



May 4, 2021

David Meyers, MD
Acting Director
Agency for Healthcare Research and Quality (AHRQ)
Office of Communications
5600 Fishers Lane, 7th Floor
Rockville, MD 20857

Re: Request for Information (RFI): Use of Clinical Algorithms That Have the Potential To Introduce Racial/Ethnic Bias Into Healthcare Delivery

Acting Director Meyers:

The American Medical Informatics Association (AMIA) is pleased to provide input on AHRQ's request for information (RFI) on the use of clinical algorithms that have the potential to introduce racial and/or ethnic bias into healthcare delivery.

AMIA is the professional home for more than 5,500 informatics professionals, representing frontline clinicians, researchers and public health experts who bring meaning to data, manage information, and generate new knowledge across the health and healthcare enterprise. As the voice of the nation's biomedical and health informatics professionals, AMIA plays a leading role in advancing health and wellness by moving basic research findings from bench to bedside, and evaluating interventions, innovations, and public policy across settings and patient populations.

We are grateful that AHRQ is discharging its mission from Congress to explore how clinical algorithms can contribute to health disparities. However, we note that in the absence of sustained remedial analysis, algorithms are only as good as the data used to train them and the developers' level of understanding of the underlying characteristics of the data. Algorithmic design cannot overcome inherent bias in the training data or faulty assumptions about what the data represents without intense focused effort. While we can have the most carefully curated data and attempts to "correct" algorithms, **data creation and collection and algorithm development, deployment, and evaluation do not occur in a value-free vacuum.** There is little evidence that algorithms that claim to be race-blind have taken the necessary steps to compensate for implicit bias in their data, especially as current measures to mitigate bias in algorithms are still nascent and inadequate. It is likely not possible for any algorithm built and trained on underlying data that are biased as a result of systemic and structural inequities to be truly free of bias.

As we reference in our detailed comments below, there are systematic steps and approaches that can be helpful in both recognizing and reducing the impact of bias on the output and use of such algorithms. These efforts, however – while vitally necessary – are not themselves sufficient steps in the wider societal effort to root out racial and ethnic bias in healthcare and beyond.

Defining Algorithm Use

We additionally note that while FDA maintains an inventory of its approved clinical algorithms, this provides an incomplete picture.¹ Innumerable homegrown and/or proprietary algorithms exist outside of this inventory, with variable scrutiny that depends on at least the developer and the institution at which they are used. In fact, even those that are FDA-approved are more likely to have originated from industry and/or proprietary sources, and thus are less likely to have been derived or validated against populations with a high proportion of underserved patients.

We believe that it would be more worthwhile for AHRQ to frame its evidence review by first examining the types of applications of the myriad algorithms, including patient-facing use cases in traditional clinical settings and for virtual care and remote patient monitoring in the home. These applications include but are not limited to: 1) prevention and detection of disease; 2) triage for the appropriate level of care; 3) diagnosis; 4) imaging/pathology; 4) prognosis; 5) treatment, including virtual therapeutics; 6) care planning and transitions; 7) discharge; and 8) remote monitoring. AHRQ should also be aware that some areas blur the line between the clinical and the operational, and between clinician and patient. For example, a health system may employ a resource allocation algorithm, which may result in information that will later be utilized or consulted by treating clinicians, or perhaps more significantly, may have a direct impact on access to care altogether.

Below, we share additional comments in response to selected RFI questions. Thank you for considering our comments. Should you have questions about these comments or require additional information, please contact Scott Weinberg, Public Policy Specialist at scott@amia.org or (240) 479-2134. We look forward to continued partnership and dialogue.

