These are the slides I used in my presentation at the University of Basel. I have annotated the overview for some context – hope it is useful.

Machine Learning and Health Inequalities

Anja Leist, University of Luxembourg

8 November 2021

University of Basel, Institute for Biomedical Ethics Lecture Series

UNIVERSITY OF LUXEMBOURG
Department of Social Sciences

Overview

1. Definition health inequalities
2. Mapping of ML approaches to research questions in the health and social sciences
3. ML to quantify health inequalities
4. *The Problem* in classic ML prediction: Health inequalities are ‘hidden’
5. Fairness in ML to acknowledge health inequalities
6. ML to reduce health inequalities

In this lecture, I wanted to give an overview of wthe different points of the research process where machine learning (ML) and health inequalities may meet.

After a definition of health inequalities and related concepts,

I would like to present a recent mapping of machine learning approaches to research questions in the social and health sciences. Here, machine learning can be used to (descriptely) assess health inequalities, (predictively) use sensitive attributes to predict health outcomes – this is where much of the ‘unfairness’ of ML algorithms is currently being hypothesized – and (through counterfactual prediction) can use sensitive attributes or their proxies as exposures to causally explain health outcomes of interest.

After showing how ML can quantify health inequalities,

I will explain The Problem in classic ML: If we don’t give enough attention to sensitive attributes like sex/gender, age, or race/ethnicity, ML may maintain health inequalities or even reinforce them. I show this with non-ML based race correction algorithms, with the note that ML algorithms may through less transparent and interpretable model decisions possibly further aggravate these issues.

I present a few recent suggestions of the extent to which we can possibly arrive at fairer ML models. Some researchers have provided evidence that it is almost impossible to arrive at true fairness in ML-based modeling.

We will end with a few examples of how ML can be used to reduce health inequalities, particularly when ML is used to facilitate diagnosis, treatment, care of currently underserved populations.

1 Definitions

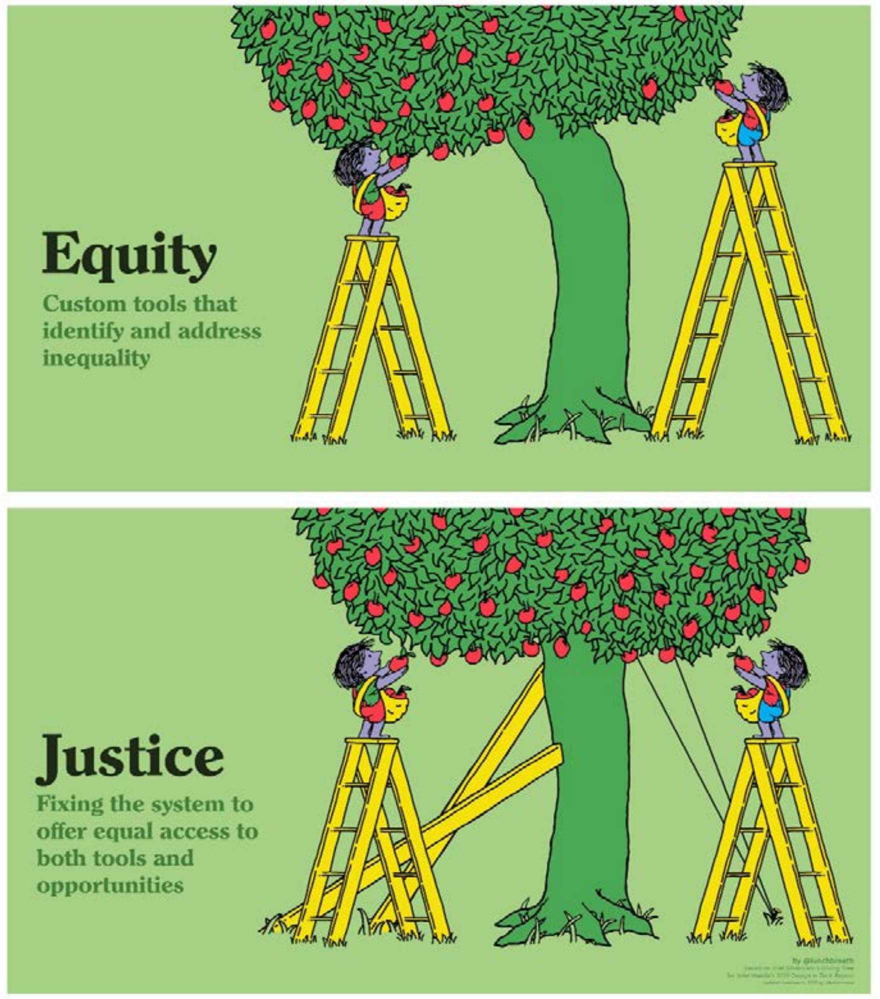
- **Health inequalities**: Differences in health between groups of people characterized by, for example, sex/gender, race/ethnicity, or socioeconomic status. Inequalities can be observed within and between countries.
- **Health inequities**: Differences in health status or in the distribution of health resources between different population groups, arising from the social conditions in which people are born, grow, live, work, and age, which are unfair and avoidable.

https://www.who.int/features/factfiles/health_inequities/en/

- **Social determinants of health**: Social determinants are the conditions in which people are born, grow, live, work and age.

https://www.who.int/social_determinants/sdh_definition/en/

- Related terms: Health and cultural capital, Health disparities (race/ethnicity)



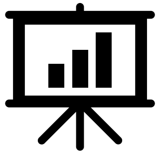
Tony Ruth’s equity series
https://cx.report/2020/06/02/equity/

Inequalities in Dementia

Inequalities in dementia can be found not only concerning risk of dementia – please see Mayeda's 2016 publication on differential dementia risk by race/ethnicity – but also regarding enrolment in clinical trials (research), frequency, intensity, start of dementia care, and dementia treatment and treatment of behavioral symptoms of dementia.

