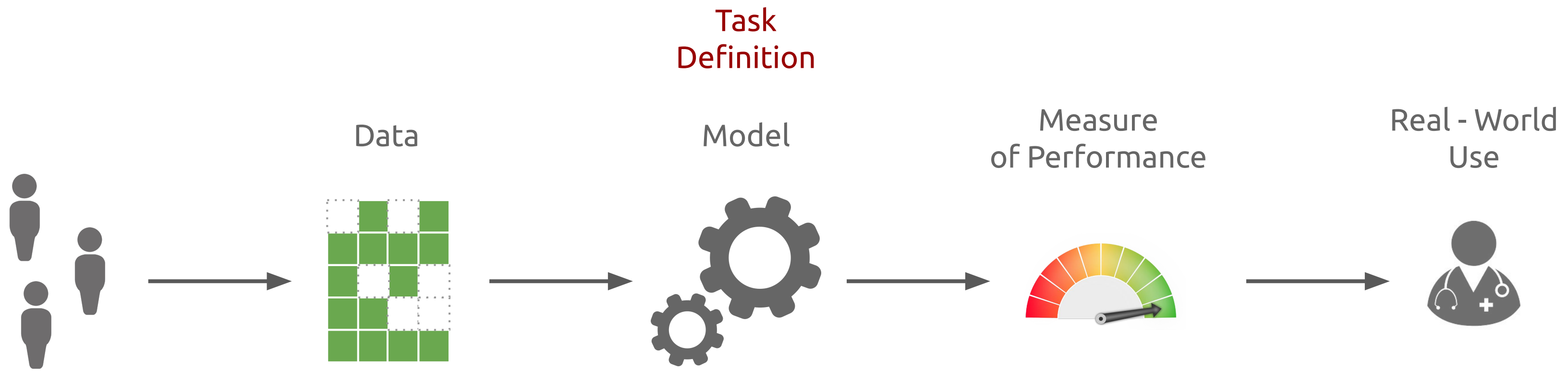


Clinical Modelling



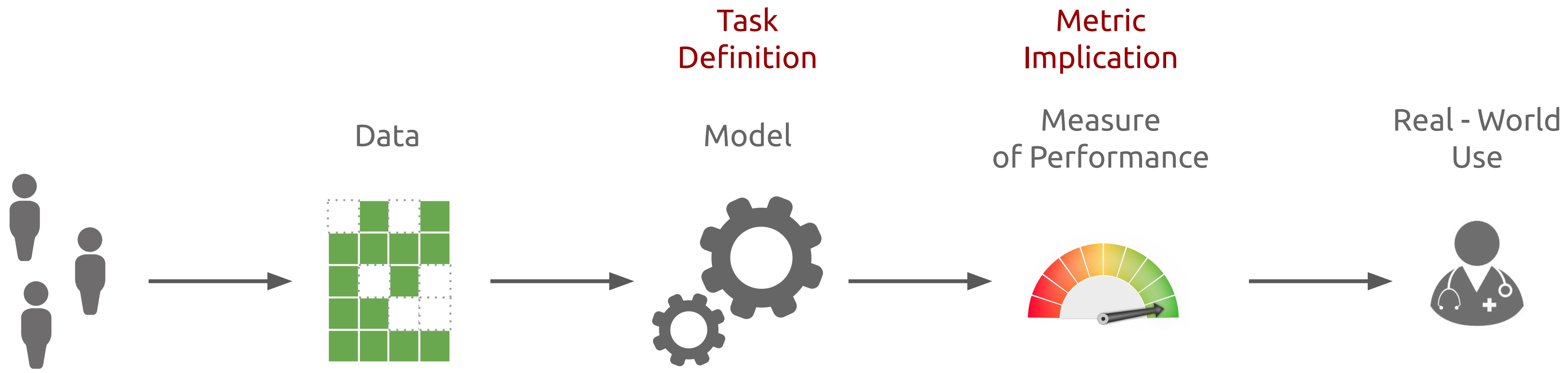
Why does it not transfer to real settings ?

Clinical Modelling



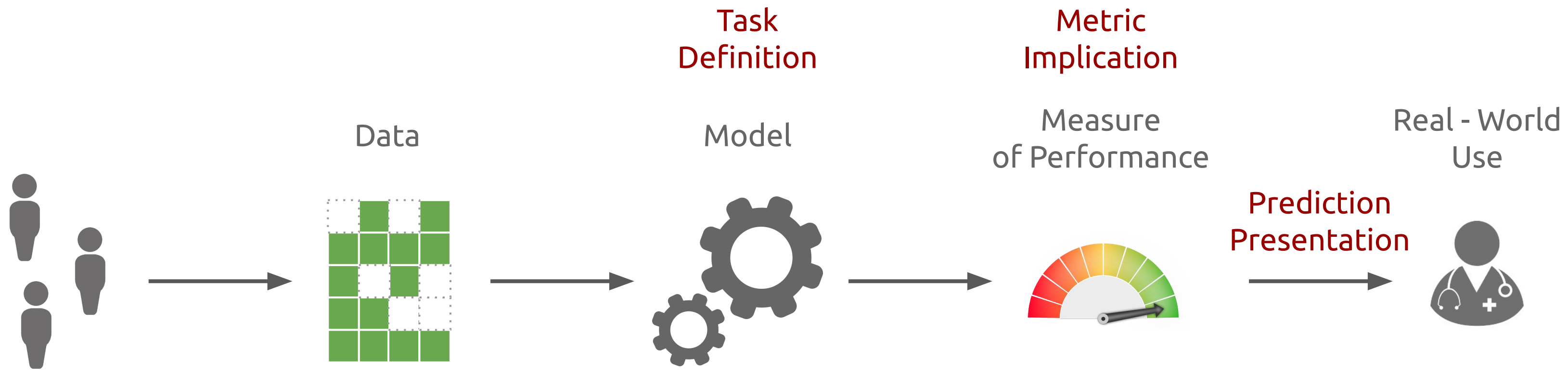
Why does it not transfer to real settings ?

Clinical Modelling



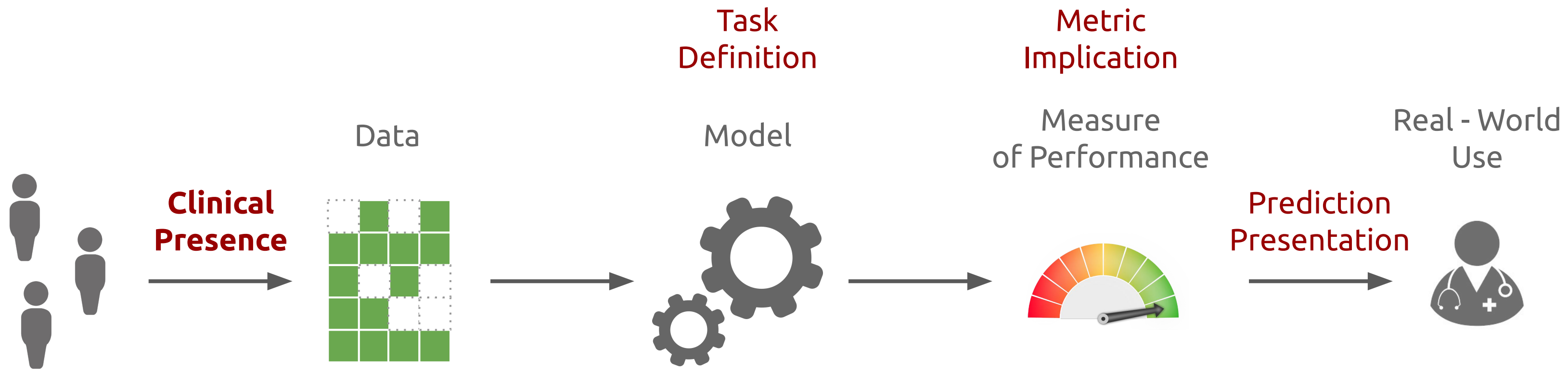
Why does it not transfer to real settings ?

Clinical Modelling



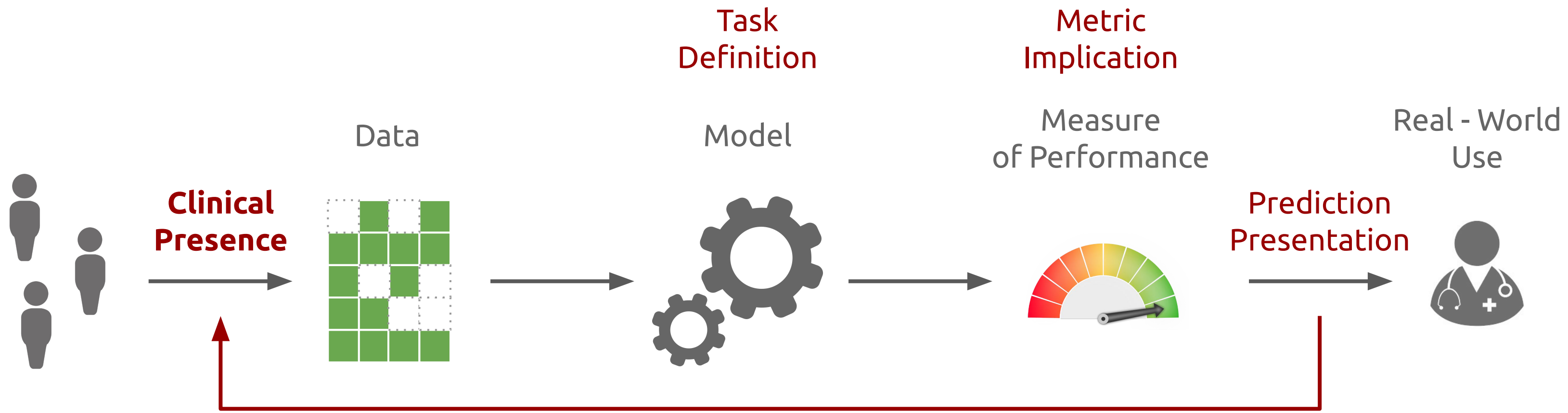
Why does it not transfer to real settings ?

Clinical Modelling



Why does it not transfer to real settings ?

Clinical Modelling

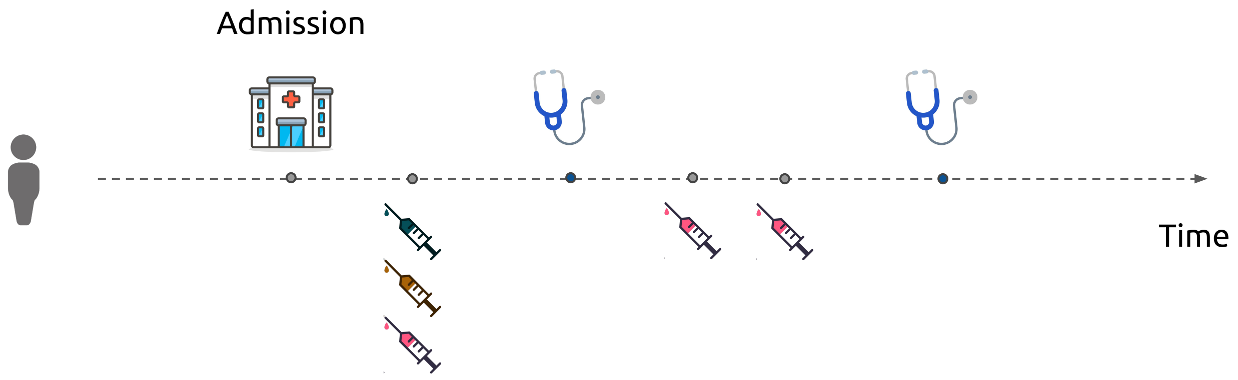


Why does it not transfer to real settings ?