Sincerely,



Patricia C. Dykes, PhD, RN, FAAN, FACMI
Chair, AMIA Board of Directors
Program Director, Research
Center for Patient Safety, Research, and Practice
Brigham and Women's Hospital

¹ Benjamins, S., Dhunoo, P. & Meskó, B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *npj Digit. Med.* 3, 118 (2020). <https://doi.org/10.1038/s41746-020-00324-0>

AHRQ Questions	AMIA Response
<p>What clinical algorithms are used in clinical practice, hospitals, health systems, payment systems, or other instances? What is the estimated impact of these algorithms in size and characteristics of population affected, quality of care, clinical outcomes, quality of life, and health disparities?</p>	<p>As noted in the preamble, there are numerous clinical algorithms currently in use, only a fraction of which are FDA-approved or even under the purview of the FDA. The impact of these algorithms on the size and characteristics of the population affected, quality of care, clinical outcomes, quality of life, and health disparities are as variable as their intended use and the health settings where they have been deployed.</p>
<p>Do the algorithms in question 1 include race/ethnicity as a variable and, if so, how was race and ethnicity defined (including from whose perspective and whether there is a designation for mixed-race or multiracial individuals)?</p>	<p>All algorithms depend upon the underlying data on which they are developed and trained, tested, validated and deployed. Thus, the definition of race/ethnicity will inevitably vary depending both on how the algorithm is built, as well as the data on which it is trained and deployed, which might not have the same characteristics.</p>
<p>Do the algorithms in question 1 include measures of social determinants of health (SDOH) and, if so, how were these defined? Are these independently or collectively examined for their potential contribution to healthcare disparities and biases in care?</p>	<p>Some algorithms do indeed include measures of SDOH, however, this too is dependent on the data source and the questions being asked. While there are a variety of ways that SDOH data may be considered, there is little maturity in standards for SDOH data which are often missing or incomplete. This variability makes it difficult to evaluate biases. This lack of standardization, in turn, leads to use of available, but incomplete and invariably flawed data.</p>
<p>For the algorithms in question 1, what evidence, data quality and types (such as claims/utilization data, clinical data, social determinants of health), and data sources were used in their development and validation? What is the sample size of the datasets used for development and validation? What is the representation of Black, Indigenous, and other people of color and what is the power to detect between-group differences? What methods were used to validate the algorithms and measure health outcomes associated with the use of the algorithms?</p>	<p>This too will depend on the setting. Electronic health records (EHR) are available and are most commonly used in health systems. Depending on the project, experts, and budget available, other data sources, most notably US Census data, are used, as well.</p>
<p>For the algorithms in question 1, what approaches are used in updating these algorithms?</p>	<p>Informatics plays a crucial role in algorithm problem definition, solution design, validation of data sufficiency and computational</p>

	<p>methods, and effectiveness studies. There are multiple levels of metrics that can be used to evaluate effectiveness: (1) how does the AI solution perform on its specific task (accuracy, sensitivity/specificity, F-1, etc.); (2) how does integrating the AI solution into clinical care affect process efficiency; (3) how does integrating the AI solution affect care outcomes for the health system (mortality, readmission, cost), provider (burden), and for the individual (morbidity, disability/disease burden, quality of life measures, as appropriate).</p> <p>Further, EHR and patient-reported outcomes data should be validated against the algorithms to see how well they responded to expected or unexpected health effects. This systematic monitoring is known as “algorithmovigilance,” which is one important way to maintain quality, minimize harm, and promote trust in healthcare AI.² Systematic monitoring should be complemented by a robust auditing trail documenting the decision points in algorithm development (e.g., choice of the issue the algorithm is intended to address, selection of datasets) and the rationale for each decision, including why some approaches were discarded. This documentation supports transparency while also looking forward to the need to create the descriptive materials needed to explain the algorithm to audiences with varying levels of specialist knowledge, from informaticians and clinicians to patients and the general public.</p> <p>We additionally point you to a recently published AMIA position paper that lays out an informatics-led policy framework for adaptive (AI/Machine Learning) clinical decision support</p>
--	---

² <https://www.regenstrief.org/article/algorithmovigilance-monitoring-healthcare-ai/>