Risk of Dementia

Dementia Care



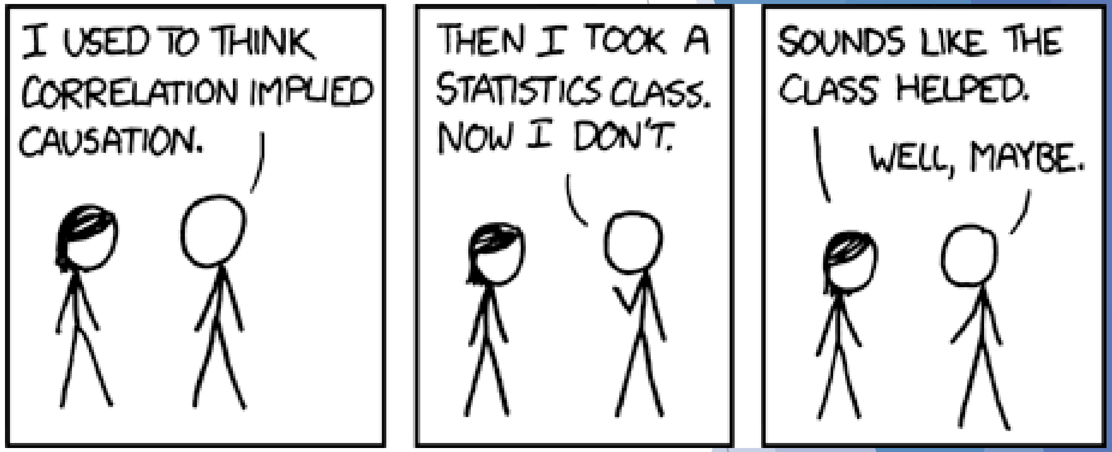
Dementia Research

Dementia Treatment

Leist, A. K. (2017). Social inequalities in dementia care, cure, and research. *Journal of the American Geriatrics Society.* doi: 10.1111/jgs.14893

2 Mapping ML approaches to research questions in the social and health sciences

- A classification of data science tasks
  - Data science for description
  - Data science for prediction
  - Data science for counterfactual prediction (causal inference)
- In some cases some conceptual or methodological overlap



Hernán, M. A., Hsu, J., & Healy, B. (2019). A second chance to get causal inference right: a classification of data science tasks. *Chance*, *32*(1), 42-49.

Classification of research questions in the social and health sciences

Description	Prediction	Causal inference

- Example research questions
- Motivation: provide optimal matches between research questions and ML methods

Leist, A. K., Klee, M., Kim, J. H., Rehkopf, D. H., Bordas, S. P. A., Muniz-Terrera, G., & Wade, S. (2021). Machine learning in the social and health sciences. https://arxiv.org/abs/2106.10716

ML for description

- Unsupervised learning, e.g. clustering
- Dimensionality reduction, e.g. factor analysis
- Descriptive approaches can be used to feature-engineer exposures for further analysis
- Predictive ML can be used to answer descriptive questions*
- Examples:

Kuan, V., Fraser, H. C., Hingorani, M., Denaxas, S., Gonzalez-Izquierdo, A., Direk, K., … & Hingorani, A. D. (2021). Data-driven identification of ageing-related diseases from electronic health records. *Scientific reports*, *11*(1), 1-17.

*Luo, W., Nguyen, T., Nichols, M., Tran, T., Rana, S., Gupta, S., … & Allender, S. (2015). Is demography destiny? Application of machine learning techniques to accurately predict population health outcomes from a minimal demographic dataset. *PloS one*, *10*(5), e0125602.

ML for prediction

- “Using data to map some features of the world (the inputs) to other features of the world (the outputs)” (Hernán et al., 2019)
- Classification or regression
- Large set of predictors = explain maximum variance in the outcome vs.
- Optimal set of predictors = balance between data requirements and variance explanation vs.
- Evaluation of candidate predictors
- Example:

Baćak, V., & Kennedy, E. H. (2019). Principled machine learning using the super learner: An application to predicting prison violence. *Sociological Methods & Research*, *48*(3), 698-721.