Clinical Presence

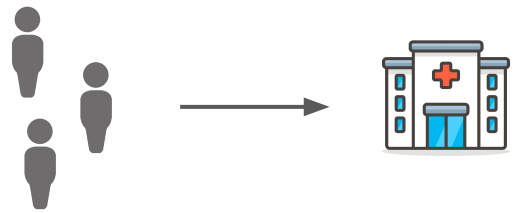
Clinical Presence

Clinical Presence

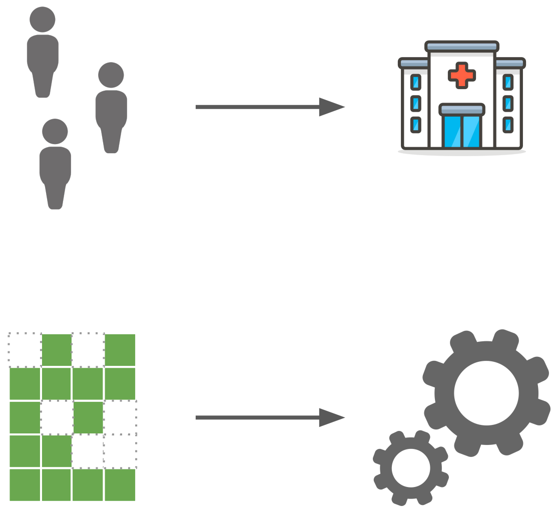


*The **observation process** is imprinted by the interaction between **patients** and the **healthcare system**.*

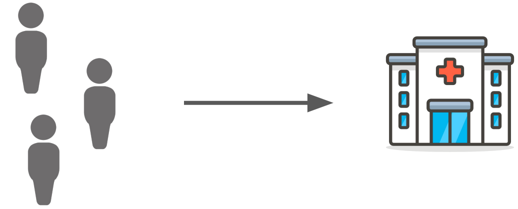
Risk of Clinical Presence

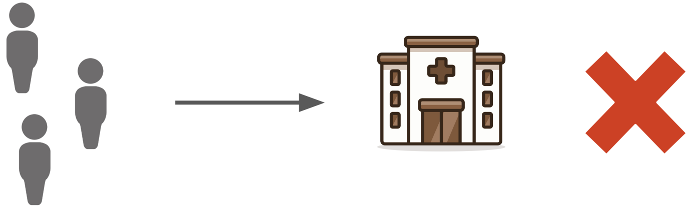


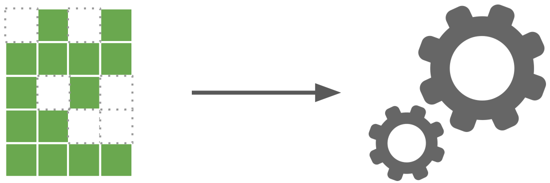
Risk of Clinical Presence



Risk of Clinical Presence

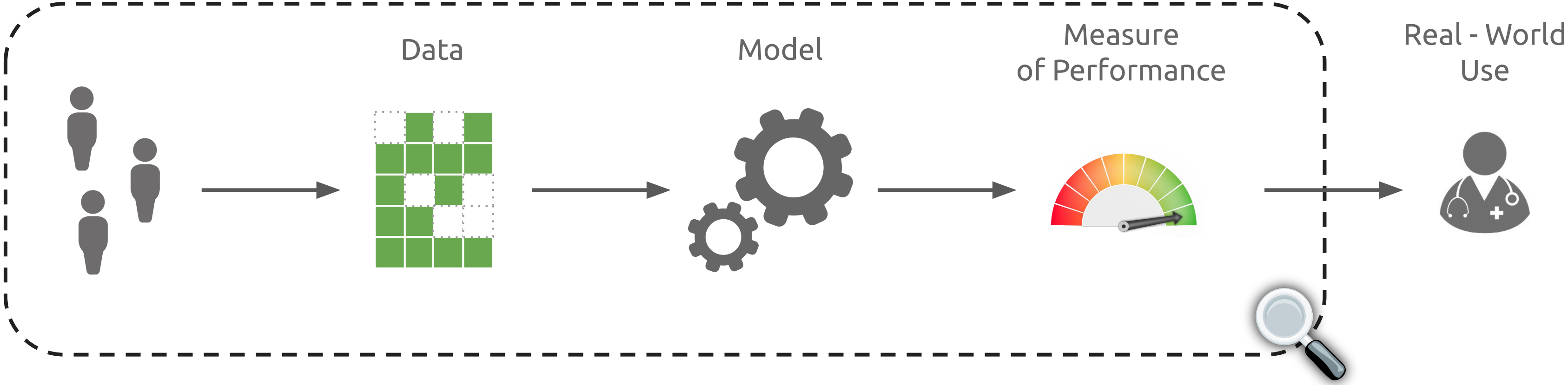






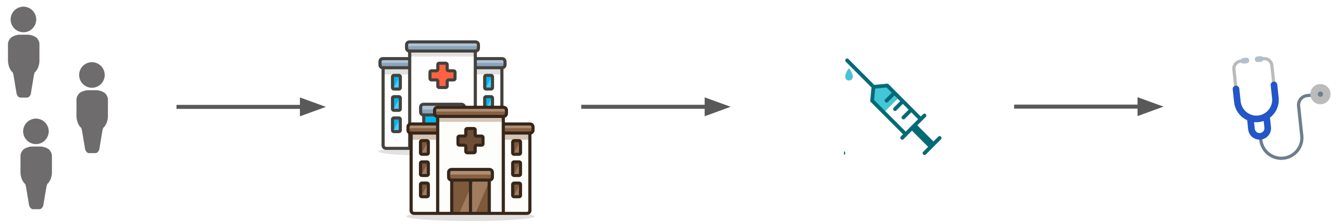
Clinical Presence presents a risk for the **generalisability** of machine learning solutions.

Clinical Modelling

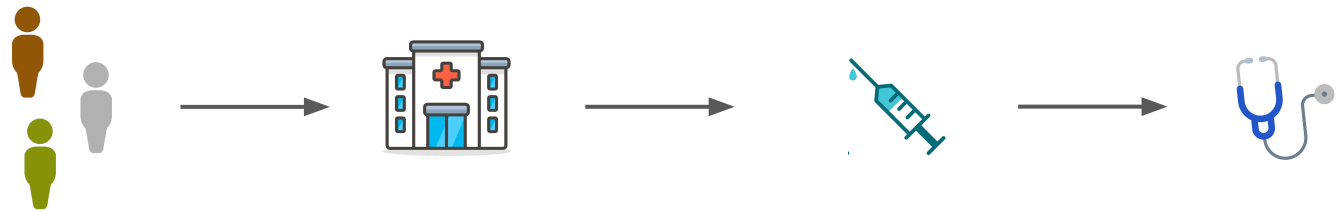


Jeanselme, V., Martin, G., Peek, N., Sperrin, M., Tom, B., & Barrett, J. (2022).
Deepjoint: Robust survival modelling under clinical presence shift.
NeurIPS 2022 Workshop on Learning from Time Series for Health.

Risk of Clinical Presence



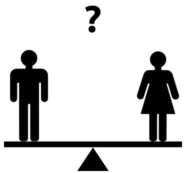
Risk of Clinical Presence



*What happens under **group-specific** patterns ?*

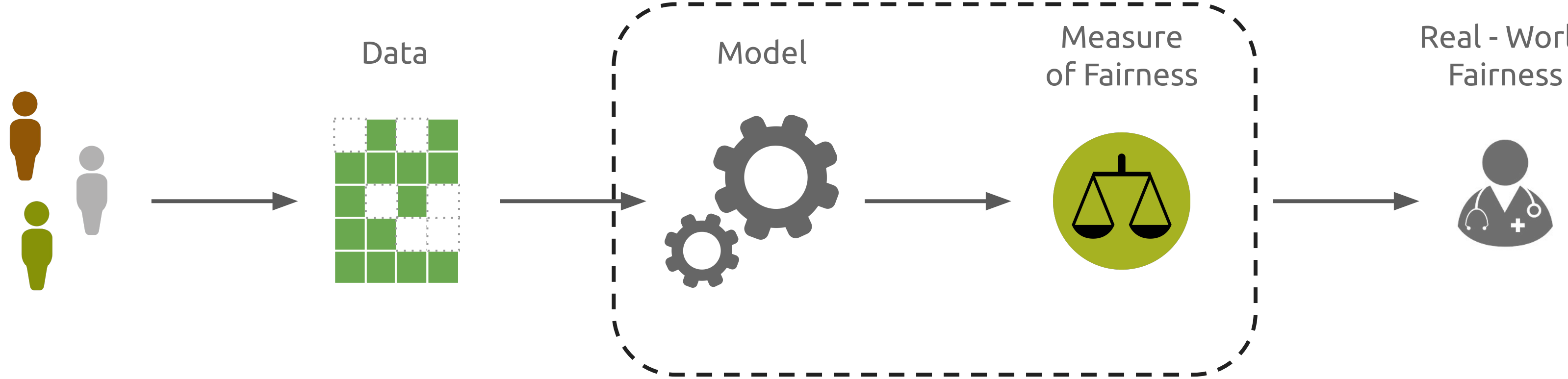
Jeanselme, V., De-Arteaga, M., Zhang, Z., Barrett, J., & Tom, B. (2022).
Imputation Strategies Under Clinical Presence: Impact on Algorithmic Fairness.
In Machine Learning for Health (pp. 12-34). PMLR. - Journal version under review in Management Science.

Algorithmic Fairness



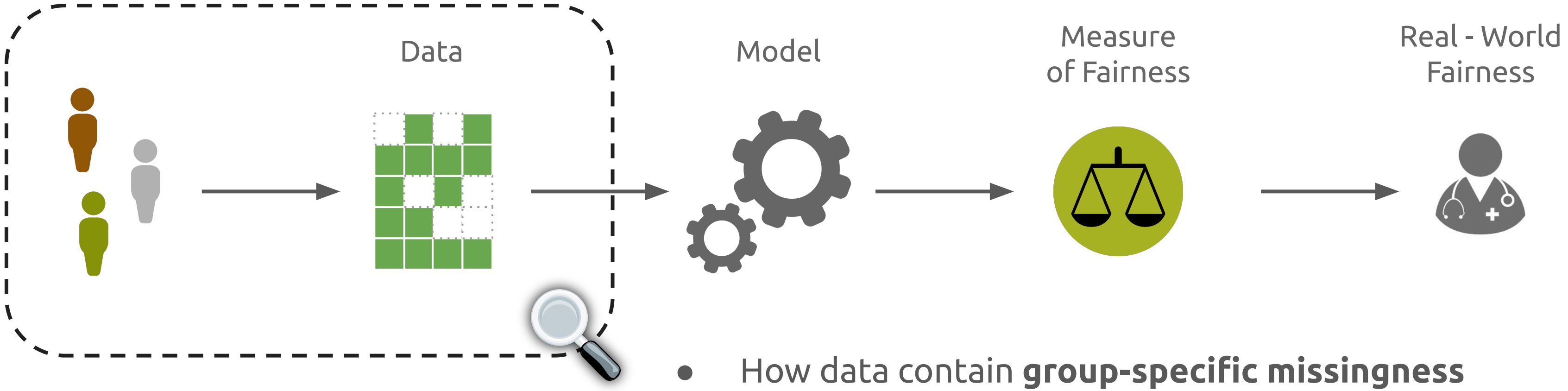
*Our work focuses on **group fairness**, measured through **equal performance across groups**, i.e. a pipeline is fairer than another with regard to a group if its performance gap is the smallest.*

