“Racism is America’s Oldest Algorithm:” How Bias Creeps into Health Care AI

In this written version of the episode, all words, including speakers, ambient sound, effects, and music, will appear in size 11 black type. SPEAKER NAMES are in bold and all caps. Music and other sound descriptions are indicated by [brackets] in regular font.

Speakers:

- **NICHOLAS ST. FLEUR ("NICK")** is the host of Color Code and a science reporter for STAT, where he often covers the intersection of race and medicine. Based in Long Island, NY, he is in his early 30s.
- **CASEY ROSS** is a health tech reporter for STAT. Based in Ohio, he is in his early 40s.
- **ALISSA AMBROSE** is the senior producer for Color Code.
- **ZIAD OBERMEYER** is an emergency medicine physician and researcher at the UC Berkeley School of Public Health.
- **CHRIS HEMPHILL** is the VP for Applied AI & Growth at Actium Health. They host their own podcast, Hello Healthcare.

CASEY ROSS
I have on my phone my copy of “Weapons of Math Destruction” right here. So I'm ready to go. It's a great book by Cathy O'Neil that I would recommend that anyone read.

ALISSA
I'll check that out.

NICK (interviewing)
Yeah. Well, Casey, thanks for joining us. So… [His voice trails off.]

NICK (narrating)
When the Color Code team started to produce this podcast, we didn’t know much about algorithms in healthcare. We’d seen the stories about the racial bias, embedded in kidney and lung testing, but all of us had a hard time wrapping our heads around what that really meant. So we turned to our colleague, Casey Ross, to start breaking it down for us. He’s a health tech reporter at STAT, who writes a lot about the ways in which medical algorithms have been “rife with bias.”

[Soft, rhythmic drums color the background.]

CASEY ROSS
The starting point for me in this is, sort of, you know: Racism is probably America's oldest algorithm, you know, and that's true in so many domains. And it's certainly true in health care
because it bleeds into decision making that's made about people’s care every day. That's the starting point for sort of how I think about it. And, you know, the technology that's available right now and that's being put to use right now can sort of turn in one of two ways: It can really aggressively combat racism, or it can spread it.

**NICK** (narrating)
A couple years ago, Casey wrote about a research study, published in 2019, that uncovered racial bias in one particular healthcare algorithm. The researchers found that the algorithm was not giving special, preventive care equally to Black and white patients, even if they had similar health problems. Each year, the algorithm was used in the care of an estimated 70 million people.

Now, 70 million – that’s an unimaginable number, so Casey wanted to see how an algorithm could affect one town, one neighborhood, or one patient.

**CASEY ROSS**
One woman, who I met there and spent a lot of time with, told me about the care of her elderly mother, who was in her seventies when she started having very serious symptoms of just feeling really fatigued. And she didn't know what was going on. She was healthy. She did yoga and exercised all her life and was thin and all that. So this seemed to come out of nowhere for her. She went to see a physician. Her physician told her that she was depressed. And kind of that was basically the working diagnosis. And then she went back, and she talked to her daughter, and they had a long conversation about it, and she's like, “I'm not depressed. I'm tired, and I'm having these symptoms.” And eventually, working with their daughter and some others – they convinced her to go to Greenville, which is the closest city with a large hospital, and get seen there. And they find out there that she has, you know, advanced heart disease. And that was the cause of her fatigue. And, at that point, she was told she had about, you know, maybe 6 months to live. And she died shortly thereafter.

I mean, what happens if she gets more aggressive care earlier, you know, if somebody recognizes her condition for what it is, when it needs to be recognized, and does something about it right away? You know, those are the missed opportunities that happen all of the time. And those opportunities get missed because of racism, because of bias, because of blind spots. And I’m not trying to suggest that doctors and physicians and people who work in medicine are overtly racist. We all have blind spots, and those blind spots get included in data and then recycled, and bad outcomes result.

**NICK** (narrating)
It’s hard to wrap my brain around the idea of one algorithm, deciding the fate of 70 million people, including this woman. But behind the scenes, these kinds of algorithms help run almost every hospital and clinic in the country. And I just keep thinking: If this is what happened to one person, how many more millions of lives could have been better cared for?
This is Color Code, a podcast from STAT. I’m Nicholas St. Fleur, a science & health reporter here. And over 8 episodes, I’m taking a look at the hidden and not-so-hidden forces behind our country’s stark racial health inequities. This is Episode 6, where we’re talking about racial bias in healthcare algorithms.

NICK (interviewing)
What would an algorithm be in this health sense, in this medical context?

CASEY ROSS
Yeah, so an algorithm is basically a set of instructions that are used to perform a specific task. That’s like your basic definition of an algorithm.

