September 10, 2021

National Institute of Standards and Technology
U.S. Department of Commerce
100 Bureau Drive
Gaithersburg, MD 20899-2000

delivered via email to ai-bias@list.nist.gov

RE: Draft NIST Special Publication 1270, A Proposal for Identifying and Managing Bias within Artificial Intelligence

To Whom It May Concern:

The National Health Law Program (NHeLP) is a public interest law firm working to advance access to quality health care and protect the legal rights of low-income and under-served people. NHeLP has a long history of advocacy regarding the use of artificial intelligence (AI) and automated decision-making systems (ADS) to determine health care eligibility and services.

NHeLP also has a long history of litigation and advocacy on issues that involve ADS in Medicaid including problematic eligibility computer systems that fail to timely and accurately determine eligibility or that wrongfully terminate eligibility; ADS that limit needed health care services; complex assessment systems for home and community based service level of care and resource allocation decisions; and a variety of ADS-related issues in between. Throughout this advocacy we commonly encounter states struggling to make changes to systems, including to provide legally sufficient notices to individuals affected by the systems. We attach and incorporate our

comments from earlier this year to AHRQ on algorithmic bias in healthcare; they discuss in more detail our experiences with bias in ADS and the impact on individuals.

Our advocacy has reinforced the importance of the meaningful due process with ADS. We have not seen an ADS that is always able to accurately determine the needs of every individual it encounters. We are particularly familiar with how the lack of transparency of many ADS impacts individuals' ability to understand and fight care denials or reductions.[1] Our experience with ADS has also led us to identify common ways in which ADS harms individuals and how decisions made along the lifecycle of a system—and the policies around its use—embed bias and other harms. We appreciate the opportunity to comment on the Proposal for Identifying and Managing Bias in Artificial Intelligence ("the Proposal") by the National Institute of Standards and Technology (NIST).[2]

Overarching Concerns

The Proposal acknowledges bias is present throughout AI systems and that the impacts of harmful biases must be managed and reduced across contexts. We appreciate this acknowledgement and support the idea that there are sources and points of bias introduction throughout the lifecycle of AI. However, our experience with ADS in Medicaid, and our knowledge of the impact of those systems in other public benefits programs, have shown us that standards or solutions that are exclusively technical are insufficient. A technical approach seems to suffer from the same issues and assumptions that many of the ADS systems themselves in terms of assumptions about being able to fully understand and identify issues with purpose, function, and impact, as well as the ability to control for bias. While we appreciate the acknowledgment that it is unlikely technology exhibiting "zero risk" can be developed, we think that a proposal for managing bias in AI needs a broader approach than just the AI systems themselves. It should include the processes and procedures that should be in place when AI is used. By way of example, this would include disclosure of the use of AI, the intended purpose or scope of the system, any limitations or known biases of an AI system, the right to dispute a decision, and how a person can ask to be an exception from the outcome of AI. We would also recommend that the lifecycle must include a very clear understanding from

---

[1] *See* Elizabeth Edwards, Nat'l Health Law Program, *Preventing Harm from Automated Decision-Making Systems in Medicaid* (June 14, 2021), https://healthlaw.org/preventing-harm-from-automated-decision-making-systems-in-medicaid/ (and the linked resources at the bottom of the page to related resources from NHeLP on the impact of ADS).
[2] National Institute of Standards and Technology, Proposal for Identifying and Managing Bias in Artificial Intelligence (SP 1270), https://www.nist.gov/artificial-intelligence/proposal-identifying-and-managing-bias-artificial-intelligence-sp-1270.

the outset about whether AI is appropriate to be used for the task. We recognize that the Proposal touches on this as part of the lifecycle, but we would recommend that it be a much more robust part of the process.

We also raise concerns about the lack of inclusion of safeguards around AI as it may be used in populations for which it may not be appropriate. This was partly identified as an issue, but we have seen significant issues with AI being used by people with disabilities or people who prefer languages other than English. Problems can range from usability of the system, to questions not translating properly or being understood well, all of which may lead to outcomes other than what should have occurred based on the rules or logic of the system. We recommend that the Proposal explicitly identify issues with language and disability that may impact the fairness of system and thus the need for related guardrails.