Fairness Pipeline



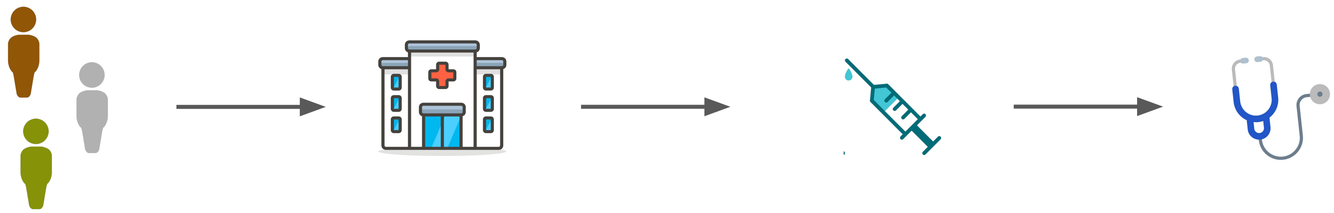
The fairness literature studies how to ***detect*** *and* ***mitigate biases*** *present the data. Current focus has been on* ***modelling*** *choices' consequences on algorithmic fairness.*

Impact of imputation on algorithmic fairness



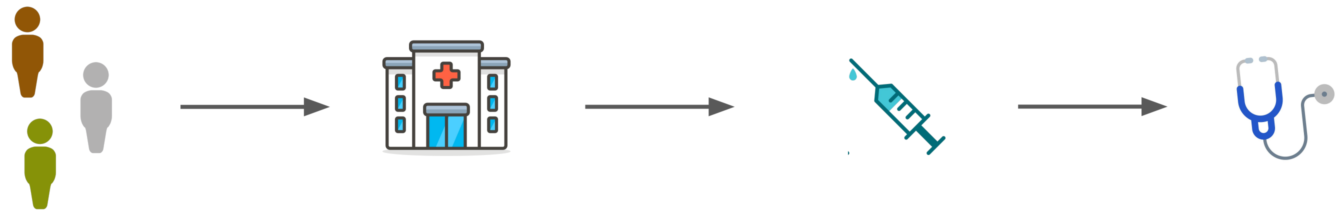
- How data contain **group-specific missingness patterns** ?
- How does imputation affects **downstream** algorithmic **fairness** ?

Identified Clinical Missingness Patterns



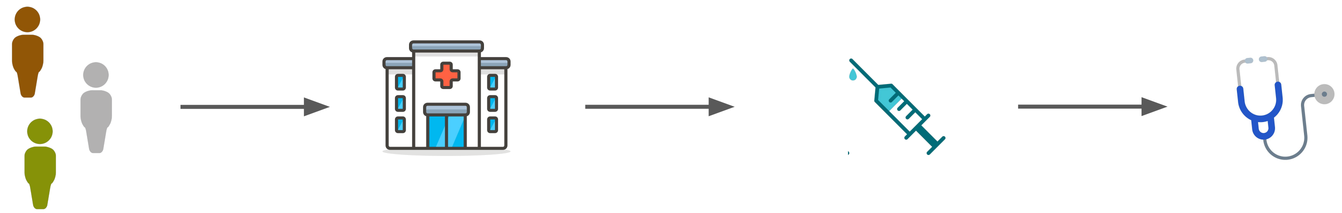
Limited access to quality care
Resulting in more missing data for some patients, e.g. as a result of less follow-up visits.

Identified Clinical Missingness Patterns



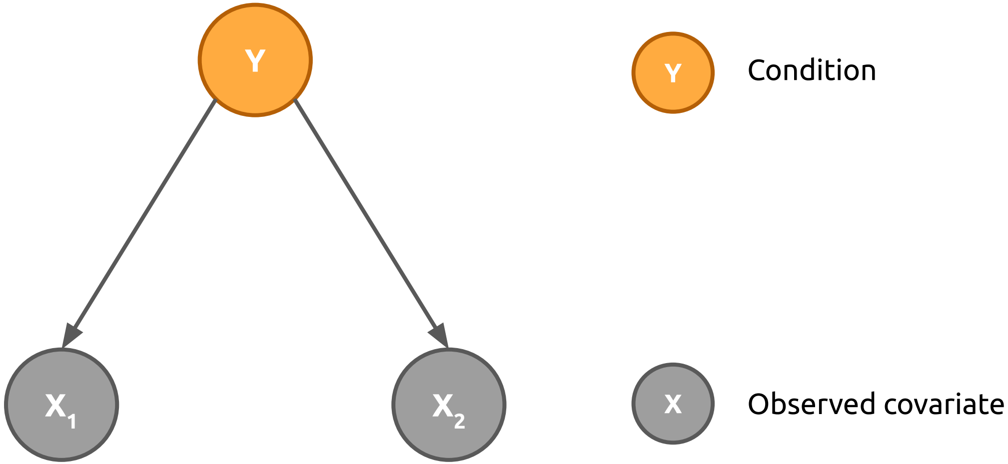
(Mis)-informed collection
Concentration of missing data for patients who do not present “standard” symptoms that trigger questions or laboratory tests.

Identified Clinical Missingness Patterns

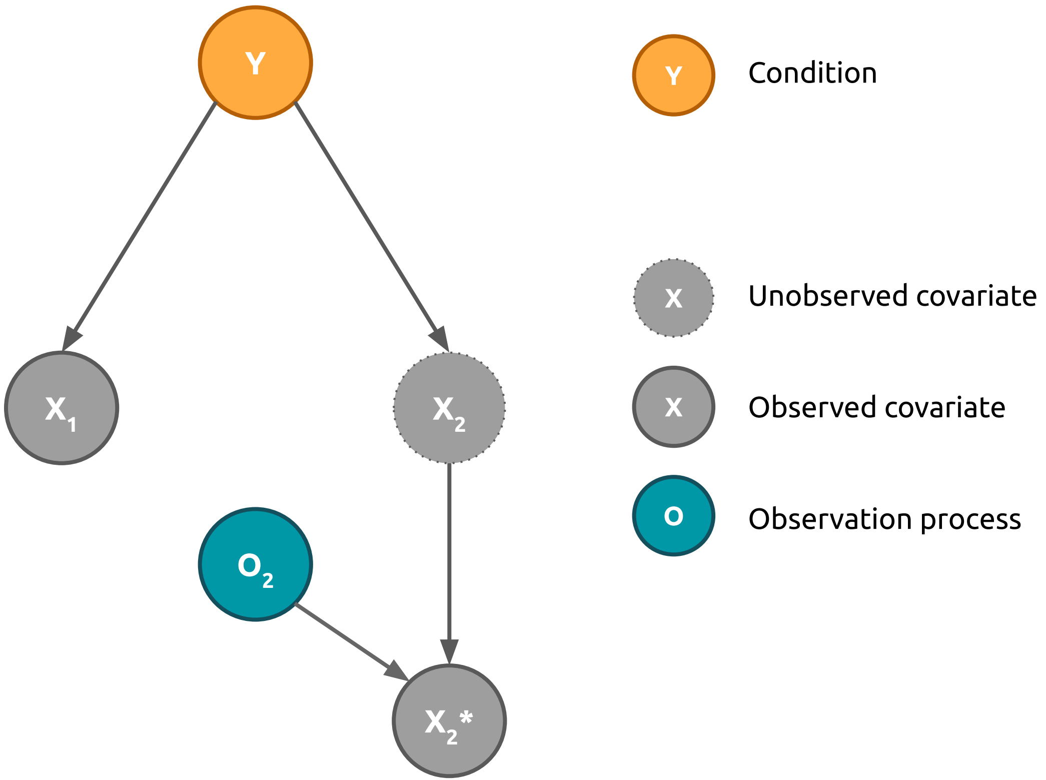


Confirmation bias
Information is only collected when a condition is expected.