	<p>(CDS) tools.³ In it, we call for identification of two policy concepts: transparency metrics and communications standards. Transparency metrics would describe how Adaptive CDS algorithms are trained, including the data acquisition processes (e.g, patient cohort selection criteria) and preprocessing or “data wrangling” steps that must be clearly documented. Communications standards articulate the components of the Adaptive CDS and describe the intended use(s) and expected user(s), similar to US Food and Drug Administration’s (FDA’s) prescription drug-labeling requirements.⁴ We believe that this should not only apply to CDS, but to any AI-related algorithms, whose development should be transparent and should be monitored publicly or to the extent possible (for proprietary options).</p>
<p>Which clinical algorithms have evidence that they contribute to healthcare disparities, including decreasing access to care, quality of care or worsening health outcomes for Black, Indigenous, and other people of color? What are the priority populations or conditions for assessing whether algorithms increase racial/ethnic disparities? What are the mechanisms by which use of algorithms contribute to poor care for Black, Indigenous, and other people of color?</p>	<p>AHRQ references the landmark Obermeyer et al. study. However, we point to more of the recent literature on recognizing and mitigating racial bias in AI algorithms:</p> <p>Vyas DA, Eisenstein LG, Jones DS. <i>Hidden in Plain Sight - Reconsidering the Use of Race Correction in Clinical Algorithms</i>. N Engl J Med. 2020;383(9):874-82.</p> <p>Marcelin JR, Siraj DS, Victor R, Kotadia S, Maldonado YA. <i>The Impact of Unconscious Bias in Healthcare: How to Recognize and Mitigate It</i>. J Infect Dis. 2019;220(220 Suppl 2):S62-S73.</p> <p>Ntoutsis E, Fafalios P, Gadiraju U, Iosifidis V, Nejdil W, Vidal ME, et al. <i>Bias in data-driven artificial intelligence systems—An</i></p>

³ Carolyn Petersen, Jeffery Smith, Robert R Freimuth, Kenneth W Goodman, Gretchen Purcell Jackson, Joseph Kannry, Hongfang Liu, Subha Madhavan, Dean F Sittig, Adam Wright, Recommendations for the safe, effective use of adaptive CDS in the US healthcare system: an AMIA position paper, *Journal of the American Medical Informatics Association*, Volume 28, Issue 4, April 2021, Pages 677–684, <https://doi.org/10.1093/jamia/ocaa319>

	<p><i>introductory survey</i>. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 2020;10(3).</p> <p>Leavy S, O'Sullivan B, Siapera E. <i>Data, Power and Bias in Artificial Intelligence</i>. arXiv:200807341 [csCY]. 2020.</p> <p>Parikh RB, Teeple S, Navathe AS. <i>Addressing Bias in Artificial Intelligence in Health Care</i>. JAMA. 2019.</p> <p>Amodei D, Olah C, Steinhardt J, Christiano P, Schulman J, Mané D. <i>Concrete Problems in AI Safety</i>. arXiv:160606565 [csAI]. 2016.</p> <p>Chen IY, Szolovits P, Ghassemi M. <i>Can AI Help Reduce Disparities in General Medical and Mental Health Care?</i> AMA J Ethics. 2019;21(2):E167-79.</p> <p>Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. <i>Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data</i>. JAMA internal medicine. 2018;178(11):1544-7.</p> <p>Paulus JK, Kent DM. <i>Predictably unequal: understanding and addressing concerns that algorithmic clinical prediction may increase health disparities</i>. NPJ Digit Med. 2020;3:99.</p> <p>Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. <i>Ensuring Fairness in Machine Learning to Advance Health Equity</i>. Ann Intern Med. 2018;169(12):866-72.</p>
<p>To what extent are users of algorithms, including clinicians, health systems, and health plans, aware of the inclusion of race/ethnicity or other variables that could introduce bias in these algorithms and the implications for clinical decision</p>	<p>The extent to which clinicians are aware of sources of bias is highly dependent on the setting. However, in our members' experience, there is currently little patient engagement or shared decision making when it comes to the use of algorithms. We</p>