ML for causal inference

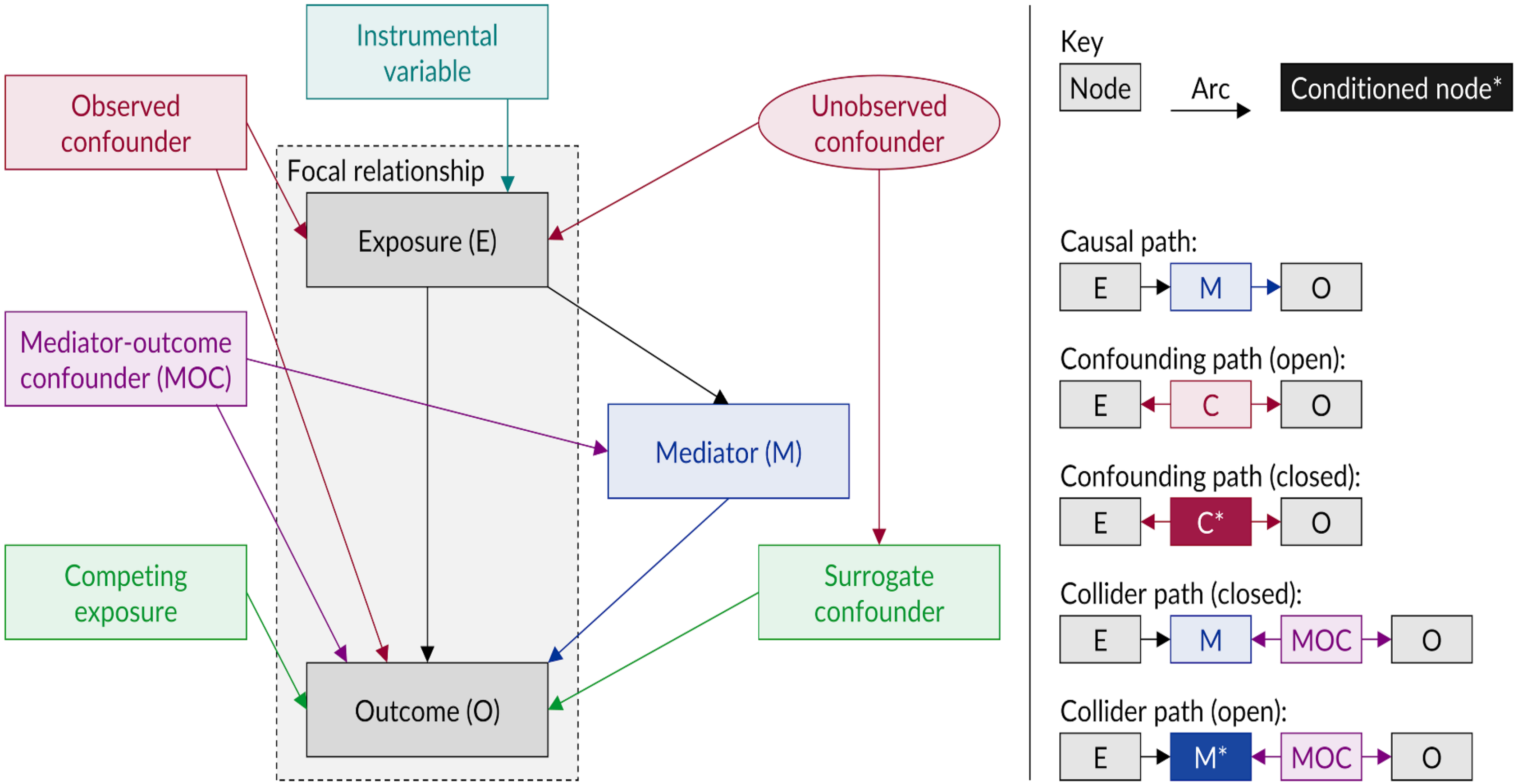
- Counterfactual prediction
- Causal inference often implicitly the main research goal in social and health sciences
- Draw a DAG!

Glymour, M. M. (2006). Using causal diagrams to understand common problems in social epidemiology. *Methods in social epidemiology*, 393-428.

Draw your assumptions before your conclusions (Miguel Hernán, edX course)

- (Causal structural learning)

Figure 1 Illustration of the main components of a DAG, the most common types of contextual variables and the most common types of paths



Tennant, P. W. G., Murray, E. J., Arnold, K. F., Berrie, L., Fox, M. P., Gadd, S. C., … & Ellison, G. T. H. (2020). Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. *International Journal of Epidemiology*.

Methods to detect causal relationships

- **“Traditional” methods** e.g. regression and more advanced methods (Table 1 taken from Hernán et al., 2019)
- **Predictive ML approaches** with properly set up data (Blakely et al. 2020)
- **ML approaches appropriate for causal inference**, e.g. BART

Blakely, T., Lynch, J., Simons, K., Bentley, R., & Rose, S. (2020). Reflection on modern methods: when worlds collide—prediction, machine learning and causal inference. *International journal of epidemiology*, *49*(6), 2058-2064.

Table 1—Examples of Tasks Conducted by Data Scientists Working with Electronic Health Records

	Data Science Task		
	Description	**Prediction**	**Causal inference**
Example of scientific question	How can women aged 60–80 years with stroke history be partitioned in classes defined by their characteristics?	What is the probability of having a stroke next year for women with certain characteristics?	Will starting a statin reduce, on average, the risk of stroke in women with certain characteristics?
Data	• Eligibility criteria • Features (symptoms, clinical parameters …)	• Eligibility criteria • Output (diagnosis of stroke over the next year) • Inputs (age, blood pressure, history of stroke, diabetes at baseline)	• Eligibility criteria • Outcome (diagnosis of stroke over the next year) • Treatment (initiation of statins at baseline) • Confounders • Effect modifiers (optional)
Examples of analytics	Cluster analysis …	Regression Decision trees Random forests Support vector machines Neural networks …	Regression Matching Inverse probability weighting G-formula G-estimation Instrumental variable estimation …

Example: Does high fruit and vegetable density in mothers-to-be prevent adverse birth outcomes?

- Food Frequency Questionnaire: Fruit and vegetable density before conception
- Outcomes see figure
- **TMLE and SuperLearner**: predictive ability of risk factor, adjusted for confounders

Bodnar, L. M., … & Naimi, A. I. (2020). Machine learning as a strategy to account for dietary synergy: an illustration based on dietary intake and adverse pregnancy outcomes. *The American journal of clinical nutrition*, *111*(6), 1235-1243.

