

Notes from webinar called “Reducing Racial Bias in Medical Artificial Intelligence” at SuccessFest23 on January 1<sup>st</sup>, 2023

By David H. Nguyen, PhD

In February of 2019, Time Magazine published an article entitled “Artificial Intelligence Has a Problem With Gender and Racial Bias. Here’s How to Solve It”. This was written by a Computer Scientist named Joy Buolamwini. She is a Black American and founder of the Algorithmic Justice League, which sounds very cool.

In her article, she talks about how AI developed by tech giants, such as IBM, Microsoft, and Amazon failed terribly at predicting gender in pictures of women and people of color.

So, what is the big deal with this problem?

Let’s imagine that law enforcement agencies want to develop an AI that looks at live video footage and calculates the likelihood that someone is a criminal based on their face. What if this AI saw you standing in line at the bank and calculated that you are not just a criminal, but most likely a violent criminal. When it calls the police, the whole SWAT team shows up. Would you want this to happen to you? Of course, not. The same bias that makes law enforcement AI dangerous also make medical AI dangerous.

Medical AI predicts:

- How much pain medication you need.
- How much oxygen you need.
- What life-saving drugs you need.
- What surgical procedures you need.

If it is biased, you and your loved ones will suffer more and die sooner.

Hi, I’m Dr. David Nguyen, the CEO of BrainScanology.

I’m a computational biologist and before that I was a cancer biologist. I got my PhD from UC Berkeley and invented an algorithm that can measure shapes in ways that area and volume cannot. It’s called the LCPC Transform and is the core algorithm of BrainScanology’s revolutionary software.

As someone who studies and creates pattern recognition algorithms, I’ve come to realize that the problem of racial and gender bias in data science is not in the algorithms, but in HOW WE USE the algorithms.

But the main reason why I feel passionately about racial and gender bias in AI is because I grew up on a low-income family of immigrants. Life was hard. The label of being underprivileged was an understatement. But I was fortunate to have great parents, teachers, and mentors, so I'm doing my part to pass things forward.

The topic of this presentation is "Reducing Racial Bias in Medical Artificial Intelligence." There are 4 main points that I want you to walk away with. After describing them, I will discuss each in detail so that you are equipped with the knowledge to ask the right questions.

Point #1. Artificial Intelligence (A.I.) is only as smart as the data that you feed it. If you feed it biased training data, then it makes biased predictions.

Point #2. Each race of people has genetics that affect how symptoms of disease manifest, but A.I. has no clue about this unless we teach it.

Point #3. Do NOT interpret Point #2 to mean that certain races are genetically inferior, because genetics cannot tell us that.

Point #4. YOU, the listener, can do something to reduce the problem of data bias by asking your local and regional lawmakers what they are doing about the problem.

---

Let's dig into Point #1. Artificial Intelligence (A.I.) is only as smart as the data that you feed it.

AI requires its creator to define something called "ground truth." For example, if you want to train AI to recognize dogs vs. cats, you have to give it a batch of images of containing only dogs and a separate batch of images containing only cats. These batches of images are called ground truth categories and is what the AI studies to learn what is true.

The same applies to training an AI to detect lungs that are infected by COVID. You need to provide ground truth pictures of healthy lungs and then ground truth pictures of infected lungs. But what if a disease affects males differently than it does females?

What if a disease causes black people to have heart attacks but white people to have strokes? If you train AI to with data only from patients who are white and male, it will misdiagnose patients who are not white and not male. Misdiagnosis means you will suffer more and die sooner. And all this simply because of your race.

So, racial bias in medical AI is going to be a huge problem. AI is the future of healthcare, so we if we don't do something about it the problem only gets worse.

Firstly, it is an ethical problem because people will be harmed.

Secondly, it is a justice problem because entire groups of people will be discriminated against.

Lastly, it is an economic problem because the US healthcare system is going to be flooded beyond its capacity. By the year 2060, two-thirds of the US population will be people of color. And if our best AI is racist, then it will harm the US economy with misdiagnoses that render people unable to live and work.