We recommend that advocates, attorneys, and individuals negatively impacted by ADS be included in these discussions. This should include not just one sector, such as policing or housing rights, but the array of subject areas as the impacts and sources of bias are varied and have different solutions. NHeLP has been involved in various national and international events convening private attorneys, legal services advocates, academics, technologists, and individuals impacted by ADS from across different issues. While the problems with ADS are often similar, the ideas about solutions and what to prioritize in terms of negative impacts to prevent differ widely. We urge NIST to consider a broader perspective that includes more socio-technical standards.

While the Proposal acknowledges this broader need somewhat in lines 369-395, we ask that greater weight be given to the social factors. We acknowledge that the natural tendency when inquiring about bias in AI is to dig into the AI system itself. However, wet recommend greater focus on why a system is being implemented, for what purpose (is it to save money or is it really designed for the benefit of the people it impacts?), and the impact on the people that will be affected centered in the analysis of the AI system. Intellectualizing the system allows reviewers of that system to become dangerously far removed from the people impacted. In our experience, a statistically insignificant source of bias or a minor error in code can have a drastic impact on an individual's health, them becoming institutionalized against their will, or losing needed medical coverage leading to irreversible damage to their health. Similarly, in other public benefits contexts an improper accusation of unemployment fraud or an inability to access benefits owed can set off the first in a set of dominoes of harms from which an individual may never fully recover. We are concerned that a continued focus on the AI systems

and developers and researchers will fail to fully include the impacts on individuals or prevent the harms caused by AI.

While we largely agree with the reasons for public distrust of the AI-related bias identified starting on line 234, we think a significant piece of the analysis of mistrust is missing. In our experience, there is significant mistrust because of the lack of transparency and explainability. Oftentimes when AI is used in the public benefits context people are not fully aware that AI has made the decision about their benefits. While this lack of transparency certainly creates a sense of confusion and mistrust of the system, it more importantly violates the constitutional rights of individuals.[3] The Proposal acknowledges why there may be mistrust and the need for explainability, but it fails to acknowledge that in some instances this is not merely a best practice, but a constitutional right. This is a major omission. When AI is used in the public benefits context, the standards that have been used to make a decision must be more than explainable at a high level. They are constitutionally required to be explained such that a person affected by the benefit can understand the decision made to deny, reduce, or terminate their benefits to a sufficient extent so that they may be able to adequately challenge such a decision. While we certainly encourage greater accountability and transparency of AI systems, we think the failure to acknowledge the constitutional requirements around the use of AI systems in certain areas is a significant missing piece of the analysis and the plan for developing safeguards around the use of AI.

Recommendations Regarding the AI Lifecycle Analysis

More Critical Questions of the Use of AI Must Be Asked at the Beginning of the Lifecycle

In some of NHeLP's advocacy, we also have taken a lifecycle approach and support much of the Proposal's identification of issues and approaches. However, we would encourage much more rigorous scrutiny at the earliest stage of the lifecycle when such systems are being used for public benefits or other "high-stakes settings." While the Proposal acknowledges that "Technology designed for use in high-stakes settings requires extensive testing to demonstrate valid and reliable performance" (lines 468-69), this assumes the acceptability of AI being used in high-stakes settings. Our first concern is what the Proposal considers to be "high-stakes settings." In our experience, an incorrect or biased decision about Medicaid benefits can be devastating to an individual's health and well-being. This is true for most public benefits as the

---

[3] Jane Perkins, Nat'l Health L. Program, *Demanding Ascertainable Standards: Medicaid as a Case Study*, 2016 CLEARINGHOUSE REV. 1 (2016), https://healthlaw.org/resource/demanding-ascertainable-standards-medicaid-as-a-case-study/.

U.S. Supreme Court has acknowledged that people have a "brutal need" for such benefits, which is why the identified constitutional due process protections apply.[4] Our second concern is that the Proposal skips to the need for rigorous testing of AI in such settings rather than asking if AI or certain types of AI should even be allowed in such settings. For example, should AI that seeks to limit benefits or seek out fraud by beneficiaries even be allowed in public benefits when evidence has shown that more people get harmed by many of these systems? Should those AI systems that are intentionally built to include difficult technological hurdles be allowed?[5] If a system is complex and cannot easily explain how a decision was made, can it even be allowed in public benefits since it cannot provide sufficiently simplified explanation to satisfy constitutional due process? Can an AI system thwart the statutory objective of the program, which in Medicaid is primarily the provision of medical assistance, if it by design limits coverage or services in ways that are contrary to the program?[6] We think much harder questions need to be asked early in the lifecycle and the Proposal's position is much too accepting of the use of AI, particularly in certain high-stakes settings.