And in medicine, you know, algorithms are formed, using different variables – pieces of information that can be combined with one another in order to assess a patient or a condition, whether they're likely to have a certain condition, or whether they're likely to benefit from a certain treatment. So you're combining all these different pieces of information that are gleaned about patients in the course of getting their care, assembling that basically into a formula that's then used to determine, you know, to perform a task, to arrive at an answer that you need to arrive for that particular person.

ZIAD OBERMEYER (interview clip)
So maybe I'll start by describing some of the work that we did with an algorithm that is extremely widely used in the healthcare system that is sold to, you know, you can imagine someone in a population health management group – someone who's responsible for the health of an entire population of people, to help them do that management task. And so that person's problem is they're responsible for a ton of people.

NICK (narrating)
Ziad Obermeyer and a team of researchers, found one algorithm, employed by Optum, Inc., that had glaring racial blindspots.
And if we knew which ones were going to get sick, we would devote a lot of attention to them now. So there's lots of things that a healthcare system can do to prevent someone from getting sick. But that stuff is generally kind of expensive. So, you know, to have a nurse practitioner come to your house or review your medications or to set aside primary care slots – you know, you can't do that for everyone. So you need to target that set of things – the people who are at high risk of getting sick that you can make a difference for.

So if you imagine there's 2 people, who have the same health, and however you want to measure health like same blood pressure, same number of conditions, whatever you're using, the algorithm scores the Black patient lower on average. And that leads to that Black patient, being deprioritized for access to all of those things that a health system can do to help. And it's a large degree of bias. So, you know, in this one system that we studied, when you look at the high-priority group – the group that's fast tracked into extra help and that is treated like, you know, this is like the VIP group. That group when we studied it was just under 20% Black, but an unbiased algorithm would have been almost 50% Black. So there's a very large degree of bias. And the question was: Why was this happening?

If you were a family doctor, who had cared for someone from infancy to adulthood, maybe you'd have an inkling of whether or not they could get *really* sick in the next year, just based on everything you know about them…

…But for an algorithm, you can't just say, “Find people, who are going to get sick.” What's sick”? What is that variable? There's no variable called “sick”; there's no variable called “health”. And so you could stitch together a ton of stuff, or you could make a simplifying assumption, which is what all these algorithm manufacturers did, which is to say, “Well, when people get sick, they generate healthcare costs. So we're just going to predict health care costs as a proxy for that complex 'health / sick' kind of thing.”

So yeah, the thinking they're having is that people who are sick are going to be expensive – are going to have cost. Is that the idea there?

Yeah, which is a very reasonable assumption. Like, when people get sick, they are more likely to end up in the hospital. They generate spending on all sorts of things. And that happens for both Black and white patients. So on its face, it's really not an unreasonable assumption, until you kind of remember that not everybody, who needs health care, gets health care.

[Light, newsy MUSIC enters in the background.]
And so 2 people, who are equally as sick – if one of them happens to be Black or poor, or in some way, that person faces barriers to accessing health care, they're going to cost less, not because they should cost less, but because they do cost less – because that's the way our health system works, because there are barriers to access. There's structural racism. There's actually, like, interpersonal racism. Doctors just treat people differently, based on the language they speak, the color of their skin.

**NICK** (interviewing)
So do you mean like they're not getting the same number of tests? They're kind of – they're not getting the same, you know, expensive care? They may not be on the list for, like, you know, whether it be like experimental drugs or something like this – like, is that what you mean for why their cost of care is less?

**ZIAD OBERMEYER**
Absolutely. It's all of those things. It's, you know – here's one very, I think, poignant statistic is when you look at the amount of money spent on psychiatric services. If you look at outpatient psychiatry – so you just go see your psychiatrist in the clinic – white patients have huge amounts of care on that relative to Black patients, whereas if you look at, you know – there's a small fraction of patients, who are unlucky enough to meet inpatient psychiatric hospitalization – that's more for Black patients, but that's rare.

And so all of these outpatient costs, routine care, you know, all of that stuff adds up. But so do medications. So do physical therapy appointments, knee replacements, all of these things – some people have more access to them than others. And so when you're predicting how much someone's going to cost, you're not just predicting their health. You're predicting their health and their level of access to health care. And that's where the bias comes in.

The assumption that all of these algorithm developers made is that health care costs are a good proxy for health, and specifically, for health needs. And even though that looks reasonable, because, in general, people who cost more money to the healthcare system have higher needs – the difference between having access to health care and having less access to health care means that 2 people can be equally healthy, but one of them is less likely to get care.

