Viewpoint

Understanding Race as a Key Analytic Dimension in AI-Enabled, Health-Related Social Engagement Online and Its Role and Impact in Health Research

Rachele M Hendricks-Sturrup[1], MSc, MA, DHSc, FACTS; Gabriella Waters[2], PhD

[1]Robert J. Duke-Margolis Institute for Health Policy, Duke University, Washington, DC, United States

[2]Cognitive and Neurodiversity AI Lab, Robotics and Digital Twin Lab, Center for Responsible AI, Virginia State University, Petersburg, VA, United States

Corresponding Author:
Rachele M Hendricks-Sturrup, MSc, MA, DHSc, FACTS
Robert J. Duke-Margolis Institute for Health Policy, Duke University
1201 Pennsylvania Ave NW
Ste 500
Washington, DC, 20004
United States
Phone: 1 2026212800
Email: rachele.hendricks.sturrup@duke.edu

Abstract

Race, ethnicity, online health information sharing, and AI-mediated environments can intersect in ways that may importantly shape individuals' willingness to share health data with researchers. In this Viewpoint, we closely examine this intersection by drawing parallels between historical and contemporary public health controversies and discussing recent empirical work centered on privacy-related experiences, data donation, and emerging evidence on AI-mediated digital disenfranchisement. We describe how AI-driven online engagement centered on race, ethnicity, and health can facilitate mutual aid and peer support while also posing privacy, reputational, algorithmic, and psychosocial risks when improperly mediated. These effects, alone or combined, can erode trust in digital health research and other engagement initiatives. Therefore, we propose 2 complementary sets of evaluation metrics in response: one focused on how race-related online experiences influence trust, consent behavior, and participation in data sharing and the other focused on race-conscious AI performance within the scope of fairness, representation, and downstream health outcomes. Together, these metrics position race and ethnicity as a key analytic dimension, rather than an optional covariate, to guide teams focused on understanding race as a key analytic dimension in AI-enabled, health-related social engagement online and its full role in and impact on health research.

(JMIR AI 2026;5:e105750) doi: 10.2196/105750

KEYWORDS

race; ethnicity; artificial intelligence; AI; digital health; privacy; psychosocial harm; data sharing; health research

Background

Personal health information sharing in the context of race and ethnicity is broadly observable online today, largely due to 21st-century communication platforms such as social media. In the 18th century, this same phenomenon was also experienced in public brick-and-mortar spaces, as showcased in "A Refutation," a Theater of War Productions dramatic reading of "excerpts from two conflicting historic accounts of Philadelphia's 1793 yellow fever epidemic." [1] This showcase presented conflicting racial community perspectives (ie, perspectives among Black and White community leaders, nurses, caregivers, and first responders) concerning yellow fever susceptibility and duty of care during that time.

Though different in time and context, this story mirrors the COVID-19 pandemic era and likely other modern public health–related controversies around race today. In such cases, context collapse, or a loss of situational boundaries with respect to disease, health management, and race, can spark or maintain spirited debates both online and offline [2-4]. In the 18th century, when, like today, free speech in the United States was protected under the Constitution, the printing press and other analog platforms amplified perspectives, biased or otherwise,

XSL•FO
RenderX

across brick-and-mortar establishments in the interest of both profit and revenue and free speech.

Today, AI technologies embedded across online platforms serve a wide range of purposes that include but are not limited to providing user recommendations, ranking algorithms, and content moderation. In some cases, however, AI technologies can amplify and propagate highly debatable claims and misinformation and rapidly stir debates online as a result. This often translates into higher levels of user engagement (ie, number of views and comments, time spent online, etc) within platforms, particularly social media platforms, and indulgence in often unwise sharing of personal information among platform users. While the intent behind such amplification is to optimize platform engagement harmful bias amplification often becomes an acceptable by-product of platform developers' goals to drive data generation and subsequent monetization of data (including personal data) shared and generated by platform users.

Across these platforms, sensitive personal or racial and ethnic community health information sharing among users (ie, private or public posts about personal, community, or family health situations; photos or videos of health service use and/or health records; etc) is observable. Broad sharing of sensitive information can come with both benefits and risks to individuals, communities, and the entities with which they share personal data. Benefits can include constructive and helpful discourse on race and ethnicity, health management, and disease surveillance, promoting a sense of belonging among platform users.