<p>making? What evidence is available about the degree to which the use of clinical algorithms contributes to bias in care delivery and resulting disparities in health outcomes? To what extent are patients aware of the inclusion of race/ethnicity or other variables that can result in bias in algorithms that influence their care? Do providers or health systems communicate this information with patients in ways that can be understood?</p>	<p>note that providers themselves first need to understand what goes into algorithms before they can explain them to patients (see our response to question 5 above regarding the usefulness of documentation). This is similarly not a topic that is currently well-covered in medical training, though we believe there is interest and need.</p>
<p>What are approaches to identifying sources of bias and/or correcting or developing new algorithms that may be free of bias? What evidence, data quality and types (such as claims/utilization data, clinical data, information on social determinants of health), data sources, and sample size are used in their development and validation? What is the impact of these new approaches and algorithms on outcomes?</p>	<p>See our response to question 5 above. However, as some of these approaches are currently still in development, it is premature to discuss effects on clinical outcomes.</p>
<p>What challenges have arisen or can arise by designing algorithms developed using traditional biomedical or physiologic factors (such as blood glucose) yet include race/ethnicity as a proxy for other factors such as specific biomarkers, genetic information, etc.? What strategies can be used to address these challenges?</p>	<p>A challenge that has arisen is the question of whether “older” algorithms need to be reassessed when their applications may not appear problematic from a bias standpoint when deployed for care, but may be more problematic under different circumstances. For example, race/ethnicity may effectively be an unintentional proxy for other factors – such as the impact of socio-economic disadvantage on health status, behaviors or outcomes. This is particularly true in communities where the rate and impact of poverty fall disproportionately on minority populations. A recent paper in the <i>New England Journal of Medicine</i> described a variety of illustrative examples in which the use of the variable “race” in clinical prediction models proved to be problematic. In many cases, other variables that were highly correlated with race may have been the true predictors, not race per se.⁵</p>

⁵ Vyas DA, Eisenstein LG, Jones DS. Hidden in Plain Sight - Reconsidering the Use of Race Correction in Clinical Algorithms. *N Engl J Med.* 2020;383(9):874-82.

	<p>A related challenge is that of misclassification/misidentification of race due to missing data. For example, race and/or ethnicity can be assumed based on name, bringing potential biases in its wake. We refer you to additional recent literature on this phenomenon:</p> <p>Bertolli J, Lee LM, Sullivan PS. <i>Racial misidentification of American Indians/Alaska Natives in the HIV/AIDS Reporting Systems of five states and one urban health jurisdiction, U.S., 1984-2002</i>. Public Health Rep 2007;122(3):382-92 doi: 10.1177/003335490712200312</p> <p>Grafova IB, Jarrín OF. <i>Beyond Black and White: Mapping Misclassification of Medicare Beneficiaries Race and Ethnicity</i>. Med Care Res Rev 2020:1077558720935733 doi: 10.1177/1077558720935733</p> <p>Jarrín OF, Nyandegé AN, Grafova IB, Dong X, Lin H. <i>Validity of Race and Ethnicity Codes in Medicare Administrative Data Compared With Gold-standard Self-reported Race Collected During Routine Home Health Care Visits</i>. Med Care 2020;58(1):e1-e8 doi: 10.1097/mlr.0000000000001216</p> <p>Labgold K, Hamid S, Shah S, et al. <i>Estimating the Unknown: Greater Racial and Ethnic Disparities in COVID-19 Burden After Accounting for Missing Race and Ethnicity Data</i>. Epidemiology 2021;32(2):157-61 doi: 10.1097/ede.0000000000001314</p>
<p>What are existing and developing standards (national and international) about how clinical algorithms should be developed, validated, and updated in a way to avoid bias? Are you aware of guidance on the inclusion or race/ethnicity, related</p>	<p>Although there are no current standards, this is a rapidly developing topic and requires careful attention to possible introduction of bias at every step, including: 1) initial formation of the question being asked or problem being addressed by the algorithm; 2) choice of data to use in its development, in</p>

<p>variables such as SDOH, prior utilization, or other variables to minimize the risk of bias?</p>	<p>particular the specific features or data elements used to operationalize the concept of interest; 3) the way the algorithm is intended to be deployed, by whom, in what settings and on what populations; 4) the appropriate application of de-biasing and bias mitigation techniques; 5) assessment of performance and impact both in the short term and over time.</p>
<p>To what extent are users of clinical algorithms educated about how algorithms are developed or may influence their decision making? What educational curricula and training is available for clinicians that addresses bias in clinical algorithms?</p>	<p>Education varies widely depending on the specific organization. Larger health systems, especially those affiliated with universities, may have access to relevant experts, while smaller ones may not have any experts or be able to hire one. Other healthcare systems might have independent research departments that could help the clinicians and even provide training. Nonetheless, it is our members' experience that virtually no clinicians and only a few medical informatics programs provide curricula and/or training in addressing bias in algorithms.</p>