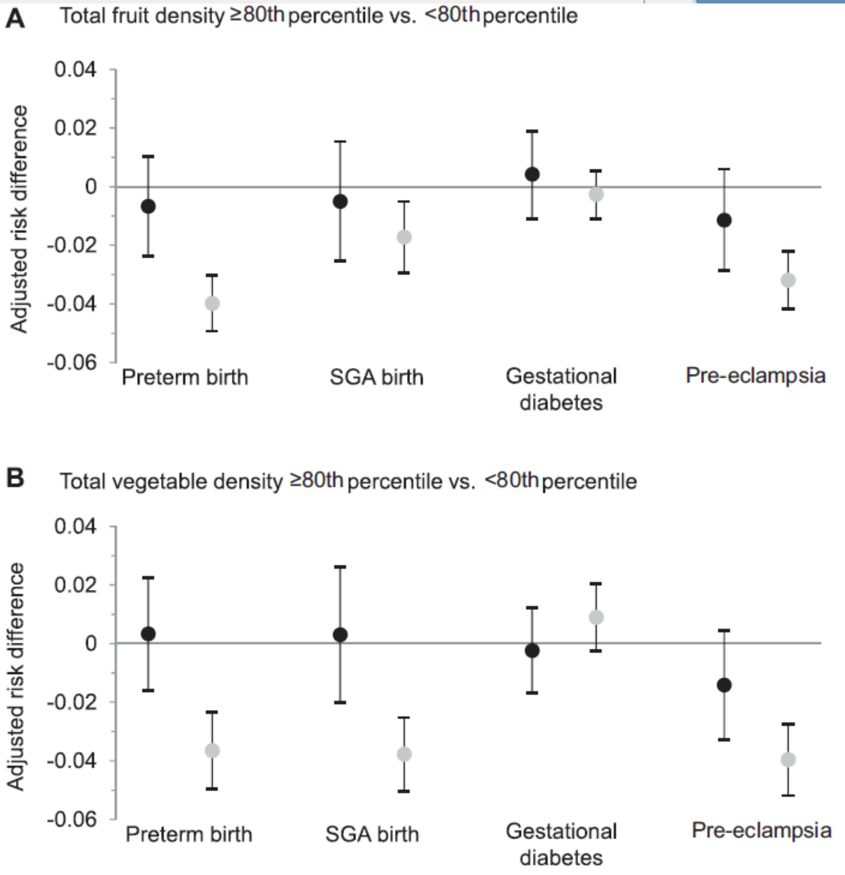


FIGURE 1 Associations between fruit and vegetable density and risk of adverse pregnancy outcomes in the Nulliparous Pregnancy Outcomes Study: monitoring mothers-to-be ($n = 7252$). (A) Adjusted differences in risk (95% CIs) of preterm birth, SGA birth, gestational diabetes, and pre-eclampsia between total fruit intake relative to energy ≥80th percentile (≥1.2 cups/1000 kcal) ($n = 1514$) compared with <80th percentile ($n = 6058$) and (B) total vegetable intake relative to energy ≥80th percentile (≥1.3 cups/1000 kcal) ($n = 1514$) compared with <80th percentile ($n = 6058$). Point estimates in black were generated from multivariable logistic regression. Point estimates in gray were generated from Super Learner with targeted maximum likelihood estimation. All models were adjusted for maternal age, race/ethnicity, education, marital status, smoking status, prepregnancy BMI, insurance, and usual dietary intake of whole grains, dairy products, total protein foods, seafood and plant proteins, fatty acids, refined grains, sodium, and "empty" calories. SGA, small-for-gestational-age.

3. ML to quantify health inequalities

- Social determinants of health
- Provide accurate and precise estimates of social health gradients or -gaps

- Example 1: Population level health outcomes for healthcare planning

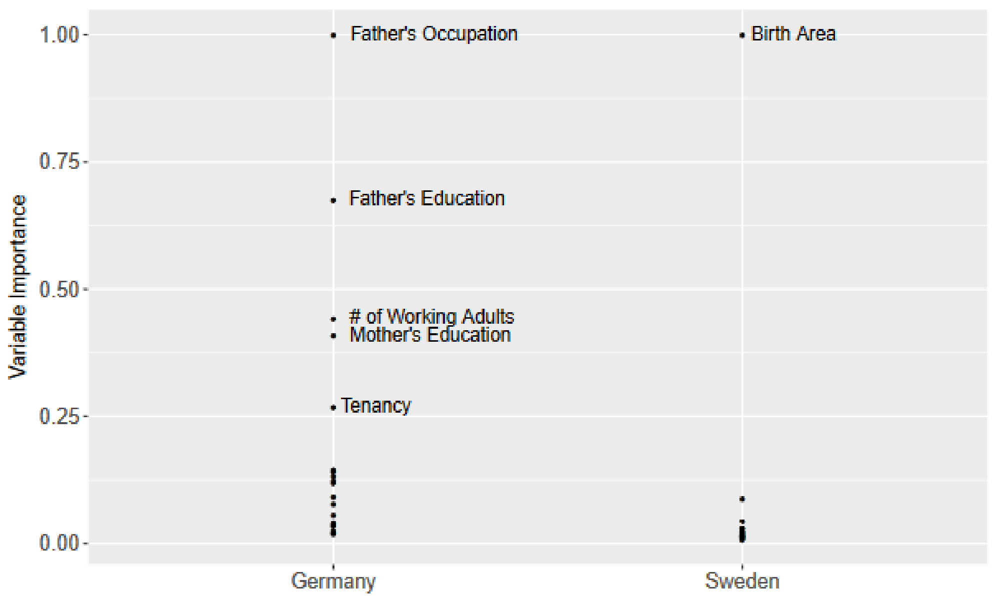
Luo, W., Nguyen, T., Nichols, M., Tran, T., Rana, S., Gupta, S., ... & Allender, S. (2015). Is demography destiny? Application of machine learning techniques to accurately predict population health outcomes from a minimal demographic dataset. *PloS one*, *10*(5), e0125602.

Example 2: Inequality of opportunity

- Determinants of income in comparative perspective
- **Regression trees** to determine Gini
- Regression trees to determine most relevant factors for IoO per country
- Comparison against traditional methods

Brunori, P., Hufe, P., & Mahler, D. G. (2018). The roots of inequality: Estimating inequality of opportunity from regression trees. *World Bank Policy Research Working Paper*, (8349). https://www.ifo.de/DocDL/wp-2018-252-brunori-hufe-mahler-estimating-inequality.pdf

Figure 5: Variable Importance for Germany and Sweden



Note: Each dot shows the importance of a particular circumstance for the predictions from our random forest. The importance of a circumstance is measured by permuting the circumstance, calculating a new MSE^{OOB}, and computing the difference in the MSE^{OOB} between the original model and the model with the permuted circumstance. The importance measure is standardized such that the circumstance with the greatest importance in each country equals one. Occupation refers to ISCO-08 one digit codes. All variables describing household characteristics refer to the period in which the respondent was about 14 years old. See Table 1 for details.