Risks, however, are major and are algorithmic, privacy-related, reputational, physical, or social in nature. For example, risks include but are not limited to loss of employment or disenfranchisement; harmful effects on immediate and long-term well-being or survival (ie, poor health due to chronic or acute stress or even suicide or suicidal thoughts following content exposure); and rapid changes in emotional regulation or perception that affect personal relationships. Risks may take on many forms across racial and ethnic and other demographic lines among users engaged online. This net effect was central to a lawsuit and a recent landmark decision within a court case against the online social media company Meta and the video streaming company YouTube, wherein the court decided that algorithms structurally embedded within those social media platforms harmed children and public health [5,6]. The original plaintiff in this case, a now 20-year-old woman, filed a case in July 2023 against Meta and YouTube in Los Angeles County Superior Court, claiming that her exposure to the social media platforms' engagement features since age 6 caused her to develop addictive and compulsive behaviors, causing her other physical and psychological harms (anxiety, depression, body dysmorphia, and suicidal ideation) [5,6]. While this case was not directly tied to race, historical and modern-day experiences allow us to reasonably predict that such harms can extend into race-based territory, where AI-enabled or -mediated harm could become racially motivated as either an intended or unintended consequence.

The net effect of such risks outweighing any benefits may affect well-meaning health research initiatives. For instance, individuals approached or invited to engage in such initiatives may have or are even likely to have experienced either real, perceived, or actual harm online and therefore might carry varied levels or degrees of negative psychosocial responses (eg, fear, anger, or withdrawal), generalized distrust, and undesired behavioral changes directed at or away from research settings. This could therefore lead to similarly varied levels or degrees of data use consent, data withdrawal, or withholding (Figure 1).

Considering this, it is critical to acknowledge how context collapse risks associated with sharing health information online across social platforms may interfere with the success of broad or local and well-intentioned digital health interventions and initiatives. Understanding whether and how individuals' health- and race-related interactions online (AI-driven or otherwise, and whether positive or negative) might shape individuals' willingness to test a digital health intervention or engage in a related program with researchers is also crucial.