---

Let's now talk about Point #2. Each race of people has genetics that affect how symptoms of disease manifest, but A.I. has no clue about this unless we teach it.

For example, some people get diseases earlier in life than others. Latin Americans get dementia 8 years sooner than non-Hispanic White Americans. That means that they have 8 fewer years to plan and provide for the families before dementia kicks in and ruins their ability to feed their families. In this case, don't you think AI should be trained to detect dementia sooner in LatinX people?

Take Alzheimer's Disease for example, 60% of Alzheimer's patients in the US are female, but the vast majority of research participants in Alzheimer's studies are male.

So, not only is racial bias a problem in medical AI, so is gender bias. Is there a reason why grandpa deserves better treatment than grandma?

---

Let's now talk about Point #3. Do NOT interpret Point #2 to mean that certain races are genetically inferior, because genetics cannot tell us that.

We know from research on identical twins that living in different environments affects their biology differently, even though they have the same DNA.

We also know this from studies of animal studies the subjected genetically identical animals to different living conditions.

Your value as a person, your intellectual potential, and your career potential are too complicated for genetics to predict in any meaningful way at the level of race. Some people do inherit mutations that cause disabilities, but this has little to do with the potential of an entire race of people.

On this subject, I've formulated a concept called Heritable Non-Genetic Information, abbreviated as HNI, which I like to pronounce as "honey," so I can say things like, "The solution to this problem, is found ... in the honey." HNI is about a biological code that you inherit, just like DNA, but that is NOT stored in DNA. Turns out there are many biological codes that can be inherited but they are based on 2D and 3D shapes.

So, the next time someone tries to bully your race based on genetics, you can ask them about environmental influences on genetics and ask, "Is it really that simple, honey?"

---

**Let's not dig into Point #4. You, the listener, can do something to reduce the problem of data bias by asking your local and regional lawmakers what they are doing about the problem.**

Lawmakers have to deal with many problems that fight for their attention: homelessness, the economy, fixing potholes in the roads, you name it. So, if you don't voice your concerns about data bias in medical AI, then they won't know that it's important to you.

I will next explain:

What you can do at the community level.

What you can do at the city and state level.

What you can do at the federal level.

### **What can you do at the community level?**

1. Talk to your religious community and social community about participating in medical research studies.

- Some studies will pay you for your participation.
- These studies are coordinated at nearby hospitals.

2. Gather people in your community and create a committee that will learn the best practices of participating in medical research. Your community needs leaders to step up, learn, and teach others.

### **What can you do at the city and state level?**

1. Contact your mayor, city council, and state legislators.

- Tell them to create plans to incentivize medical schools and research centers to collect diverse patient data in their studies. These reward incentives can be in the form of tax breaks and tax exemptions.

2. Join research advocacy organizations and attend their events. Ask local researchers what they are doing against data bias.

Both lawmakers and scientists need to hear a collective concern from society about data bias.

### **What can you do at the federal level?**

1. Contact your representative in congress and your senator.
  - Congress controls the annual budget of the National Institutes of Health, which funds a lot of medical research at universities.
  - Congress is able to create laws requiring the collection of diverse patient pools in medical studies.
  - Congress is able to create penalties against university research labs that don't follow funding regulations that come along with grant funding.

To summarize point #4. Change needs to come from the top. For major shifts in behavior to happen, change has to come from the people who control the research funding.

### **We've now reached the end of this presentation, so I want to reiterate the four main points.**

1. A.I. is only as smart as the data that you feed it.
2. Each race of people has genetics that affect how symptoms of disease manifest, but A.I. has no clue about this unless we teach it.
3. Genetics cannot tell you that one race is inferior to another.
4. YOU, the listener, can do something to reduce data bias by consistently asking your local and regional lawmakers what they are doing about the growing problem.

If you have questions, feel free to reach out to me via email. My address is [dave@tsg-lab.org](mailto:dave@tsg-lab.org)