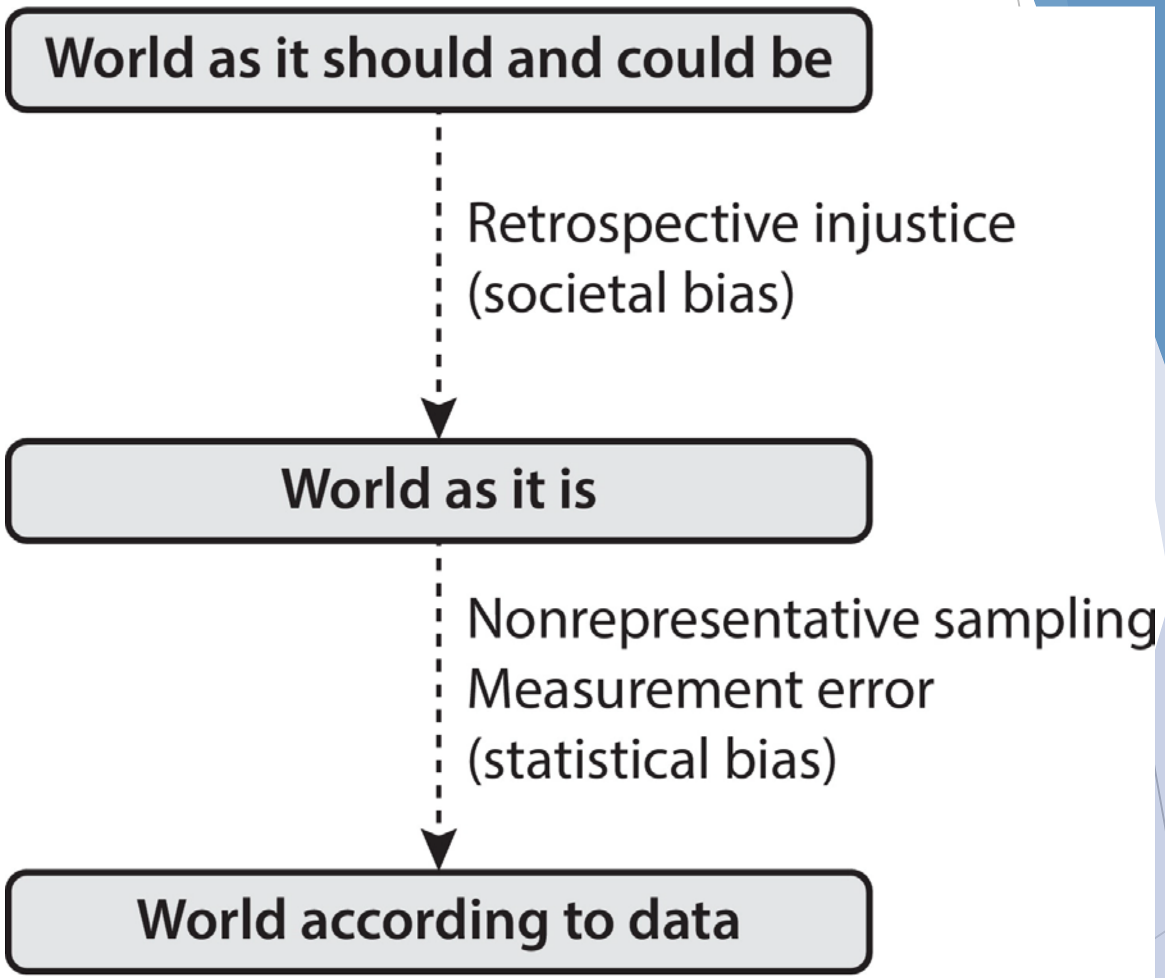
4. *The Problem* in classic ML prediction: Health inequalities are 'hidden'

- Race correction in clinical algorithms (ML- or non-ML based)
  - Black patients receive less favorable risk scores in various clinical applications that may deter them from further diagnostic assessment or have them receive less favorable treatment (Vyas et al. 2020)
  - Questions the meaning of 'race'
  - Data inequality
  - Systemic discrimination
  - Reductionist, unaware of possibly complex, intersectional background

- Race correction will oftentimes result in steering healthcare away from minority groups: favoritism of already advantaged groups, possibly exacerbating existing health(care) inequalities (Vyas et al. 2020)

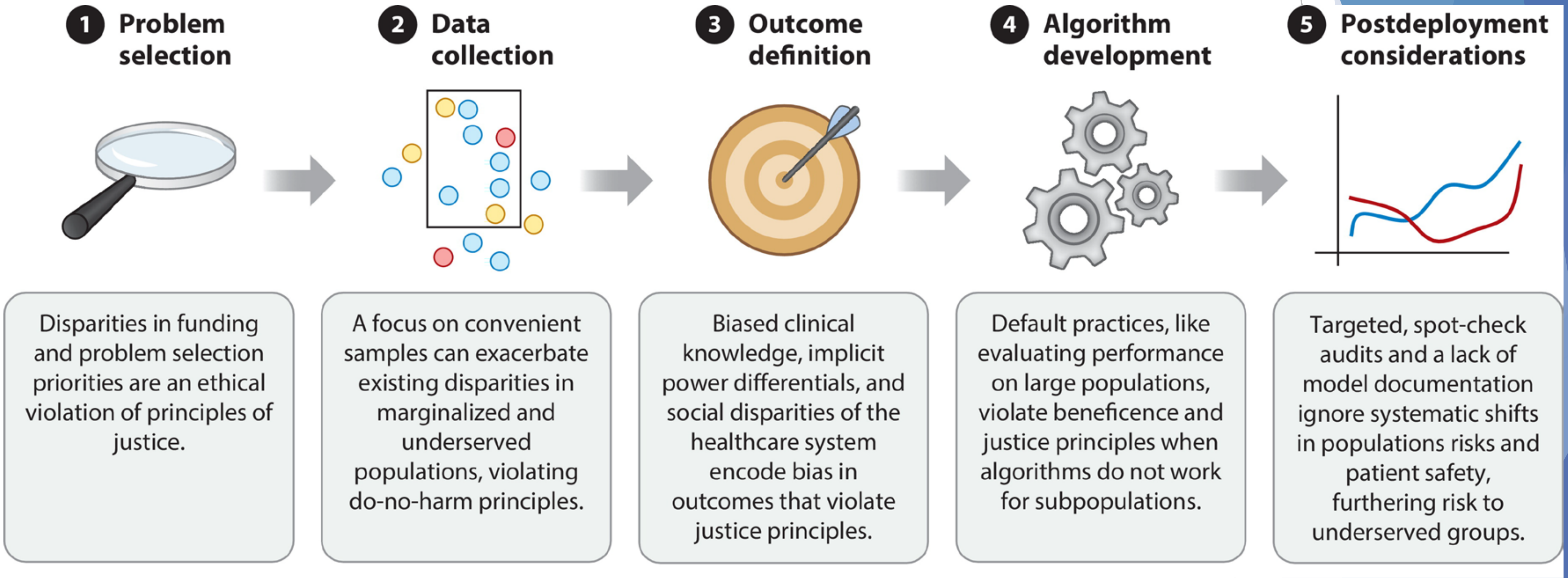
Vyas, D. A., Eisenstein, L. G., & Jones, D. S. (2020). Hidden in plain sight—reconsidering the use of race correction in clinical algorithms. New England Journal of Medicine, 383, 874-882.