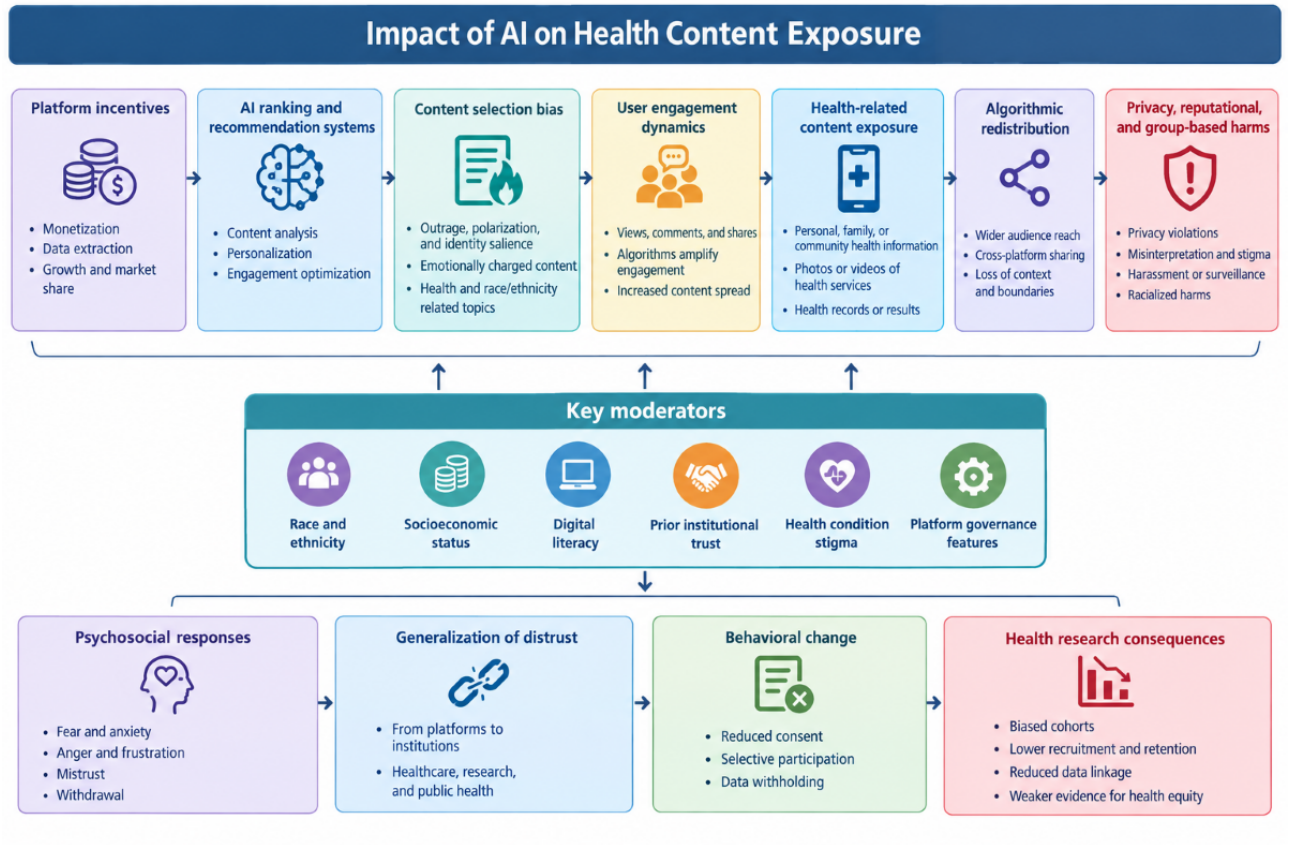
Other risks include what is called "digital disenfranchisement" or unequal access to digital tools and systemic exclusion from fair representation in data, equitable model behavior, meaningful consent, and the benefits of AI-enabled health innovations [7,8]. The ubiquity of hype, influence, and perceived or demonstrable utility of AI technology and AI-driven systems today causes AI to increasingly sit between individuals and institutions. This especially involves institutions that curate, interpret, and act upon individuals' health data across health systems and digital platforms. As the influence, presence, and implementation of AI technology grow, the likelihood of knowledge and experiences around how AI technology can reproduce and amplify race-based discrimination will increase as well. However, the black box or unexplainable nature of AI technology, when and where applicable, will make race-based discrimination claims challenging to prove. This leads to further challenges for patients and communities to legally contest AI-driven, race-based discrimination, despite feeling, experiencing the effects of, or seeing race-based discrimination in real-world settings, including but not limited to online settings [9,10]. Therefore, exploring AI-mediated digital disenfranchisement in broad health data exchange ecosystems is also vital, requiring that race and ethnicity be treated as a central analytic dimension rather than an optional covariate [9].

In this article, we explore and reflect on empirical work focused on privacy-related and AI-mediated experiences, willingness to donate data to researchers amid those environments, and phenomena online and across health data ecosystems, altogether substantiating the need to evaluate and mitigate the risk of unintended harm. With these phenomena in mind, we propose and discuss metrics to begin evaluations of both individuals' race-related experiences online and race-conscious AI performance. We believe our reflections and proposed evaluation metrics can be useful to initiatives and communities seeking to build safe and high-quality digital solutions, enterprises, and ecosystems while accounting for and controlling for real-world race- and ethnicity-related social phenomena. We offer our perspective, informed to date by emerging literature and both historical and modern lessons learned, on mechanisms that can be applied today and moving forward to

XSL•FO
RenderX

support the practice of digital health data sharing in real-world, observational research settings. We believe our perspective could offer inspiration and inform early thinking among programs seeking to test or implement strategies that account for or wish to be sensitive to individuals' potentially varied exposures to AI-mediated, race-related content online. Importantly, we believe our perspective could contribute to broader thoughts exchanged around ways in which researchers could improve engagement with communities participating in health research in our modern era. While our perspective is anchored in past and present phenomena occurring or discussed within the United States, we believe our perspective could also be useful outside of United States–based contexts.

Figure 1. Mediators and moderators across engagement-optimized AI systems embedded in online platforms that may affect health-related information content exposure and disclosure.



