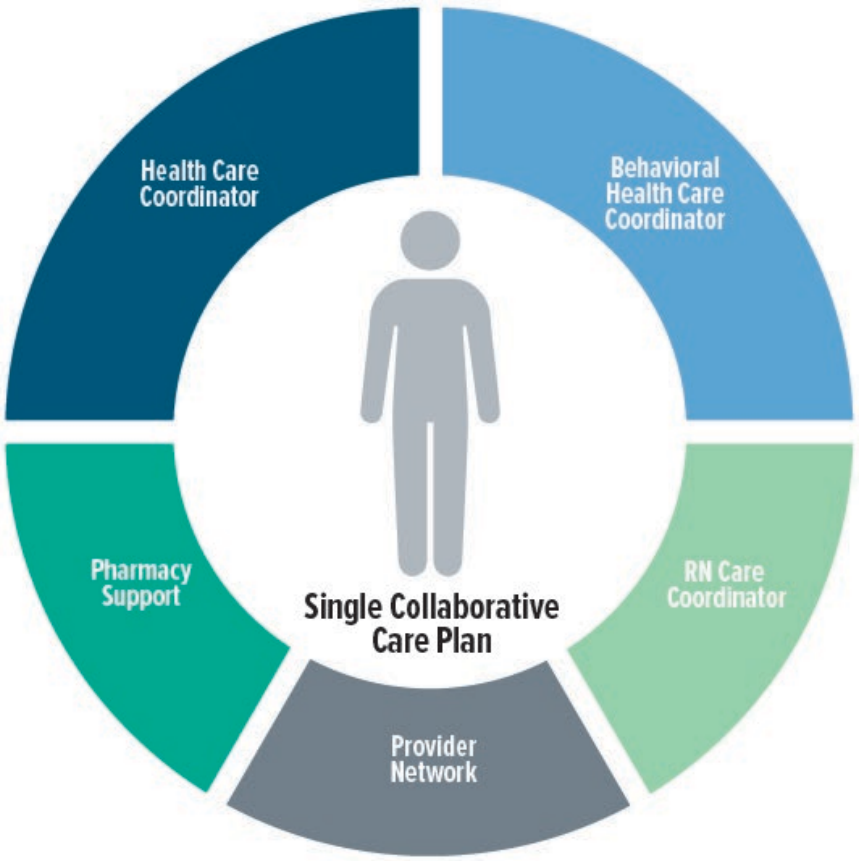
Two Stories of Race and Algorithms

Sendhil Mullainathan

Care Coordination Programs



Patients who will be sick receive dedicated resources

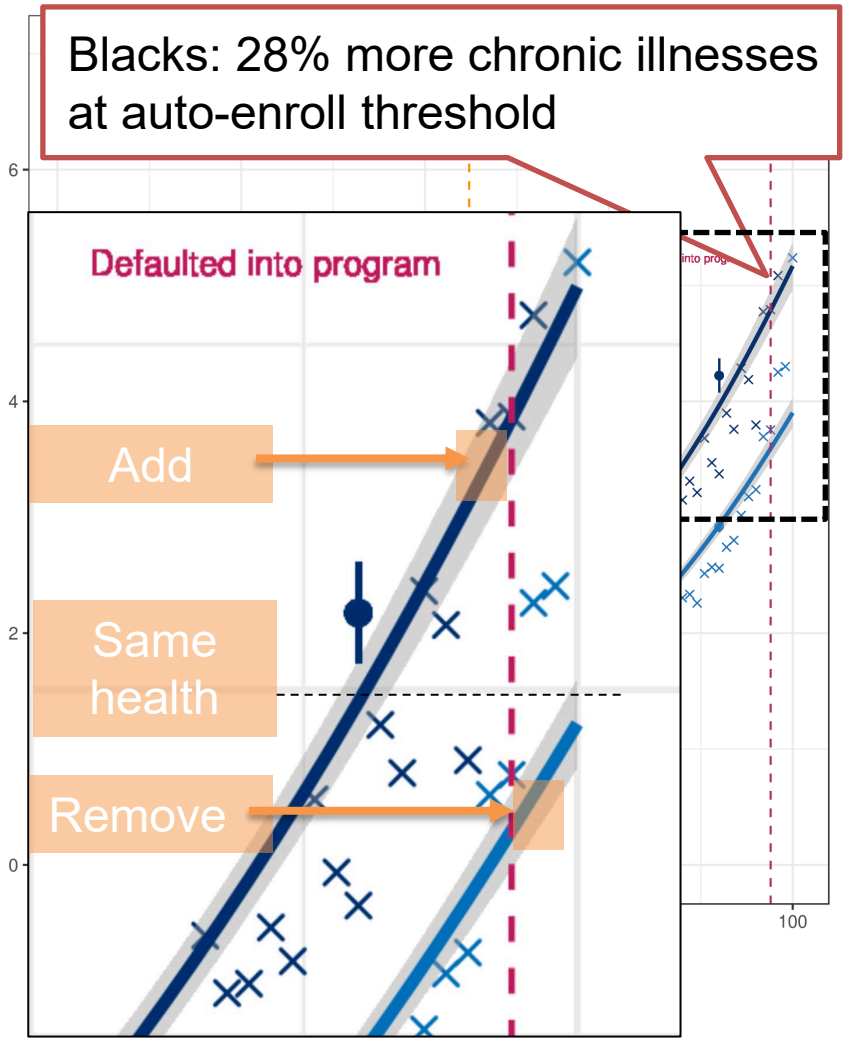
How do we identify who *will* be sick?