**NICK** (narrating)
Ultimately, the algorithm fast-tracked white patients ahead of Black patients in the queue for medical care.

Ziad and his team reached out to the company. They helped them, pro bono, because they just wanted patients to get better care. Together, they have done work to eliminate the racial bias present in the earlier algorithm.

If it weren't for their research, this algorithm might’ve continued to magnify health disparities on a national scale. And even still, if the researchers hadn't offered to help for free, who knows if the algorithm would’ve been corrected at all?
This new lens, that’s aimed toward health equity in the world of healthcare algorithms – it’s relatively new.

CHRIS HEMPHILL
I grew up – I’m a child of the nineties – well, I’m a child of the eighties, grew up in the nineties. And a lot of the decisions that were made, and a lot of like how people process things wasn't based on data; it was based on intuition and gut feel and how people had been trained and programmed all their lives.

NICK
Chris Hemphill is the Vice President of Applied A.I. & Growth at Actium Health. They think often about a topic called “responsible A.I.” Like Ziad, their work seeks to improve the algorithms of healthcare companies.

CHRIS HEMPHILL
It’s developing A.I. with an understanding of how you might be able to serve and address the needs of the underserved. And responsible A.I is just asking the question: Well, are we perpetuating, like, existing biases that have failed people in the past?

NICK
Chris gave us a rundown on the history of the field.

[An electric electric keyboard plays sparse MUSIC in the style of a waltz.]

CHRIS HEMPHILL
The research is pretty recent. There's a website called arXiv that keeps track of, like, the latest and greatest machine learning and data science papers. And there were basically no papers being published on that site – very low number of papers being published on that site between 2013 and 2019, very low, kind of flatline. 2019, it popped up. And then in 2021, the number of papers coming out has doubled. So responsible A.I. – it's a budding field. It's recent. So it's exciting to be, like, part of this newer field.

NICK
The field has grown so much that, in April, the University of Chicago held a conference called “Responsible A.I. in Health Care,” which Chris attended.

CHRIS HEMPHILL
What shocked me was just like: They had people from the White House in attendance and the Federal Trade Commission. Dr. Alondra Nelson, for example – she is the director of the Office of Science and Technology Policy and gave the opening keynote speech. So to see it elevated to that level was exciting, and the fact that they’re bringing in people not only from that policy and government level, but they brought in, like, big tech, little tech health care leaders and stuff like that. It was fantastic to have all these people collaborating. And as somebody who's a
practitioner and working on this problem, kind of, boots on the ground, fantastic to see the level
of interest – to see that level of interest coming from all these different angles and have the
ability to connect with other practitioners as well.

It becomes kind of a political battle, where there's a certain way that a department may have
been doing something, and they have to kind of meet in the middle on. It's not just about
developing, like, a fair algorithm, but it's also determining how it's going to be used, and how it
impacts other lines of the business. So you can tell the conversation's maturing when it's not
just people just focused on the technology, but overall how it's going to be used in the business,
what the regulatory landscape is going to be.

NICK
As health tech and A.I. in healthcare grow faster and faster, Chris hopes hospital administrators
keep “responsible A.I.” in mind.

CHRIS HEMPHILL
I want them to start thinking about: Well, how do we make health equity a KPI, a key
performance indicator, and then how do we enlist our technology to help support that
organizational goal? So the more healthcare leaders that are, first of all, learning about this stuff
and then becoming excited and then introducing it as an overall strategic goal, I think the more
likelihood that we have to help address some of the issues, like people being invisible to
algorithms and such.

NICK
There's one algorithm that Ziad told us about that is trying to combat the racial biases we see on
these massive scales. It's for arthritis in people's knees. When a radiologist looks at an x-ray for
arthritis, they see how small the joint space has become with age, or whether there were any
new bony projections that could be signs of disease. Using that information, they would score
just how bad that person's arthritis is: 0 being the best, and 4 being the worst. That scale is
called the Kellgren-Lawrence system.

ZIAD OBERMEYER
And that is a system of grading the degree of arthritis in someone's knee, according to very
established criteria. And they're so established that actually, they go back to the 1940s and 50s
when doctors were studying coal miners in Lancashire, in England, and looking at what their
knees looked like on an x-ray and correlating that to pain scores. And when you go back to
those original studies, you find that they didn't even comment on the sex and race breakdown of
those patients because they were all the same. They were all white men. And so that might
make you worry that, well, maybe that Kellgren-Lawrence grade is missing things that affect
nonwhite people, that affect different populations from the one that was being studied.