Exploring Race-Related Experiences and Individual Willingness to Share Data With Researchers

In our prior work [11-13], we evaluated individuals' willingness to share various sources of real-world data pre- and postmortem with health researchers, as well as their personal privacy-related experiences that have shaped their willingness to share (or not share) real-world data, via an electronic survey. Our results showed that, among 402 (adults aged ≥18 years) survey respondents, around half had their accounts accessed by someone without their permission (54%) and/or experienced a violation of the privacy of their personal information (49%). When adjusted for age range and education level, we also found that participants whose reputations were negatively affected because of information posted online were more likely to share electronic medical record (EHR) data (odds ratio [OR] 2.074, 95% CI 0.986-4.364) and genetic data (OR 2.302, 95% CI 0.894-5.93) than those without this experience [11,12]. Finally, respondents were more likely to donate some data after death (vs no data) if they had longer duration (in years) of using online medical websites (adjusted relative risk ratio=1.22; *P*=.04; 95% CI 1.01-1.48) [13].

Our survey, however, was conducted without the considering race and ethnicity as demographic variables (for reasons discussed in our published study methods) [11-13]. Also, the generally wide CIs observed in cases where participants were more willing to share EHR and genetic data if their reputations were negatively affected because of information posted online indicate a need for more contextual knowledge. This sparks further questions and hypotheses on whether our findings (1) are reproducible in the context of race and ethnicity, (2) provide necessary nuance for race- and ethnicity-focused digital health interventions and initiatives that rely on personal health information sharing, and/or (3) might indicate patterns in human behavior associated with information sharing as a direct response to external social stigma or privacy-related phenomena.

Further evaluation of this topic is needed to build high-quality digital solutions that rely on community health information sharing. Ideally, this should involve repeating our survey among a larger responder cohort to validate our quantitative findings, exploring qualitative context and meaning, and building on or

XSL•FO
RenderX

conducting a scoping review of existing literature. For instance, a scoping literature review should capture the study by Marrie et al [14] who reported, among their North American study cohort (n=4662), that respondents reporting Black race (n=104) showed an unwillingness (OR 0.41, 95% CI 0.20-0.82) to link clinical trial data to administrative health databases. Schooley and Feldman [15,16] reported that race was significantly (though modestly in some cases) associated statistically with willingness to share COVID-19 exposure or positive test information with health care providers (*P*=.06; app notification of positive test); employer or school (*P*=.02; exposure notification); city, county, or state public health dept (*P*=.09; app notification of positive test). A study reported by Chavarria-Soley et al [17] explored preferences among a Costa Rican sample population, finding that most respondents (89%) are willing to donate their information for research purposes, except to for-profit researchers (34% willingness to donate). Relevant hypothetical harms were described to respondents during the survey (eg, "being stigmatized and labeled in some way online" and "ethnic identification and racial discrimination") [17]. Finally, Pratap et al [18] found no statistically significant differences regarding willingness to share data upon comparing racial and ethnic minority groups with nonminority participants in their sample.

Exploring AI-Mediated Digital Disenfranchisement in Health Data Ecosystems

Emerging empirical work highlights that racial- and ethnic-based observations of significant effect are not incidental side effects of health AI but recurrent patterns that emanate from biased data generation processes, model design choices, and deployment practices where harm is an acceptable by-product [9,10]. Systematic reviews show that algorithms used across diverse domains, including risk prediction, imaging, and allocation of services, are more likely to misclassify, undertreat, or deprioritize minoritized populations, even when race is not explicitly included as a predictor [9]. Mbakwe et al [19] argue that fairness assessment in health AI cannot be reduced to generic model performance metrics because the "ground truth" against which models are evaluated is itself often shaped by historical undertreatment and racial tensions in living systems. Similarly, algorithmic fairness scholars have shown that health care algorithms can encode harmful race-based decisions through proxies such as prior expenditure or use, resulting in the decreased allocation of resources to Black patients with comparable underlying health needs [10,20].