Algorithm – deployed > 100m people

How good are these algorithms?

Racial Disparities

Chronic medical conditions vs risk score

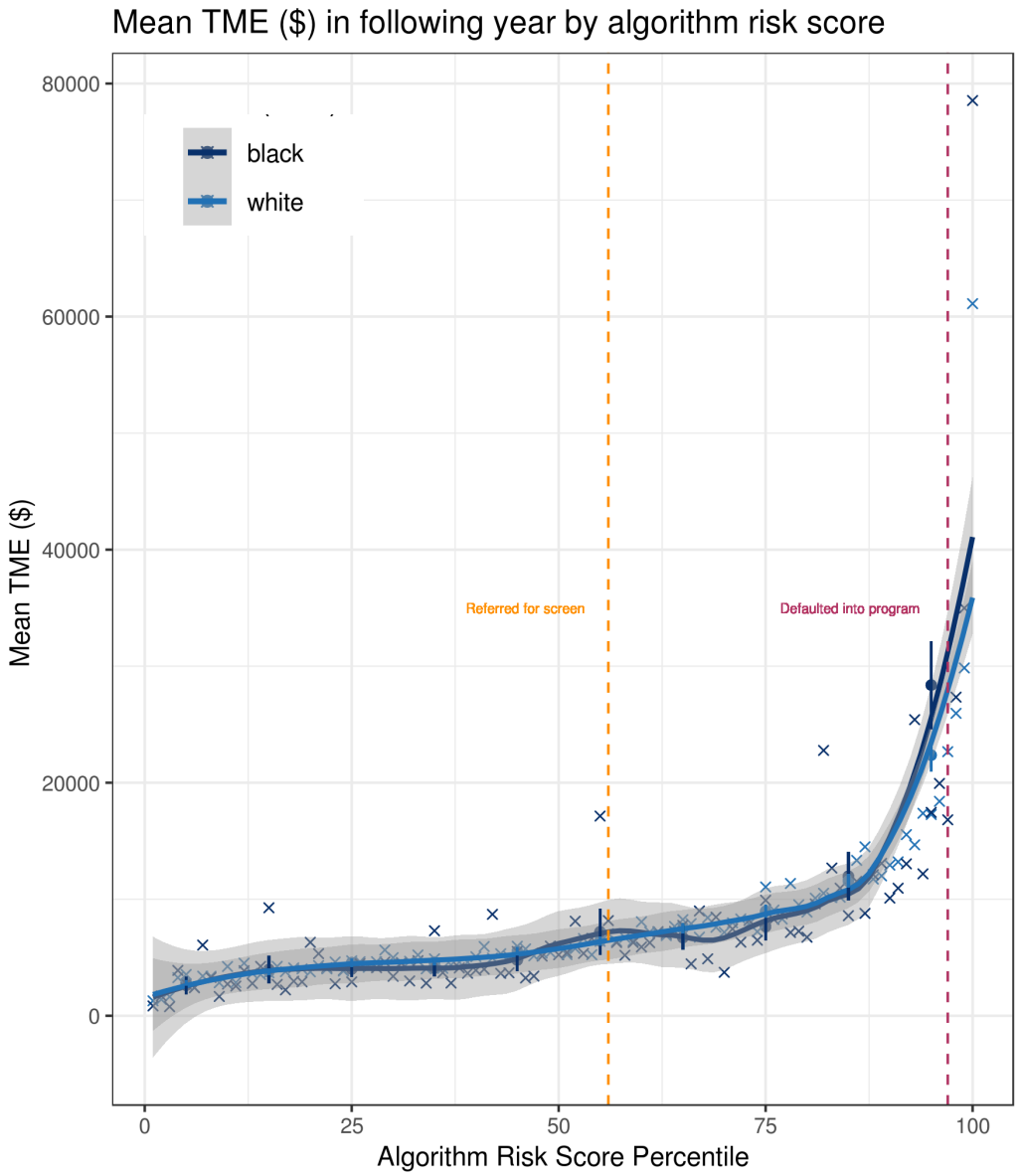


Simulate: replace healthier whites with less healthy blacks until marginal patients equally healthy