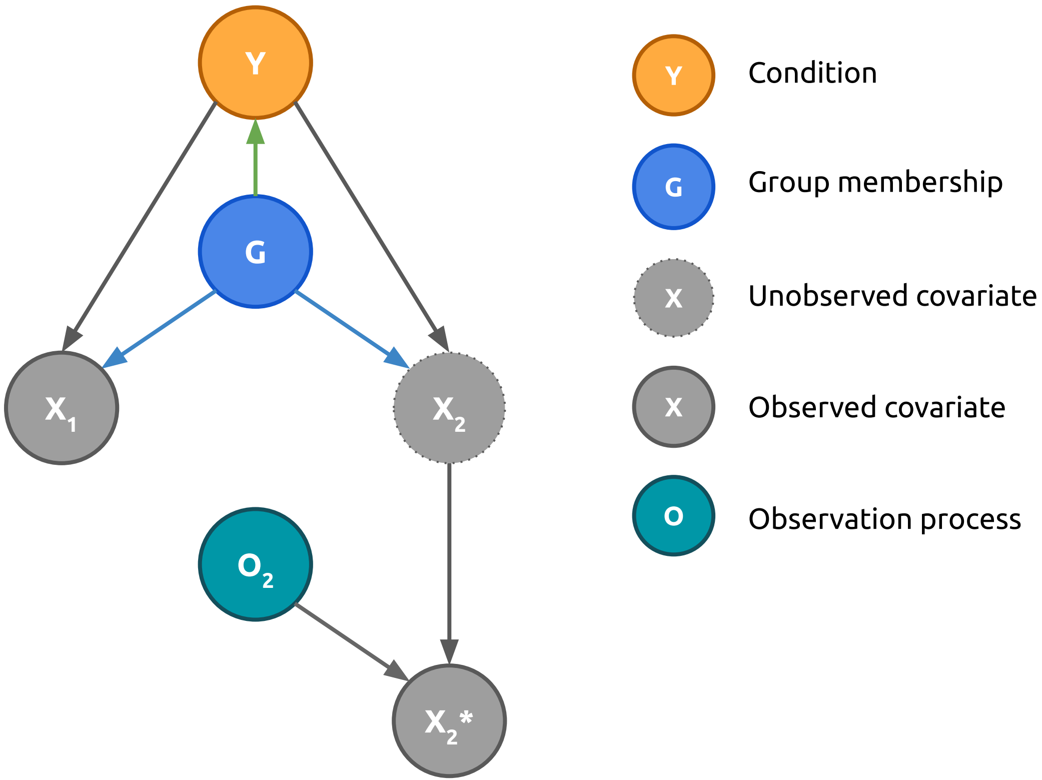
Formalisation



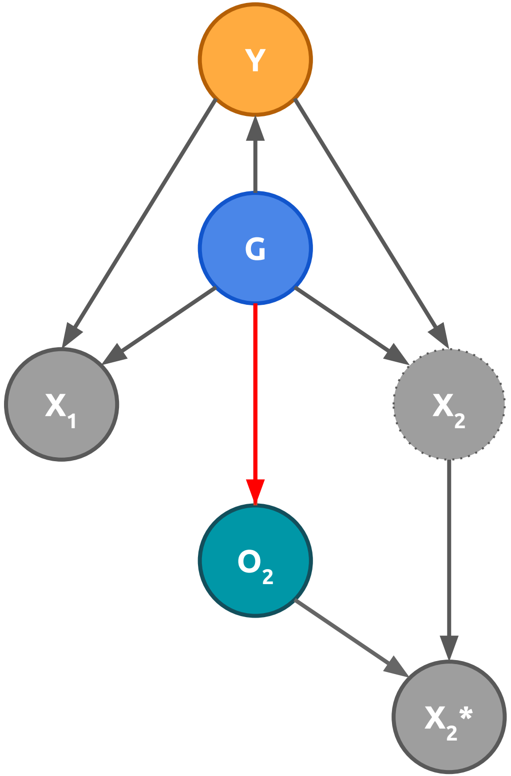
Formalisation



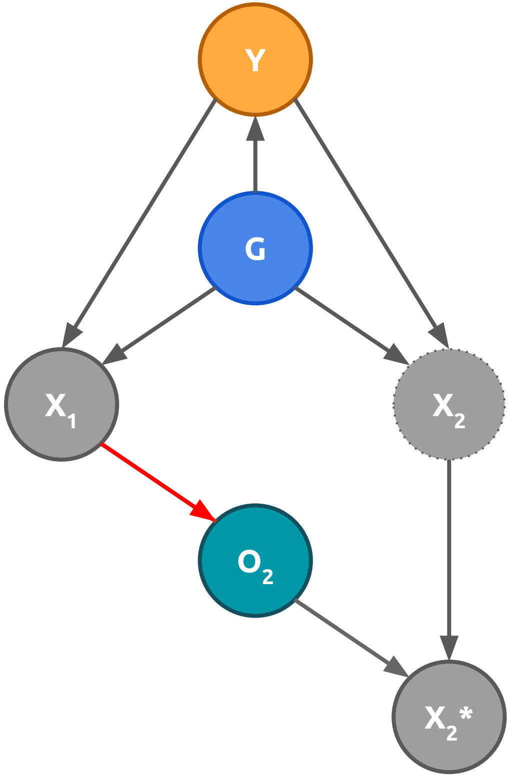
Formalisation



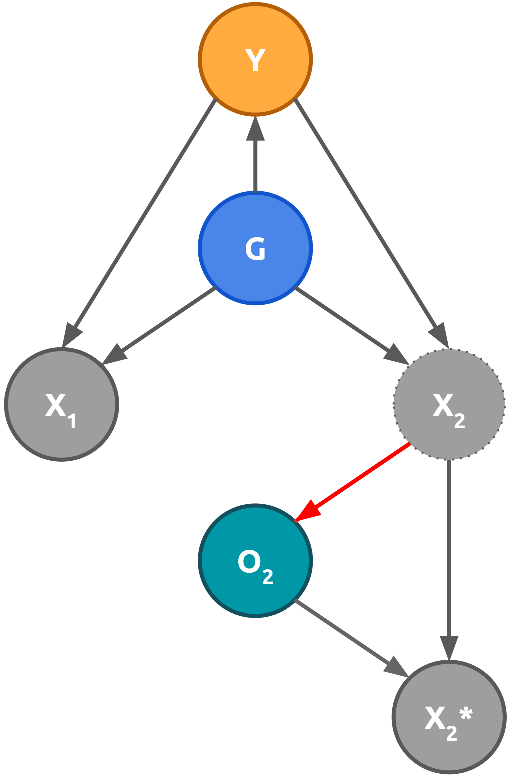
Limited access to quality care



(Mis)-informed collection



Confirmation bias

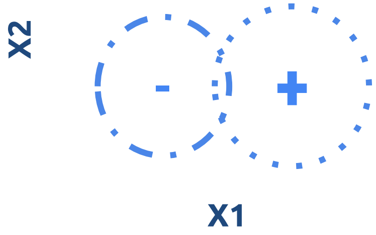


Simulations

Outcome | **Group**

Positive

Negative | Marginalised



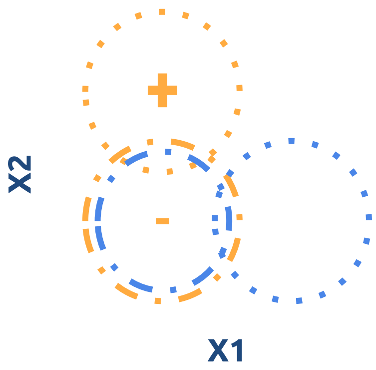
Ground Truth

Simulations

Outcome **Group**

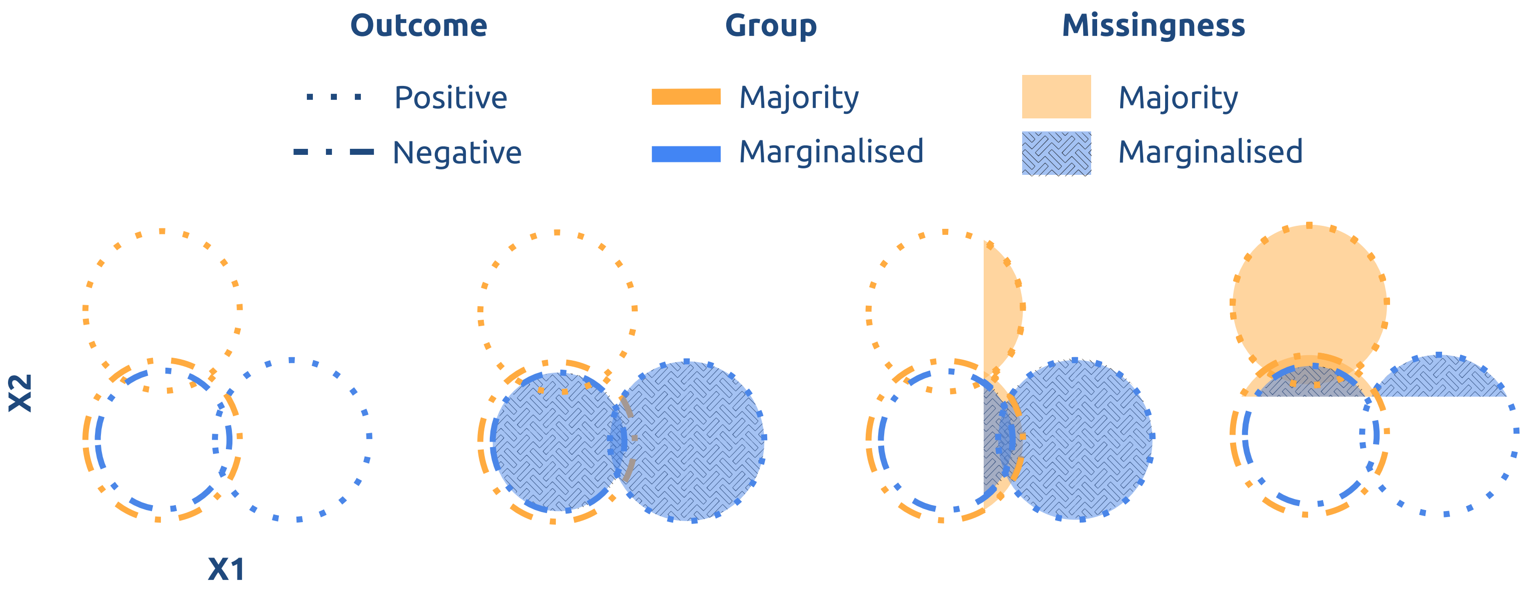
Positive — Majority

Negative — Marginalised



Ground Truth

Simulations



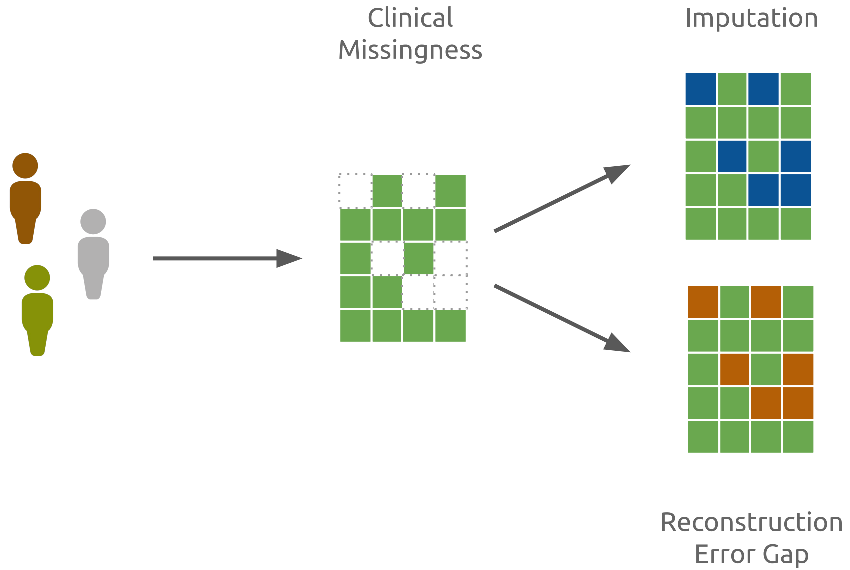
Ground Truth

(S1)
Limited access
to quality care

(S2)
(Mis)-informed
collection

(S3)
Confirmation
bias

Reconstruction Error



Imputations

➔ **Single mean imputation** (Population Mean) - Missing data are replaced by the population mean.

Imputations

➔ **Single mean imputation** (Population Mean) - Missing data are replaced by the population mean.

➔ **Group mean imputation** - Missing data are replaced by the group-specific mean.

Imputations

- **Single mean imputation** (Population Mean) - Missing data are replaced by the population mean.
- **Group mean imputation** - Missing data are replaced by the group-specific mean.
- **Multiple Imputation using Chained Equation** (MICE) - Missing covariates are iteratively drawn from a regression model built over all other available covariates with median initialisation.
- **Group MICE** - Group membership is added to render the MAR assumption more plausible

Imputations

➔ **Single mean imputation** (Population Mean) - Missing data are replaced by the population mean.

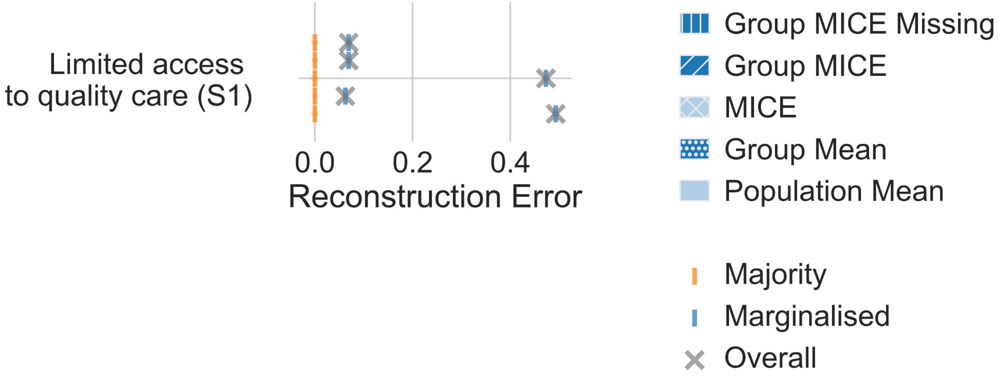
➔ **Group mean imputation** - Missing data are replaced by the group-specific mean.

➔ **Multiple Imputation using Chained Equation** (MICE) - Missing covariates are iteratively drawn from a regression model built over all other available covariates with median initialisation.