5 Fairness in ML



Mitchell, S., Potash, E., Barocas, S., D'Amour, A., & Lum, K. (2021). Algorithmic fairness: Choices, assumptions, and definitions. *Annual Review of Statistics and Its Application*, *8*, 141-163.

Ethical pipeline



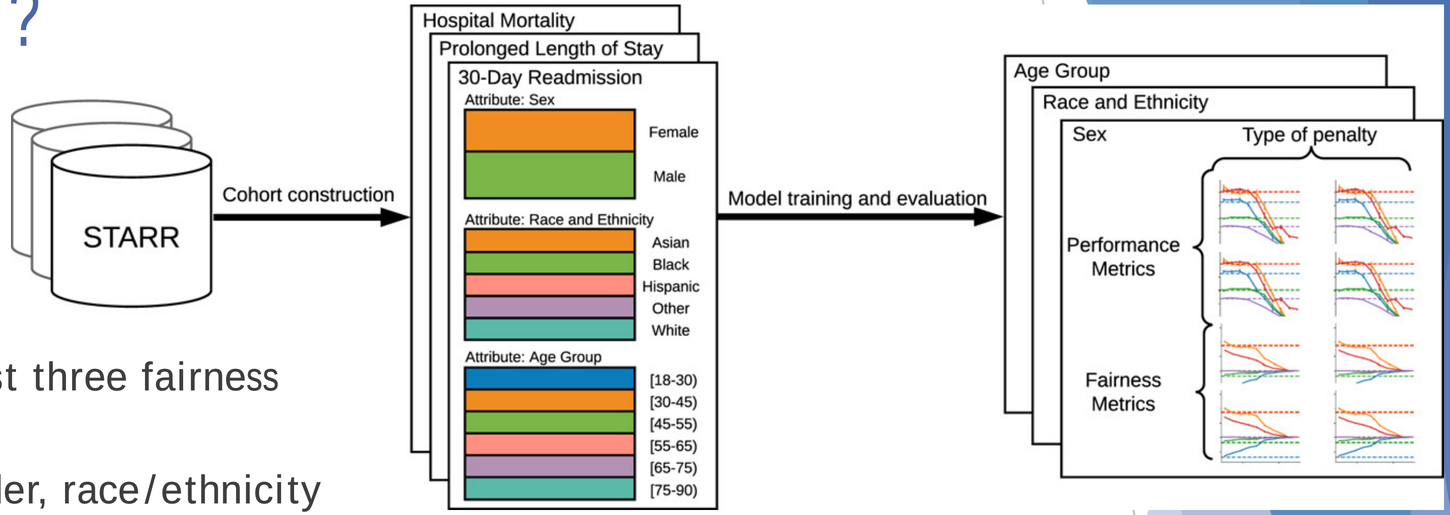
- Chen, I. Y., Pierson, E., Rose, S., Joshi, S., Ferryman, K., & Ghassemi, M. (2020). Ethical Machine Learning in Healthcare. *Annual Review of Biomedical Data Science*, *4*.

What to consider to make algorithms ‘fair’ (1)

- Possibility /willingness to include sensitive attribute in ML model or not
- Machine learning in healthcare: Dramatic cost and insurance implications (see Sherri Rose’s talk at JSM)
- Most applications so far: Mixture learning
- Example: Prediction of probable dementia for disparities research

Gianattasio, K. Z., Ciarleglio, A., & Power, M. C. (2020). Development of algorithmic dementia ascertainment for racial/ethnic disparities research in the US Health and Retirement Study. *Epidemiology (Cambridge, Mass.)*, *31*(1), 126.

Example: What is the best procedure to arrive at a fair model?