NICK
To this day, the Kellgren-Lawrence system is widely used for patients of all backgrounds,
despite the lack of diversity in the original study.
ZIAD OBERMEYER
And so one of the hypotheses that we had was that that might be responsible for a kind of mystery in the literature, which is that, if you compare two patients – one Black, one white – whose knees look the same to a radiologist, Black patients on average report more pain, even though their knees look the same. So there’s this gap in pain where Black patients seem to be reporting extra pain that can’t be accounted for by what the radiologist is seeing.

And so what we did is we actually trained an algorithm to not just spit back out what the radiologist would say about that x-ray, but to listen to the patient and to look at a knee and make a prediction about: Does the patient report pain in the knee? And so by switching that variable that the algorithm is learning from, we’re switching the algorithm from something that learns from the doctor to something that listens to the patient. And when we did that, what we found is that we could explain a lot more of the extra pain that Black patients were feeling. So this algorithm was finding something in the knee that radiologists were missing that was disproportionately accounting for the pain that Black patients were reporting, but that radiologists weren’t finding a reason for.

NICK
I feel encouraged by talking to Ziad, someone who is helping to pave a path forward for algorithms without bias. And according to my colleague Casey, other researchers are trying to do the same.

CASEY ROSS
Institutions across the country are, I think, right now trying to excavate the racism that’s within the algorithms that are in use by their institutions. The University of Utah is doing that right now. I mean, they’re looking through microfilm, you know, to sort of find, you know, yellowing copies of studies published decades ago to figure out, “Okay, this was actually not validated on a diverse group of people. We need to go back and examine how its use is impacting recommendations for care.” And that kind of work is really crucial, and I think it’s beginning to happen more.

But the problems are it’s hard to get the data. You know, it’s hard to get enough data on enough patients of diverse-enough populations to be able to validate algorithms that, you know, sort of an early testing have showed promise. Like, one example is a breast cancer risk prediction algorithm that was developed by folks at Mass General at MIT. And this is actually a really positive example because they basically, literally went to the ends of the earth to validate the thing. I mean, they went to Brazil; they went to Taiwan; they went to Emory University in Georgia; they validated it in Boston – you know, all kinds of, you know, but that’s what it takes, right?

NICK
Even if an algorithm is developed with a diverse population of subjects, problems can still arise later.
CASEY
Yeah. No, absolutely. It's a huge problem, you know, because time passes, and populations change, and diseases change. New diseases pop up. Coronavirus. You know, all these other confounding factors can kind of change the underlying data in a way that renders the output of the algorithm incorrect or skewed in some way.

NICK
About $6 billion dollars were invested in digital health startups in the first 3 months of 2022 – so the money is there, and this area of healthcare is booming. But let’s not forget that since these technologies can affect people all around the world, it’s important to get this right. Ultimately, these algorithms are only as good as the data that we pour into them. As the saying goes, if it’s garbage in, then it’s garbage out.

So, to ensure that healthcare of the future works for every person, that data should be representative of every person. And it all starts with who’s invited to take part in these studies in the first place. In 2 weeks, we’ll be taking a look at diversity in clinical trials and speaking with experts about how to make sure that science doesn’t fall short.

[Full, uplifting music, complete with a keyboard, percussion, and more, rises. THEME MUSIC closes the episode and takes us to the credits.]

Thank you for listening and being part of our Color Code community.

Our team here at STAT is Alissa Ambrose, Hyacinth Empinado, Theresa Gaffney, Crystal Milner, and me, Nick St. Fleur. Kevin Seaman is our engineer, and Tino Delamerced is our intern. Our theme music is by Bryan Joel.

Special thanks to Casey Ross, Chris Hemphill, and Ziad Obermeyer.

Thanks to the Commonwealth Fund for supporting this podcast.

After every episode, we’ll have a bunch of photos and some more reading related to the episode’s topic at STATnews.com, so please: go check it out! We’ll have a new episode in two weeks.

If you like the podcast, please leave a review, and subscribe! And if you have any thoughts for us, you can reach us at ColorCode@statnews.com.

[THEME MUSIC ends on a final note.]

CASEY ROSS
Like, machine learning is dumb. It's dumb; it is dumb; it is dumb. It's not that smart. Everybody thinks it's, like, “Oh, machine learning!” It's stupid. Fundamentally, it really is. It's not like, “Think
of a pipe. It's like, okay, you put the data in here, and then it goes through the pipe, and it gets processed, and it comes out the other side, and poof! There's your answer."

But it's like, the algorithm and the AI is stupid. Like, if you showed my two year old a cat once, she would know it, man! She knows – it doesn't matter if the cat is in a tree or under a bush or in a hole. She knows it's a cat. The machine learning algorithm will be like, “I think it's a dog this time.” It's not a dog!