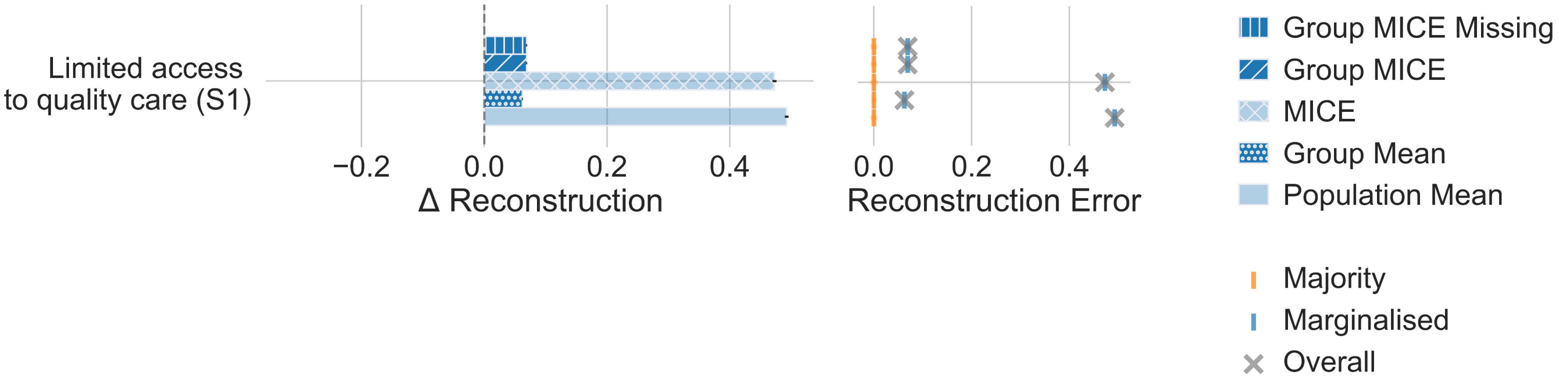
➔ **Group MICE** - Group membership is added to render the MAR assumption more plausible

➔ **Group MICE Missing** - Missingness indicators are concatenated to the input data to leverage informative missingness.

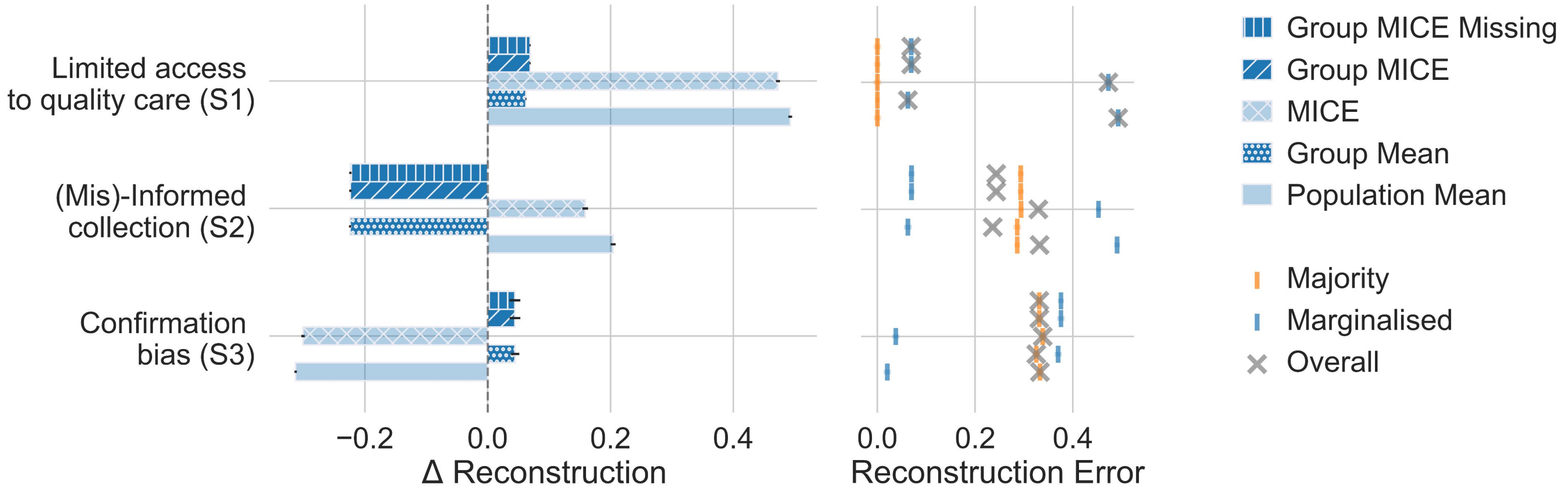
Reconstruction error



Reconstruction error



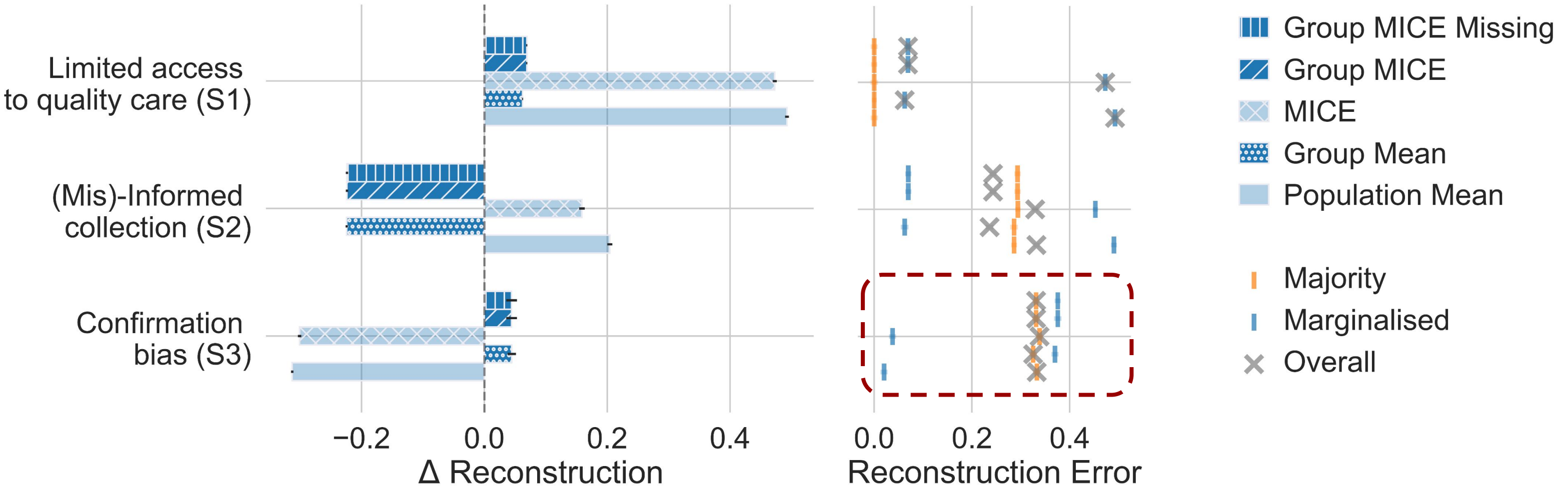
Reconstruction error



Reconstruction error

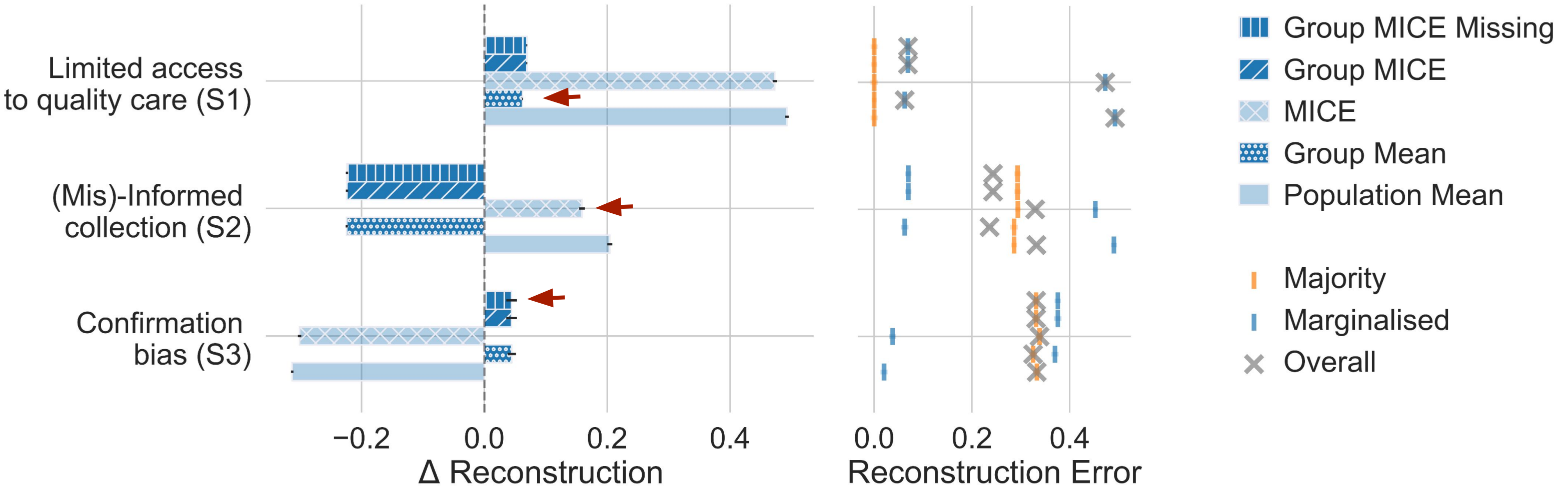
1

Different imputation strategies may have equal reconstruction errors at the population level while having different group reconstruction gaps.



Reconstruction error

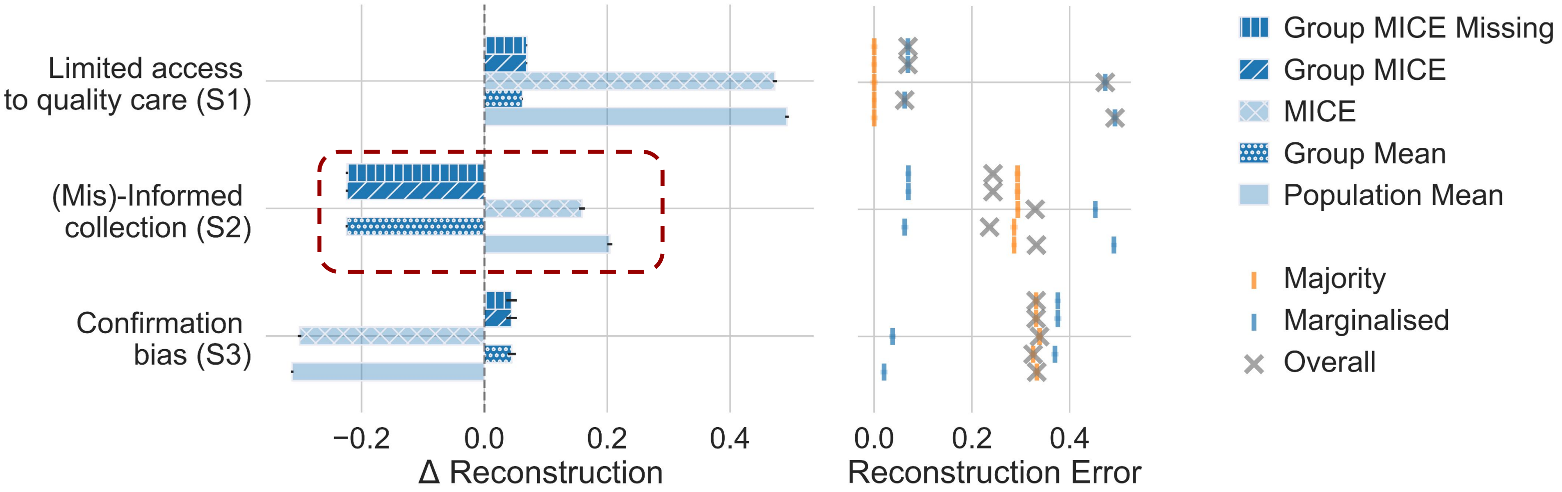
1 Different imputation strategies may have equal reconstruction errors at the population level while having different group reconstruction gaps.