Double fraction of auto-enrolled blacks

17.7% → 36.1%

Where is algorithm going right?



Algorithm well calibrated by race for total health costs

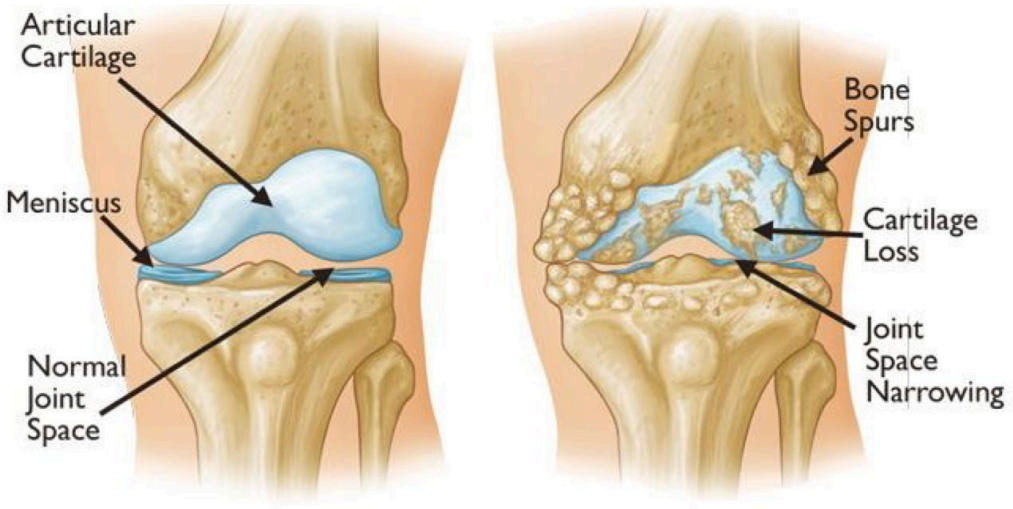
Whites have better access to health care