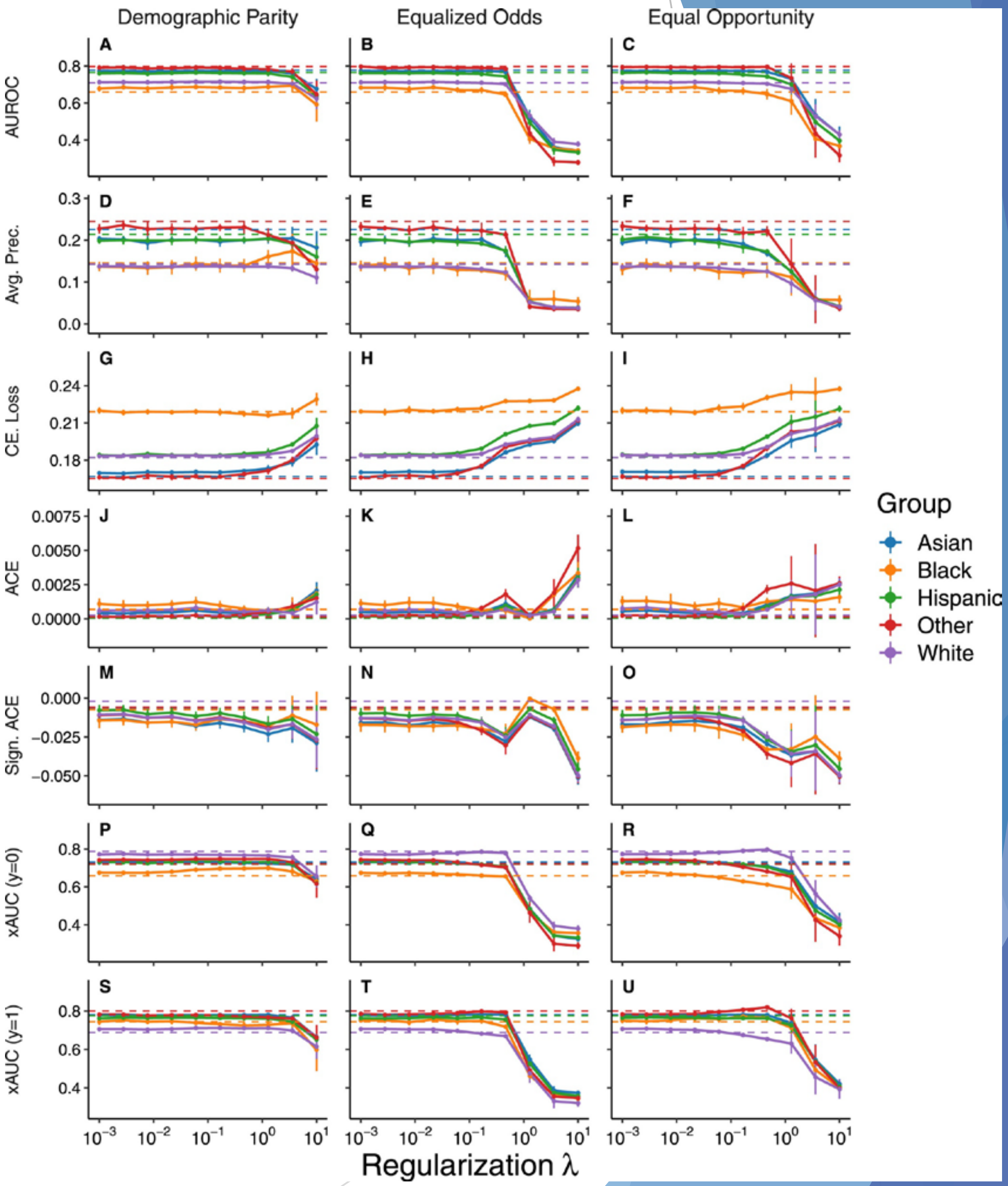


- Comparison of model performance against three fairness criteria
- Three sensitive attributes: age, sex/gender, race/ethnicity
- Increasing levels of regularization to penalize unfair decisions
- Heterogeneity of results = no one-size-fits-all approach → no standard recipe to improve fairness of ML algorithms
- Stakeholder involvement in evaluating algorithms

Pfohl, S. R., Foryciarz, A., & Shah, N. H. (2021). An empirical characterization of fair machine learning for clinical risk prediction. Journal of biomedical informatics, 113, 103621.

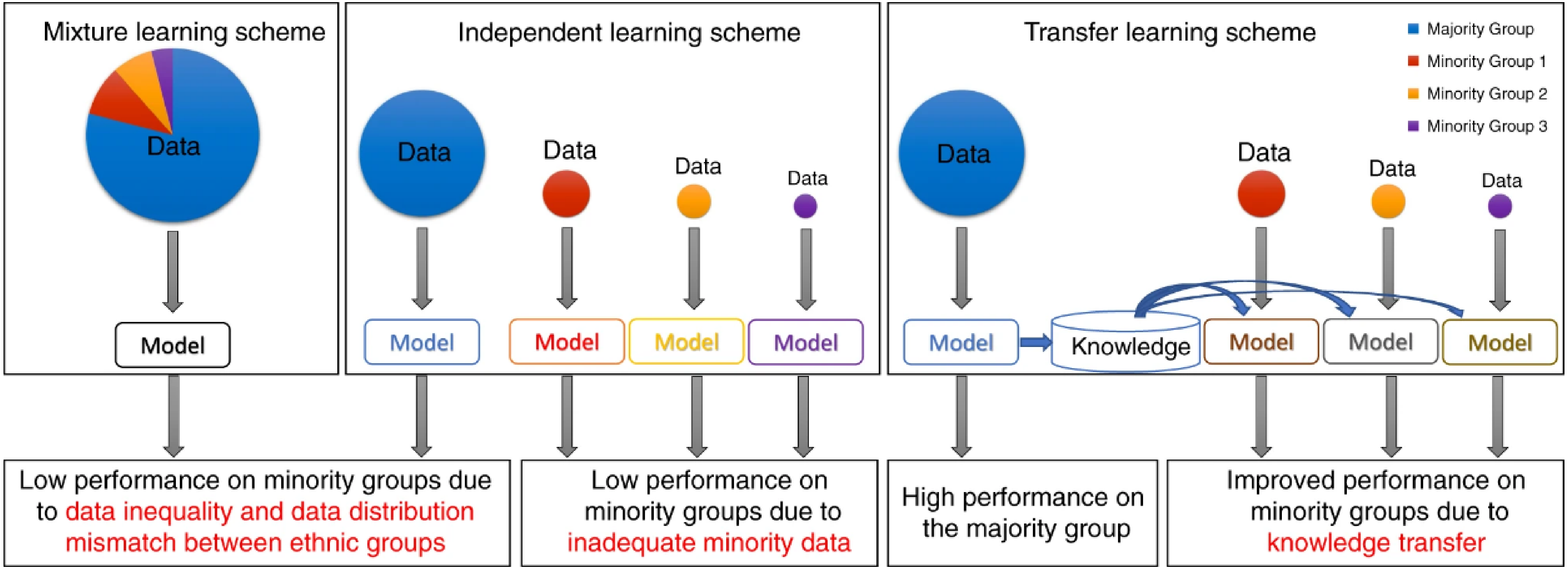
Example: What is the best procedure to arrive at a fair model?