2 No imputation strategy consistently outperforms the others across clinical presence scenarios.

Reconstruction error

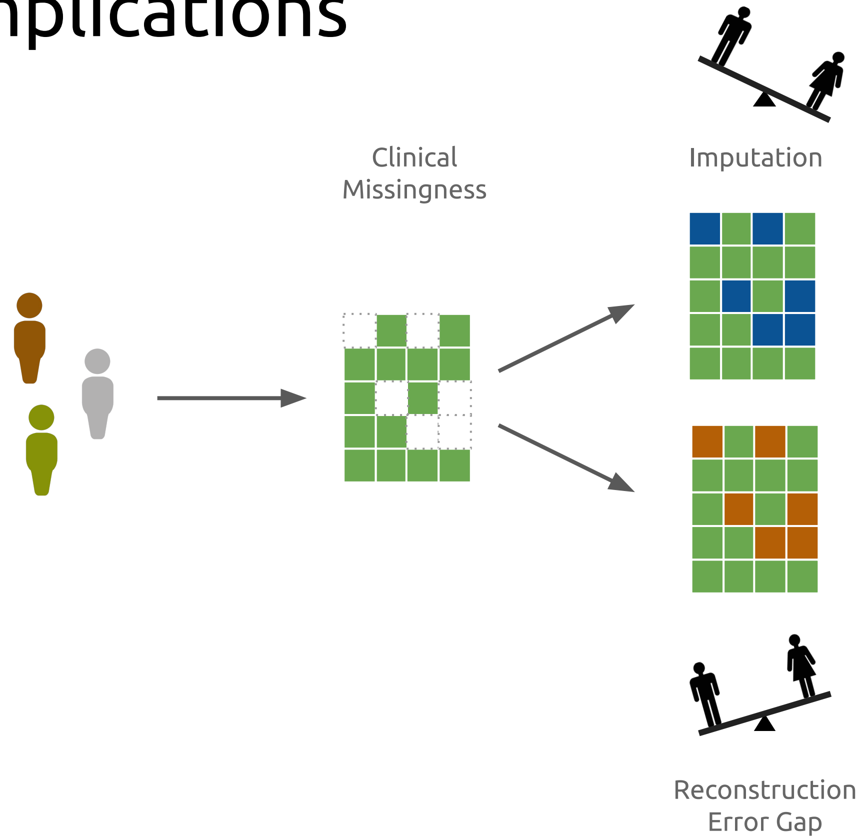
1 Different imputation strategies may have equal reconstruction errors at the population level while having different group reconstruction gaps.



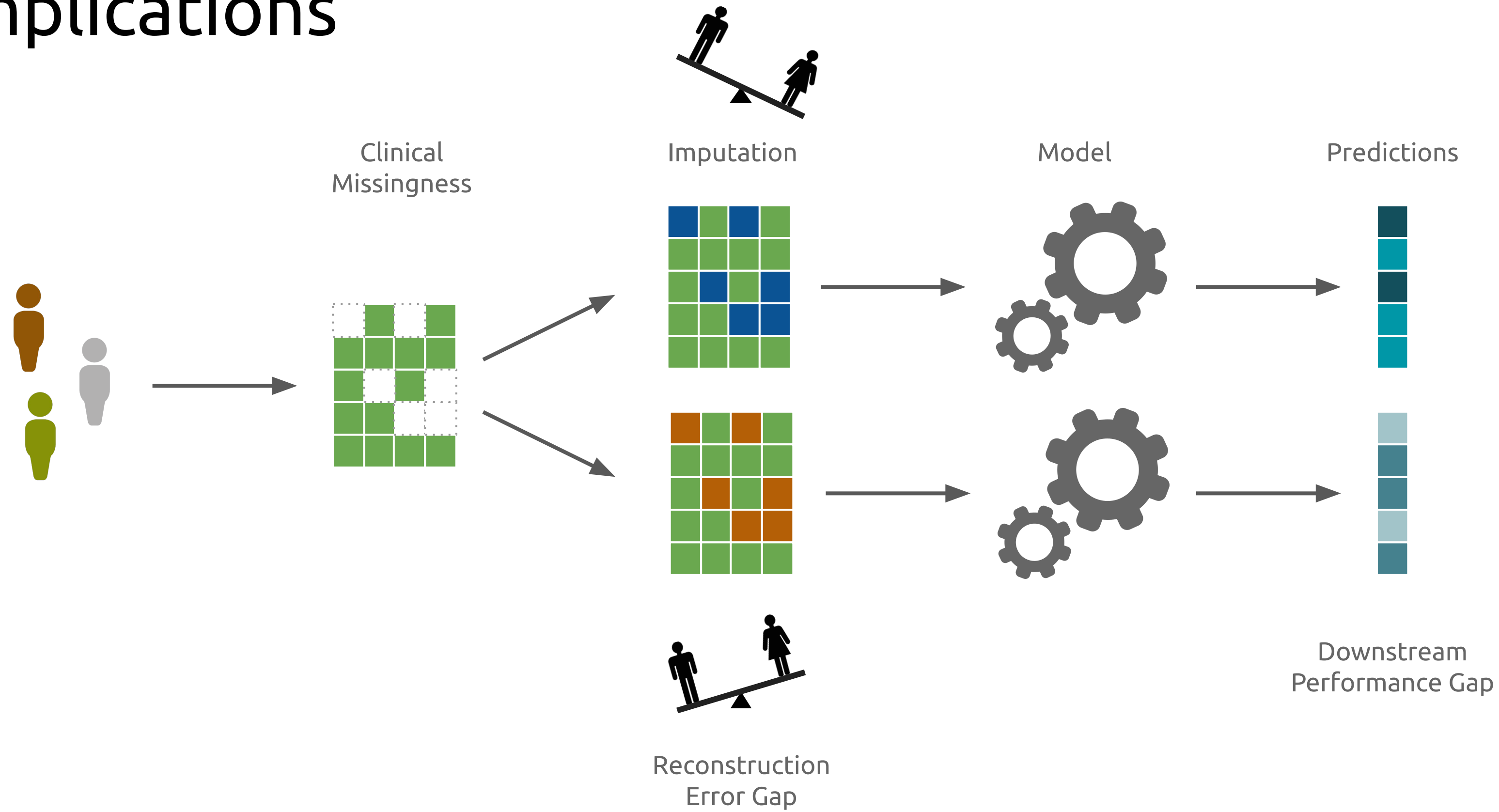
3 Current recommendations for group-specific imputation can increase the reconstruction gap and yield a worse reconstruction error for the marginalised group