Greater Involvement of People Impacted is Needed

The Proposal acknowledges the need for addressing potential societal impacts and trying to address biases early in the problem formulation. We appreciate that that the Proposal acknowledges that such efforts are complicated by the role of power and decision making. NHeLP's experience with AI shows that the AI systems compound existing austerity approaches to public benefits. Where open cuts to benefits may not be politically palatable, AI has been used to surreptitiously deprive people of benefits by increasing the rationing of resources, falsely identifying widespread fraud, and frustrating enrollment through technology-enabled administrative burdens. If people impacted by AI systems in Medicaid and other public benefits programs were designing such systems, they would likely take drastically different approaches or perhaps not even allow them at all. But oftentimes public benefits beneficiaries do not have meaningful seats at the table or the knowledge necessary to meaningfully evaluate an AI system. The benefit of transparency is limited without the time or expertise to

---

[4] Goldberg v. Kelly, 397 U.S. 254 (1970).

[5] *See, e.g.,* Tampa Bay Times Editorial Board, *More Evidence that Florida's Unemployment System Was Designed to Fail,* TAMPA BAY TIMES (Mar. 23, 2021), https://www.tampabay.com/opinion/2021/03/23/more-evidence-that-floridas-unemployment-system-was-designed-to-fail-editorial/.

[6] 42 U.S.C. § 1396-1; 1396d(a); *see also Stewart v. Azar,* 366 F. Supp. 3d 125, 138 & 144 (D.D.C. 2019) (finding that the Medicaid Act is "designed . . . to address not health generally but the provision of care to needy populations").

NATIONAL
HEALTH
LAW PROGRAM

analyze what is available. Moreover, they often lack the political power to influence, much less stop a dangerous AI system.

While greater transparency in AI is necessary and helpful, part of the lifecycle process must incorporate the perspective and needs of the people impacted by the AI system. Importantly, it should center the needs of the persons the system affects, as opposed to how a developer, purchaser, or end user of an AI system perceives the needs of the person impacted. We recommend that the Proposal make changes to the practices discussed in lines 545-58 to move beyond the inclusion of subject matter experts and practitioner end users and include methods that will ensure the needs of those who are impacted are properly included, and not just from the perspective of consultants or researchers who may study those needs. We also recommend that the Proposal include mechanisms that would clearly identify the intended use and scope, including any limitations, of a given AI system. Labelling and other forms of disclosure, which should include disclosure to people impacted by an AI system, would more clearly identify known issues, bias, or other limitations of a given system.

The Deployment Stage Should Be Inclusive of Ongoing Quality

As described in the enclosed comments, NHeLP has experienced significant problems with the deployment stage of AI systems. Based on our experience, AI used in Medicaid is often not well-tested or validated for the particular use or may have significant errors that are not caught until well into implementation and sometimes cannot be easily fixed, causing significant harm to those impacted. We would suggest greater robustness in the Proposal delineating meaningful guardrails for the pre-deployment and deployment phase of AI systems. We know that many AI systems are not built with sufficient audit trails or other mechanisms that can easily identify issues or run reports to find problems. We have also found that oftentimes an AI system cannot easily be turned off or overridden. NHeLP's experience, and that of the state advocates with whom we work, has also shown us that there is often little ongoing evaluation of the quality and use of AI systems. We recommend that the Proposal be more robust in the lifecycle of AI beyond deployment.

The Need for Systems and Protections Around AI Systems

As the Proposal acknowledges, creating a zero risk AI system is highly unlikely. Thus, the safety net processes around the AI systems are critically important. For Medicaid and other public benefits, these safety net processes are constitutionally required. Our work has shown that although required, these protections are often not well implemented. The experiences of how the due process requirements have worked and failed to protect individuals from AI harm

should inform the Proposal. And relatedly, the Proposal must recognize that there are existing legal requirements certain AI systems must meet. If we acknowledge that AI is not zero risk, then the policy solutions and protective frameworks around these systems must include more than just methods of making the AI better.

We would be happy to discuss these comments and our experiences further. We are also connected to legal services advocates and others who have experience with the impact of AI. Please contact Elizabeth Edwards (edwards@healthlaw.org) for any follow up to these comments, including the attached comments we previously made to AHRQ.

Sincerely,

Elizabeth Edwards
Senior Attorney

Attachment: NHeLP AHRQ Comments on Algorithmic Bias in Health Care