Blacks utilize less health care, cost less

So accurate cost prediction = biased health prediction

Key point: Label bias

Control for severity

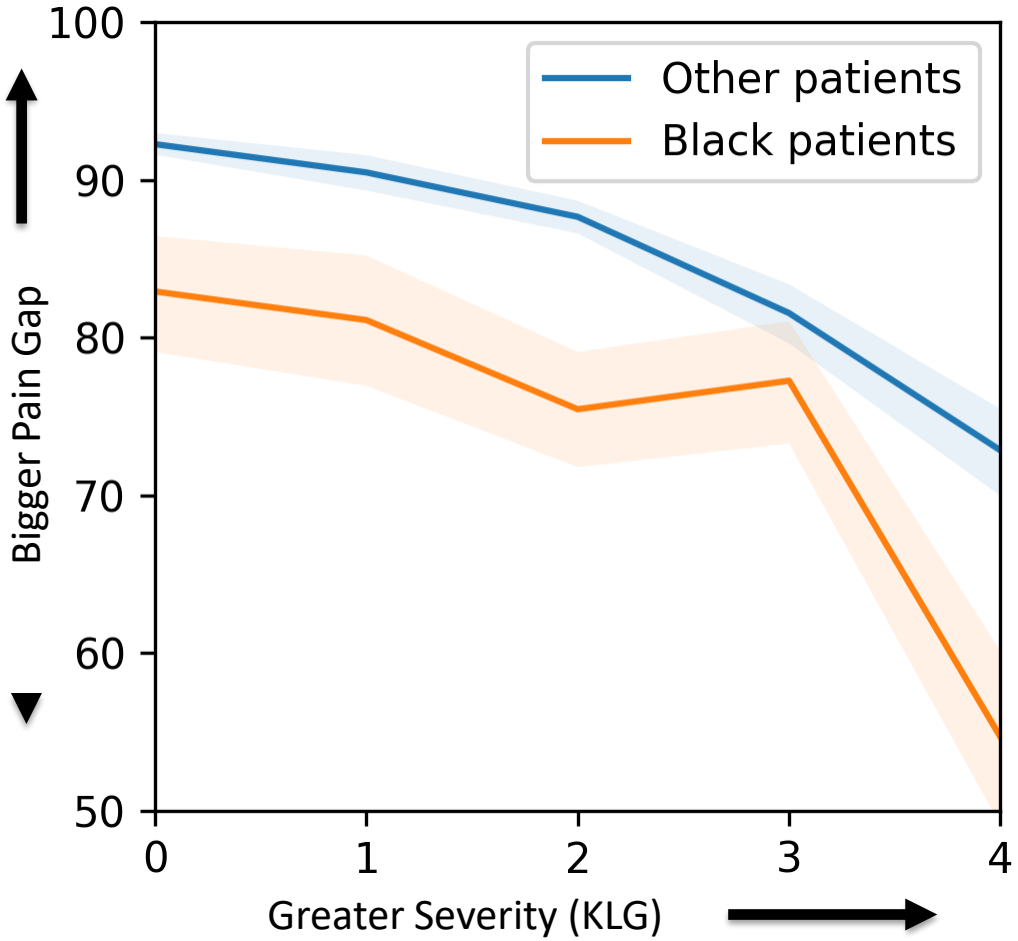


Osteoarthritis most common joint disorder in US

Pain especially severe amongst disadvantaged patients

“Inside their knees”

“Inside their heads”



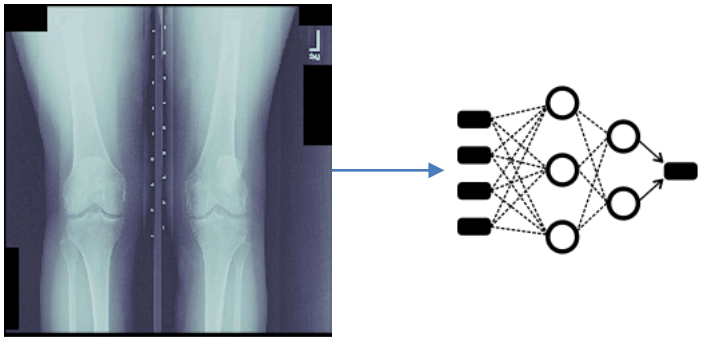
Notice this controls for severity

Suggests problem is not "inside their knees"

Often dealt with as "inside their heads"

But does it?