AI-mediated disenfranchisement can manifest through several interlocking mechanisms that position race and ethnicity as indispensable to analysis within health data ecosystems [9]. First, data used to train and evaluate health AI frequently underrepresent or misclassify racial and ethnic minority populations, which can lead to models that are poorly calibrated or less accurate for these groups [20,21]. For example, fairness reviews of clinical AI repeatedly find that imaging and diagnostic models achieve lower sensitivity for conditions in patients with darker skin tones due in part to training sets dominated by patients with lighter skin tones and unstandardized race and ethnicity fields [9,20].

Second, even when race and ethnicity are not explicitly encoded as variables, AI models can infer these attributes from correlated signals, such as imaging artifacts, language patterns, or geolocation, using this information to derive predictions in ways that remain opaque to users and clinicians [8,20]. Third, deployment decisions often prioritize efficiency, engagement, or cost reduction over equity. This means that racially biased models are deployed at scale without stratified reporting or ongoing monitoring for disparate impacts [7,8,19]. Together, these dynamics converge to limit racial or ethnic communities' abilities to shape how their data are used, to receive accurate and respectful care, and to benefit from digital health innovations, which is the core of AI-mediated disenfranchisement [9,10].

As a result, race and ethnicity cannot be treated as optional or "nice-to-have" covariates in evaluating the performance of AI-driven or AI-supported systems. They are foundational axes along which disparate performance, harm, and exclusion must be measured and mitigated [8,9]. Mbakwe et al [19] call for routine stratification of AI model performance across patient subgroups, including race and ethnicity, as a minimum requirement for assessing clinical readiness, and argue that failure to do so obscures inequities in how patients are diagnosed, treated, and billed. Translational fairness frameworks emphasize that race-conscious evaluation is necessary to identify and correct algorithmic behaviors that exacerbate poor population health in a similar way, particularly where poorly placed race-conscious efforts have shaped both data and care pathways [8,20]. Systematic reviews of negative AI-driven effects further highlight that many studies either omit race and ethnicity data altogether or report them without subgroup analyses, which limits the field's ability to understand how AI may affect persons of various races differently [9]. In this sense, race and ethnicity are not simply another demographic descriptor; they are central organizing principles in the production and experience of health inequity, and therefore nonnegotiable dimensions for any serious evaluation of AI-mediated digital health [7,9,19].

Proposed Metrics to Evaluate the Effect of Race- and Ethnicity-Related Experiences Online on Individual Willingness to Share Data With Researchers

As mentioned above and discussed using Figure 1, social media and similar online platforms can foster individual connectedness to enable group-level support and awareness of important personal and public health issues. However, when AI-enabled engagement mechanisms are applied with little or no moderation against social context collapse with respect to race and ethnicity (or other demographic factors), provocation and health information sharing risks can enter or become magnified. Clinical and translational research and digital innovation initiatives inherently rely on health information sharing among

XSL•FO
RenderX

people of diverse backgrounds. Therefore, those working within such initiatives should care to evaluate whether or how individuals' health information sharing experiences online engenders downstream negative consequences (eg, privacy, reputational, and/or racialized harms). That is, as negative consequences, race- and ethnicity-related or otherwise, may or may not affect individuals' willingness or trust to engage in data sharing initiatives with health researchers.