- Pfohl et al., 2021: Group-level model performance measures as a function of the extent that violation of the fairness criterion is penalized when race/ethnicity is considered as the sensitive attribute for prediction of 30-day readmission in the STARR database.
- Results
  - Mean SD for the area under the ROC curve (AUROC)
  - Average precision (Avg. Prec)
  - Cross entropy loss (CE Loss)
  - Absolute calibration error (ACE)
  - Signed absolute calibration error (Sign. ACE)
  - Cross group ranking performance (xAUC)
- For each group for objectives that penalize violation of threshold-free
  - Demographic Parity
  - Equalized Odds
  - Equal Opportunity with MMD-based penalties
- Dashed lines = mean result for the unpenalized training procedure.



Pfohl, S. R., Foryciarz, A., & Shah, N. H. (2021). An empirical characterization of fair machine learning for clinical risk prediction. Journal of biomedical informatics, 113, 103621.

Fairness in ML: Transfer learning



- Gao, Y., & Cui, Y. (2020). Deep transfer learning for reducing health care disparities arising from biomedical data inequality. *Nature communications*, *11*(1), 1-8.
- Y. Cui: New cancer research project on transfer learning to address data inequality in omics data

What to consider to make algorithms ‘fair’ (2)

- Mixture learning or independent learning? Transfer learning?

Interesting development in mixture learning:

- *Counterfactual fairness*; *multi-level fairness* (Mhasawade & Chunara, 2021)

Mhasawade, V., & Chunara, R. (2021, July). Causal multi-level fairness. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 784-794).

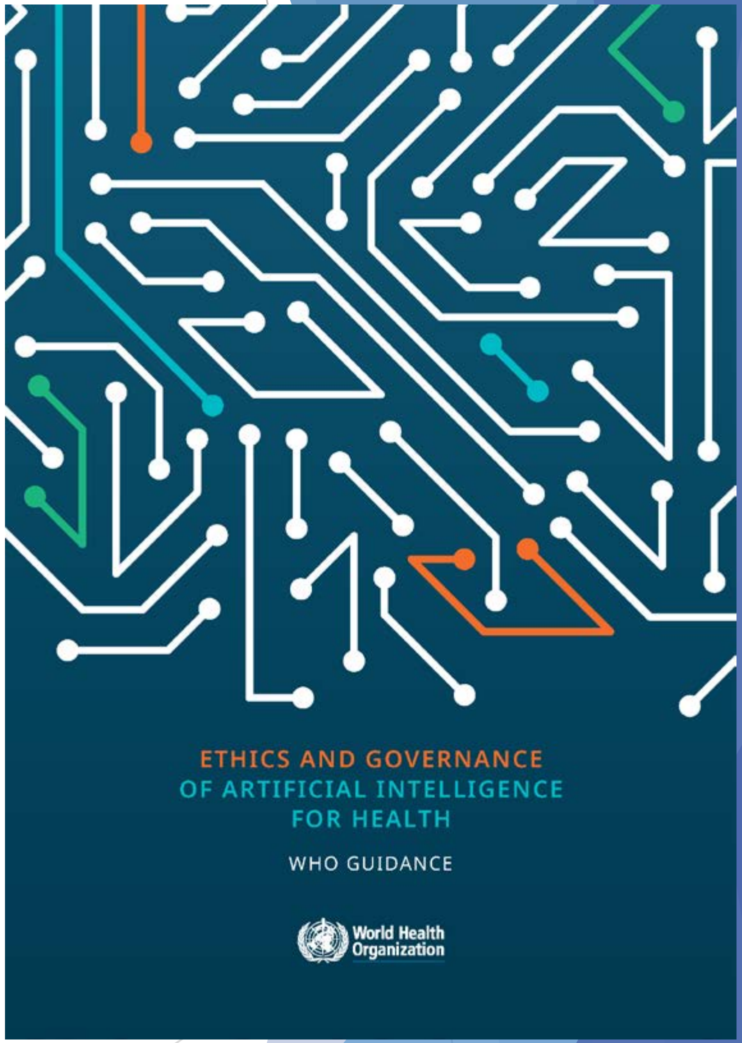
ML to reduce health inequalities

- AI to analyse imaging data and easy-to-collect questionnaire data - easier access to diagnosis, e.g. probable dementia, and care
- Identify new risk factors and putative causes, identify systematic biases in healthcare, population health monitoring etc.

Mhasawade, V., Zhao, Y., & Chunara, R. (2021). Machine learning and algorithmic fairness in public and population health. *Nature Machine Intelligence*, *3*(8), 659-666.

Leist, A. K., Klee, M., Kim, J. H., Rehkopf, D. H., Bordas, S. P. A., Muniz-Terrera, G., & Wade, S. (2021). Machine learning in the social and health sciences. https://arxiv.org/abs/2106.10716

World Health Organization. (2021). Ethics and governance of artificial intelligence for health: WHO guidance. Retrieved 28 October 2021 from https://apps.who.int/iris/bitstream/handle/10665/341996/9789240029200-eng.pdf



Thanks!

UNIVERSITY OF LUXEMBOURG
Department of Social Sciences

cognitiveageing.uni.lu

anja.leist@uni.lu

This research was funded by the European Research Council, grant agreement no. 803239

European Research Council
Established by the European Commission