Medical knowledge is grossly incomplete



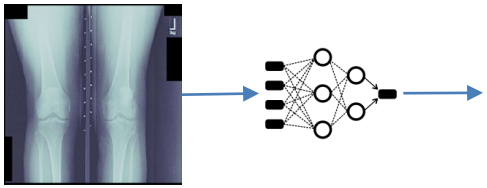
Clinical judgment
KLG Score

Ground Truth
Pain Score

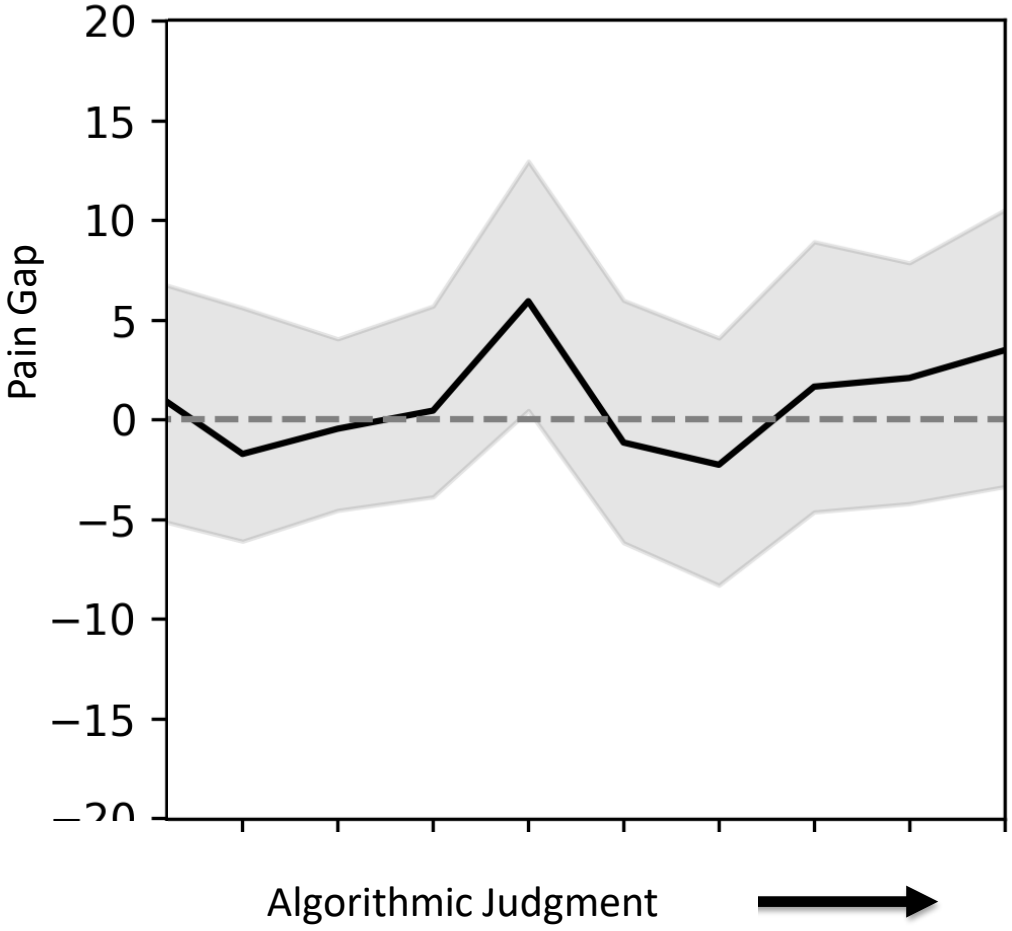
Automation

Discovery

Key: Algorithm
only sees X-rays



Ground Truth
Pain Score



Controlling for algorithmic judgment…

Eliminates much of pain gap

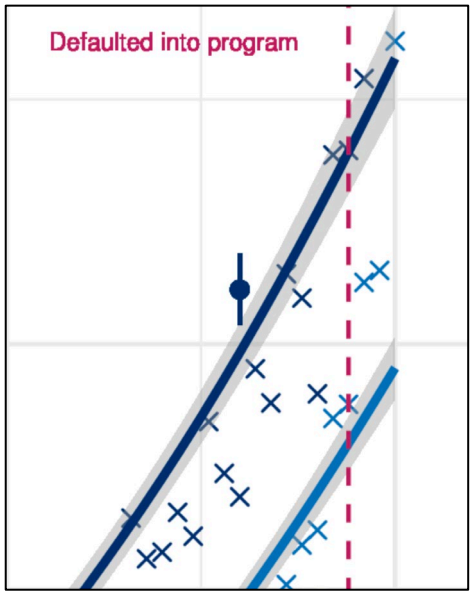
Tragic implication: In their knees all along

Algorithms in one sense better at listening

Algorithms a force for equity

Bias was in the structure of medical knowledge

Notice connection between these two



Racial bias

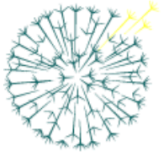
Force for equity

Choice of label dictates outcome

Broader Point

- There is a lot to love and a lot to lament about algorithms
- There is a lot to love and a lot to lament about people
- Poorly designed algorithms can magnify the worst of our tendencies
- But well-designed algorithms can cover for them?

Coda



Dandelion

Data Team Ethics News Contact

Free Public Service for Algorithm Audits