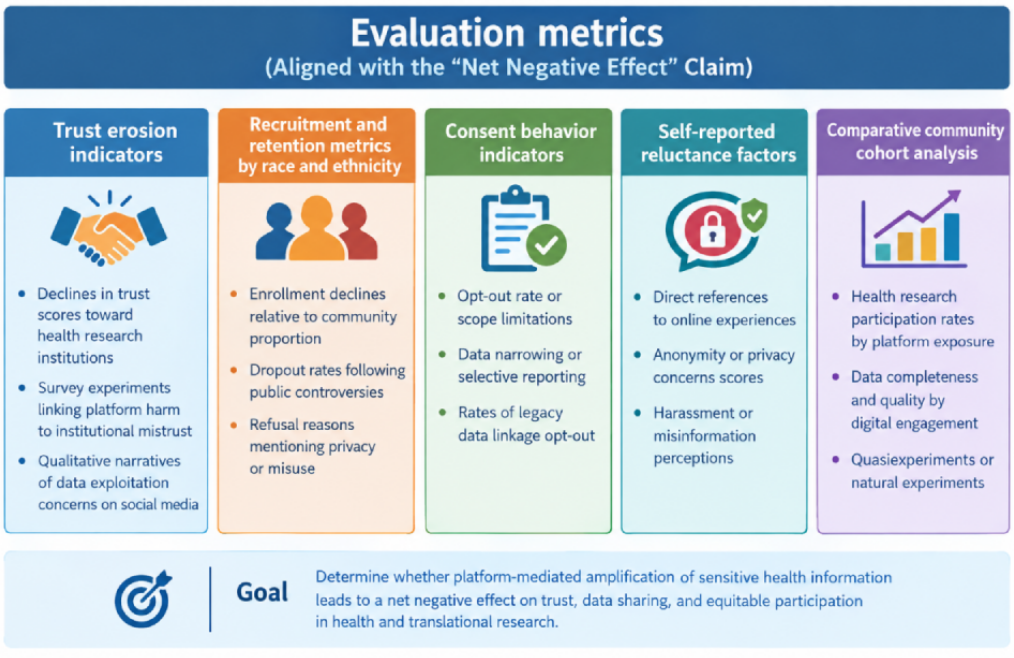
With this in mind, we encourage researchers to empirically assess whether privacy, group, and/or stigma-related harms outweigh any mutual aid, advocacy, or other support or opportunity benefits associated with online health information sharing. Our proposed evaluation metrics to support empirical assessments that consider the effects of self-reported race and ethnicity within the scope of willingness to engage in data sharing with researchers are (Figure 2) as follows:

- Trust erosion indicators—qualifying and quantifying factors that contribute to declines in trust toward research institutions.
- Recruitment and retention metrics by race and ethnicity—measuring engagement (eg, study enrollment, dropout, completion, or refusal).
- Consent behavior indicators—observing and quantifying study opt-in or opt-out rates, narrow or selecting data reporting, and willingness to link data across multiple sources.
- Self-reported reluctance factors—explicitly stated or expressed reasons or rationale for unwillingness to engage in data-driven research following adverse experiences online.
- Comparative community cohort analysis—observing and assessing participation rates among a given cohort, measuring data completeness and response quality across digital data collection tools, and observing outcomes from quasiexperiments or natural experiments.

Above, we described some studies that have examined the role of race and ethnicity and individuals' willingness to share data with researchers. Additional studies, along with a scoping review of published work reported in the literature, on this topic can also support the development and use of these proposed metrics in the context of race and ethnicity. For example, a cross-sectional study used self-administered questionnaires to assess concerns about privacy breaches and misuse of data [22]. Also, a structured survey was used to capture attitudes (eg, perceived discrimination, deception, and exploitation) toward research conducted in emergency situations [23]. Random digital telephone surveys have also been used to assess, using standard measures, several factors, including individuals' knowledge of informed consent procedures, previous participation in research, health care experiences, social support, risk perception, and trust [24]. Finally, phone-administered surveys have been used to conduct mixed methods assessments of patients' (1) baseline attitudes and opinions about research and information privacy and (2) researchers' access to and use of medical records.

In today's AI-driven engagement environment online and in digital settings, we also recommend that clinical and translational research and digital innovation initiatives consider, alongside the proposed evaluation metrics, additional metrics to evaluate race-conscious AI performance. Doing so might offer a clearer picture and thus understanding of the role and influence of race-ethnicity on individuals' behaviors, choices, and motivations to engage in data sharing with researchers.

Figure 2. Proposed evaluation metrics to empirically assess net negative effects resulting from health and race-related interactions online.



XSL•FO
RenderX

Proposed Metrics to Evaluate Race-Conscious AI Performance

Several families of metrics can be used to interrogate how AI models perform with respect to race and ethnicity and digital disenfranchisement, to move beyond descriptive claims toward accountable practice [8,19,20]. Group fairness metrics such as statistical parity, equal opportunity, and equalized odds provide structured ways to quantify whether different racial and ethnic groups receive similar model outputs, conditional on need or risk on a technical level [8,20]. Statistical parity assesses whether the probability of a favorable prediction (eg, being flagged for a beneficial intervention) is similar across groups, while equal opportunity and equalized odds examine whether true positive and false negative rates are comparable across groups with the same underlying condition status [19,20].