2 No imputation strategy consistently outperforms the others across clinical presence scenarios.

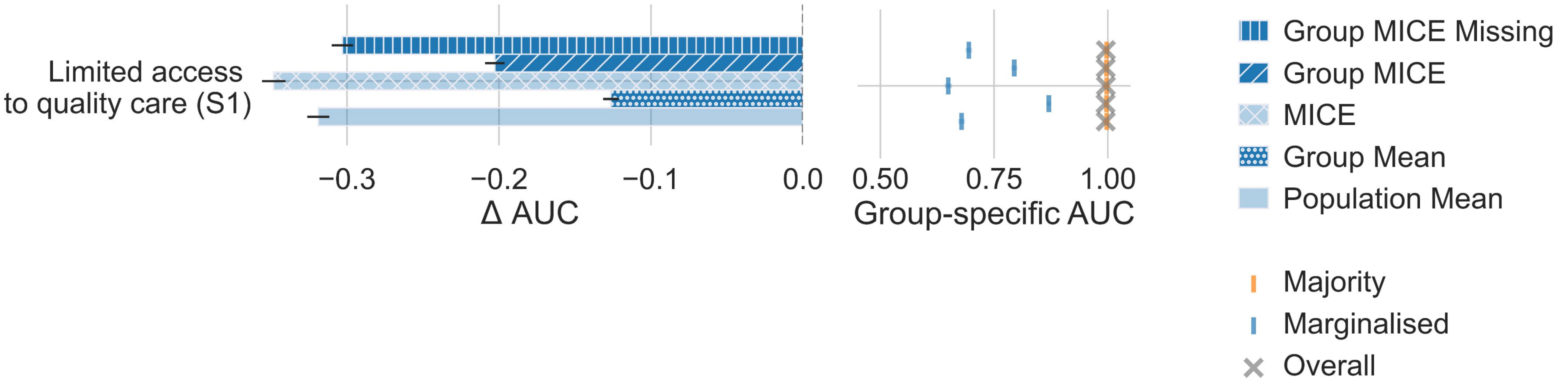
Implications



Implications

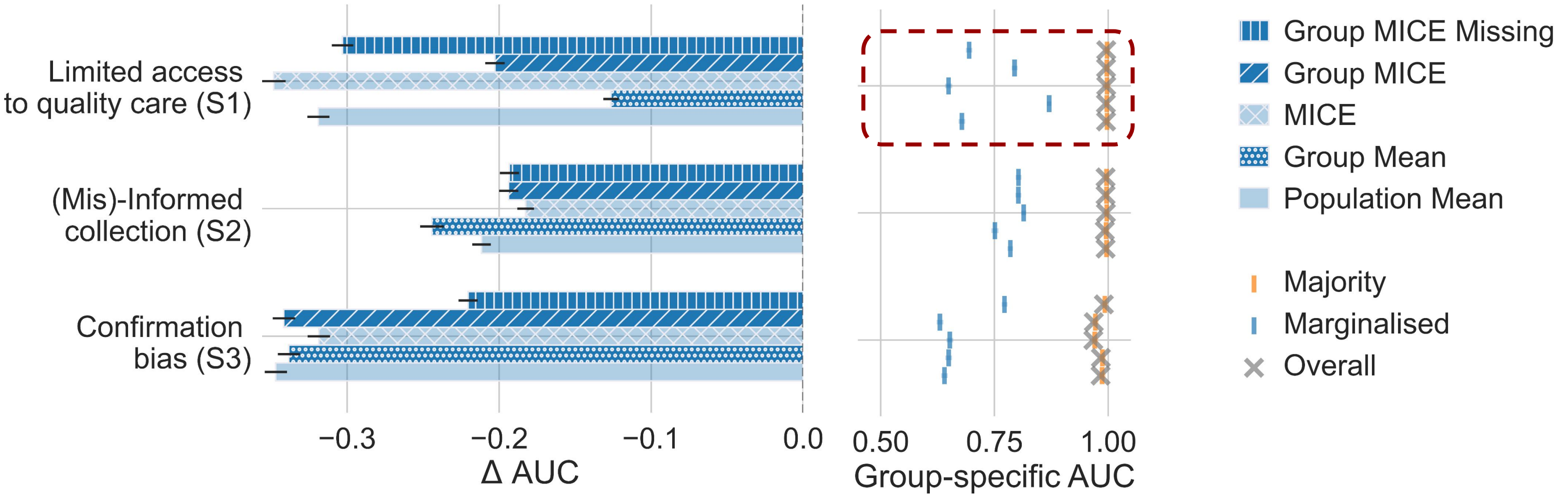


Downstream performance

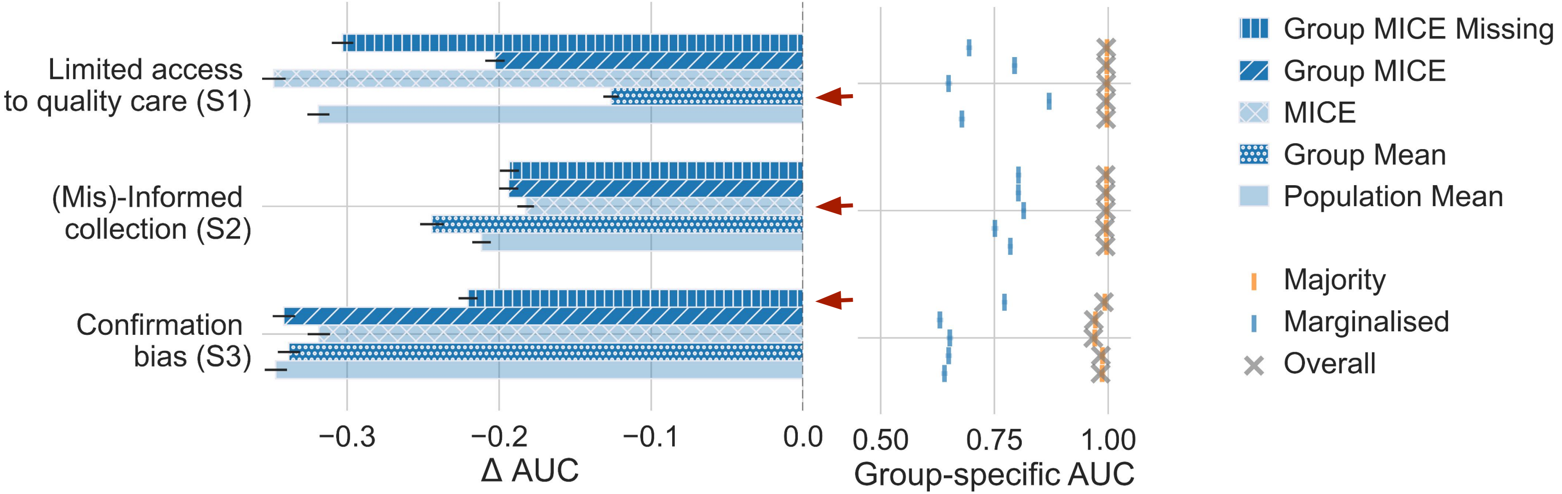


Downstream performance

1 Different imputation strategies may have equal downstream performance at the population level while having different group performance gaps.



Downstream performance

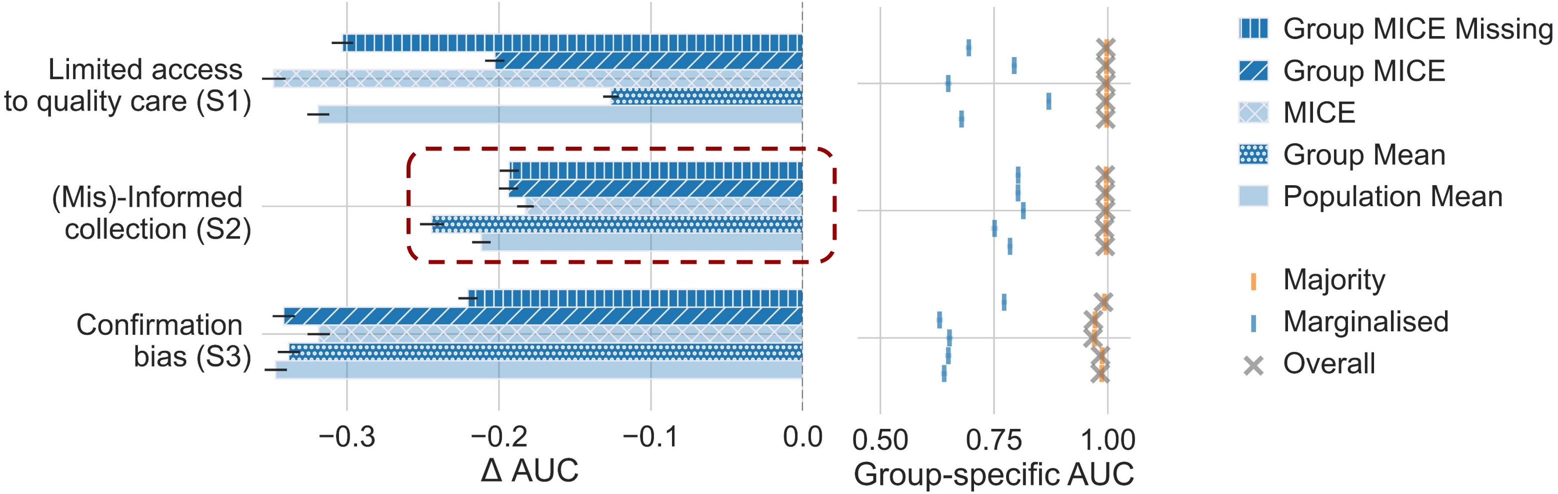


1 Different imputation strategies may have equal downstream performance at the population level while having different group performance gaps.

2 No imputation strategy consistently outperforms the others across clinical presence scenarios.

Downstream performance

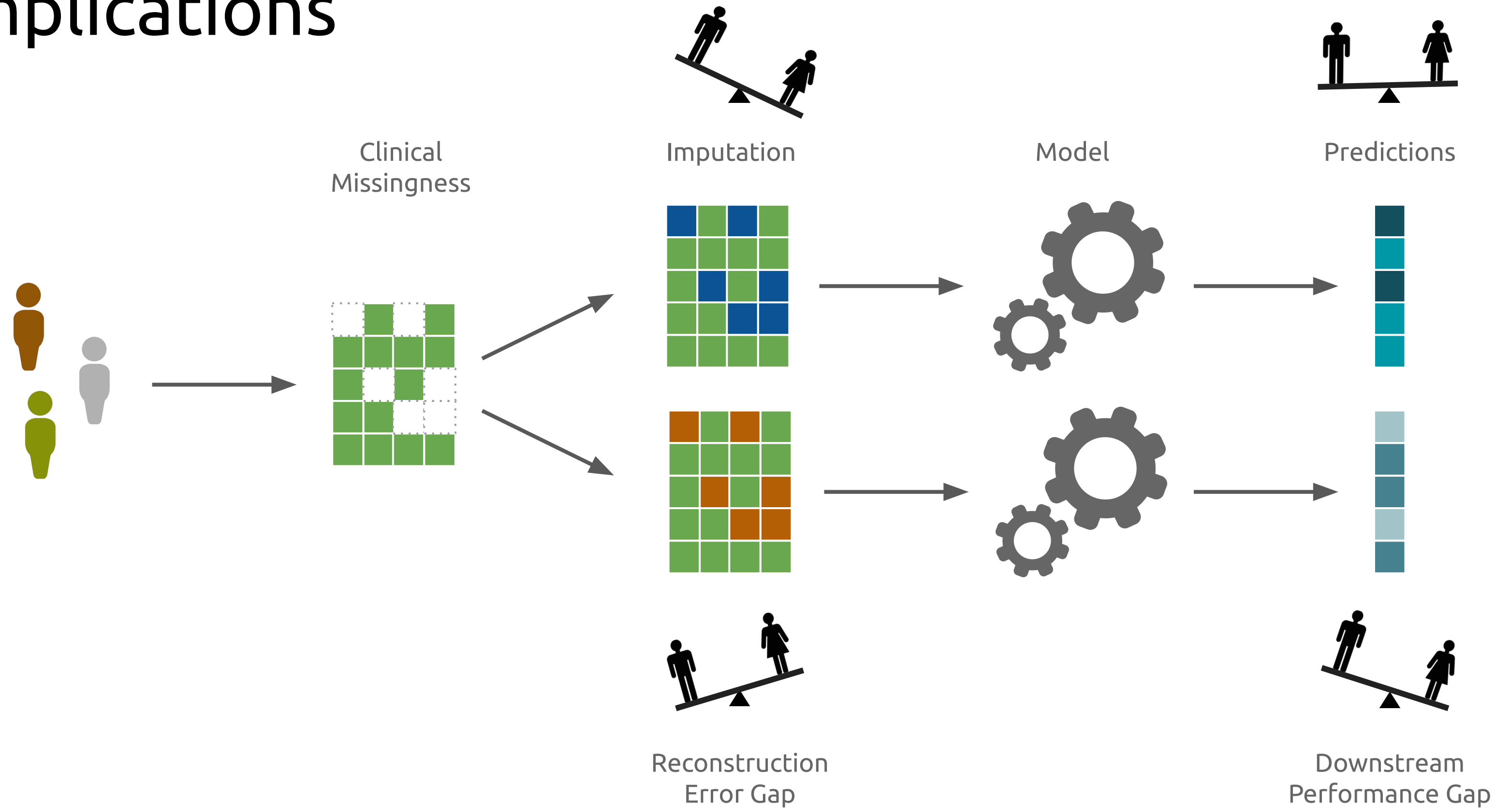
1 Different imputation strategies may have equal downstream performance at the population level while having different group performance gaps.



3 Current recommendations for group-specific imputation can increase the performance gap and yield a worse performance for the marginalised group

2 No imputation strategy consistently outperforms the others across clinical presence scenarios.

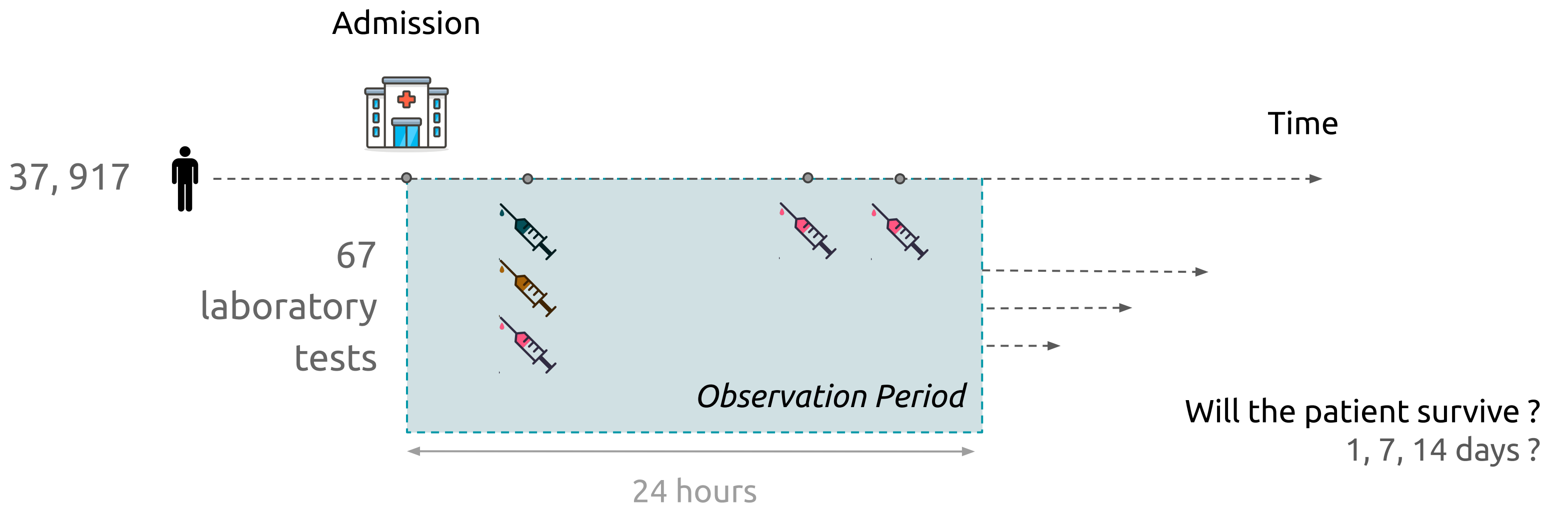
Implications



Implications

Hypotheses	Imputation quality	Predictive performance
Equally performing approaches at the population level have similar algorithmic fairness properties	✗	✗
Imputation properties are consistent across missingness processes	✗	✗
Controlling/stratifying on group results in improved group performance	✗	✗
Controlling/stratifying on group reduces group disparities	✗	✗

More than theoretical?

