

Case Studies

Scenario 1: AI-Assisted Screening for a Rare Condition Predominantly Affecting African Populations

Background: An AI tool is being developed to assist doctors in screening for a rare but serious condition, which mostly affects people of African descent and which can be fatal unless treated promptly. Screening currently relies upon a team of experts reviewing reports of symptoms as well as follow-up MR imaging data. Using this approach about 10% of positive cases are missed resulting in significant mortality. There is an expensive but reliable blood test that can be used for subjects flagged up by the screening process.

Development and Training: The AI tool will be trained using the symptoms and MR imaging as input data and blood test results as output labels. Training and evaluation will be based upon data acquired from a network of 5 hospitals across the USA. It is intended that the tool will be made available globally to improve screening for the rare condition.

Questions to consider:

- What dangers/risks of the use of AI for this problem can you identify at this stage?
- How would you go about addressing these?
- What fairness metric(s) do you think might be appropriate when assessing the AI tool for potential bias?

REMINDER - DEFINITIONS OF FAIRNESS

- *False Negative Rate (FNR): the rate at which positive cases are missed by the classifier*
- *Demographic parity - equal chance of being classified positive for each protected group*
- *Equalised odds - equal true positive rate (TPR) & false positive rate (FPR) for each protected group*
- *Equal opportunity - only equalise either FPR or FNR, not both*