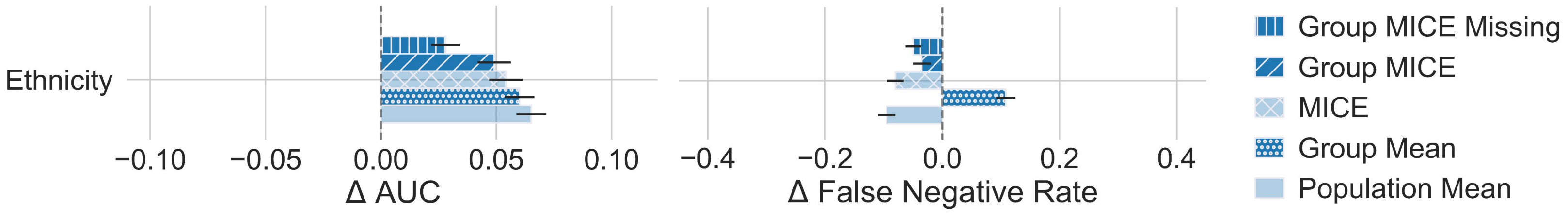


More than theoretical?

	Orders
Alive	**5.68**
Dead	7.57
Black	**5.24**
Other	5.86
Female	**5.54**
Male	6.03
Public	**5.67**
Private	6.11

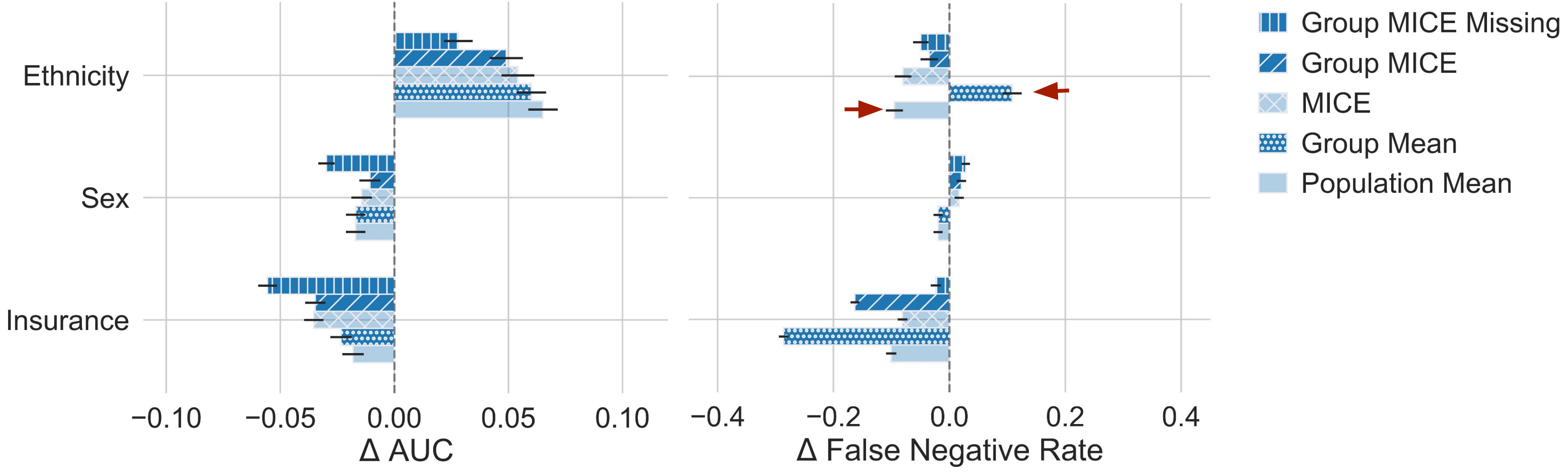
Real-world data presents group-specific observation processes.

More than theoretical?



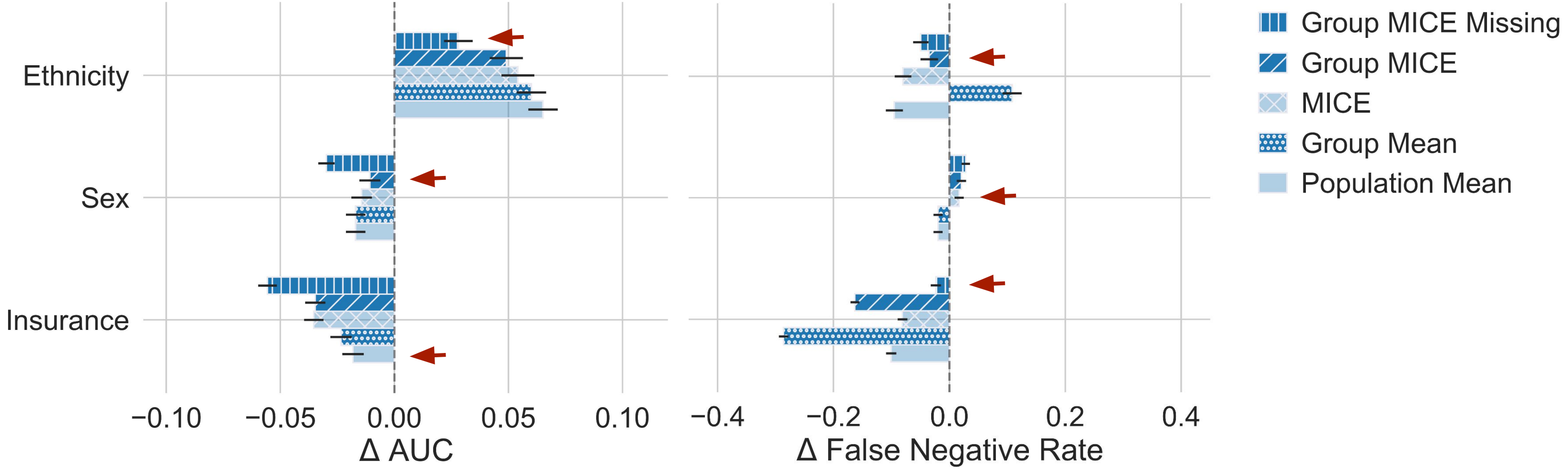
More than theoretical?

1 Different imputation strategies may have equal prediction performance at the population level while having opposite group performance gaps.



More than theoretical?

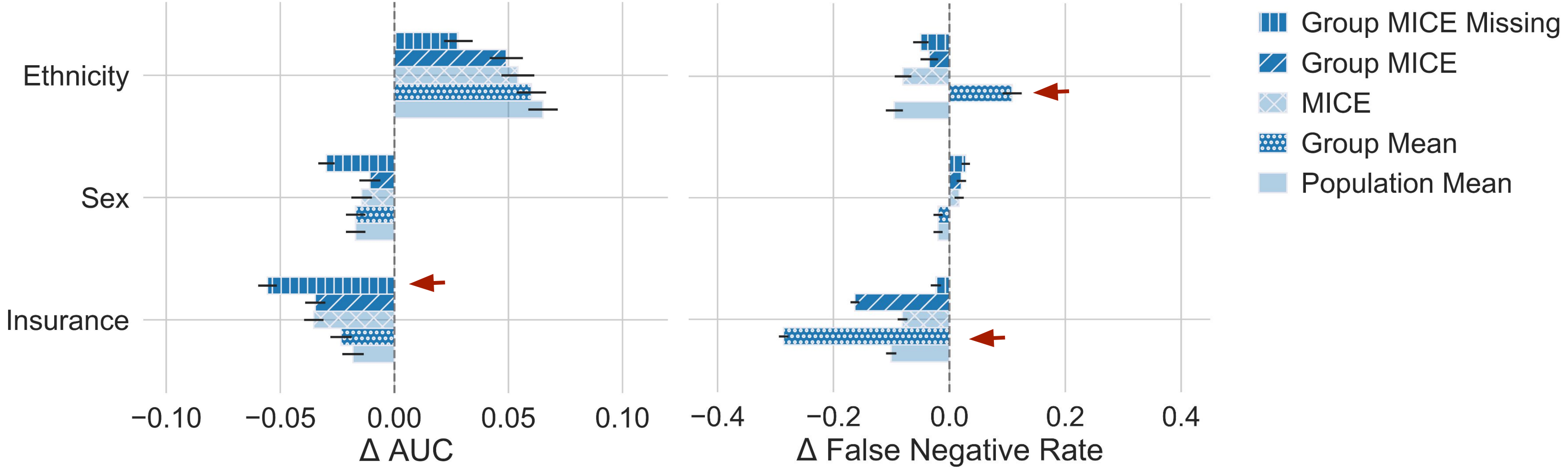
1 Different imputation strategies may have equal prediction performance at the population level while having opposite group performance gaps.



2 No imputation strategy consistently outperforms the others across clinical presence scenarios.

More than theoretical?

1 Different imputation strategies may have equal prediction performance at the population level while having opposite group performance gaps.



3 Current recommendations for group-specific imputation and use of missingness indicators can increase the performance gap and yield a worse performance for the marginalised groups.

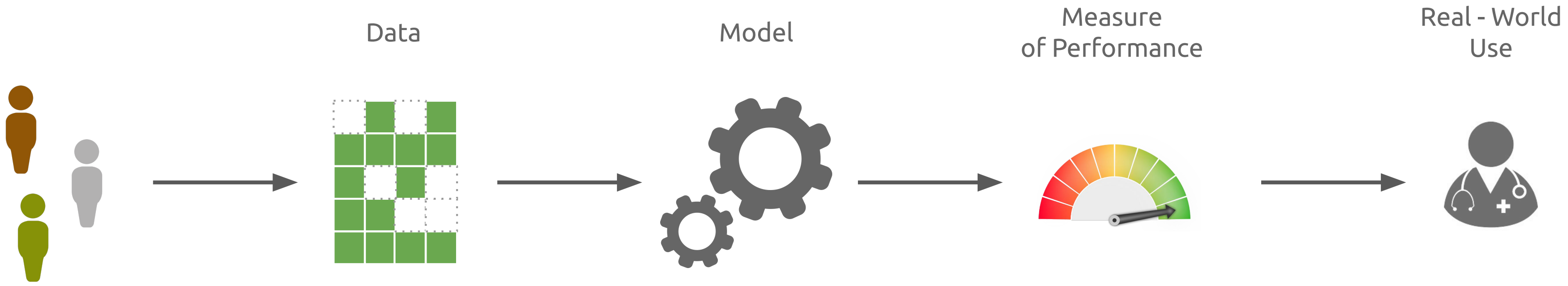
2 No imputation strategy consistently outperforms the others across clinical presence scenarios.

Recommendations

- **Study** the missingness process.
- State the missingness **assumptions**.
- Consider **differences** in the missingness process between **training** and **deployment**.
- Evaluate the impact of **different imputation strategies**.

For more details

Conclusion



Contact
vincent.jeanselme@gmail.com