Reviews of fairness in medical AI emphasize that these metrics should be computed and reported by race and ethnicity to detect flaws in diagnostic accuracy, triage decisions, or referral patterns that would otherwise be hidden by aggregate performance measures [8,9,19]. In addition, subgroup calibration metrics, comparing predicted risk with observed outcomes separately for each racial and ethnic group, are critical in settings such as risk scoring and resource allocation, where systematic underestimation of risk for one group directly translates into undertreatment [9,10,20].

Individual fairness metrics are also relevant to AI-mediated disenfranchisement beyond group-level statistics. Individual fairness approaches ask whether "similar individuals are treated similarly," which is especially important when race and ethnicity are imperfect or politically charged proxies for underlying social position and risk [8,25]. A scoping review of algorithmic individual fairness in health care finds that these methods can help detect cases where patients with comparable clinical profiles, but different racial identities receive systematically different predictions or recommendations [25]. This kind of divergence may signal that racialized features or proxies are driving the model's behavior in ways that disadvantage certain racial and ethnic groups, even when group-level metrics appear acceptable [19,25]. Combining group and individual fairness metrics offers a more nuanced picture of how AI systems distribute benefits and burdens across racial and ethnic lines [8,20,22-25].

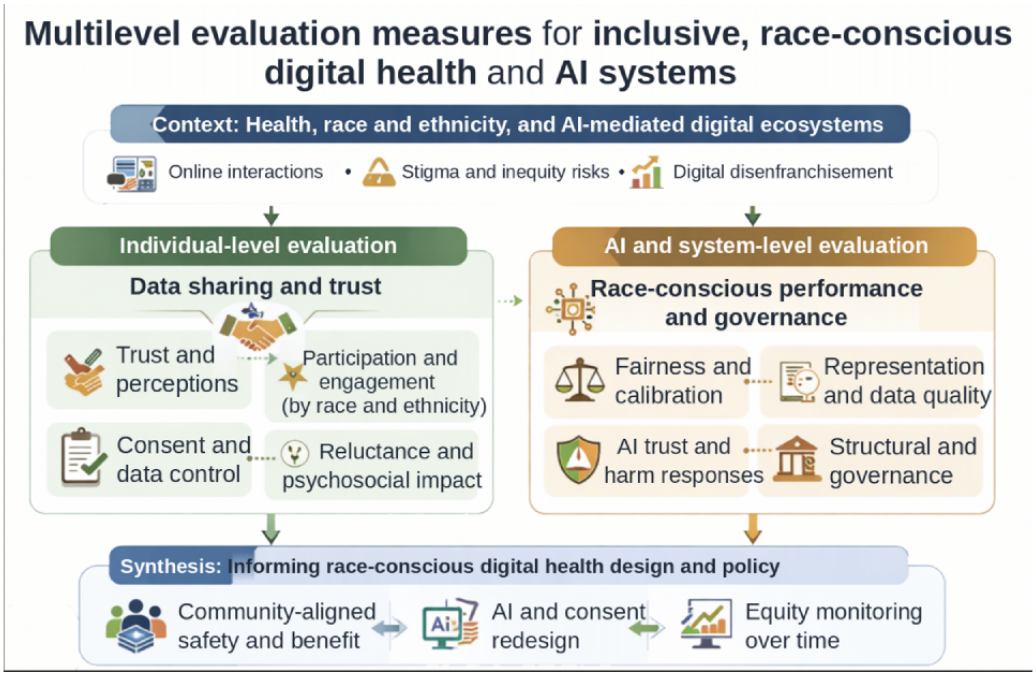
Representation and participation metrics that quantify whose data inform AI models and whose experiences are reflected in evaluation are equally important [9,19]. Fairness reviews stress the need to report the racial and ethnic composition of training and test datasets, as well as the coverage of key intersectional subgroups (eg, Black women with specific chronic conditions), and to compare these distributions with relevant source populations [19,20]. Underrepresentation of racial and ethnic subgroups in training data increases the risk of poor performance and contributes to epistemic exclusion, as communities' illness narratives and care pathways are not fully captured [7,9]. Metrics that track the completeness and accuracy of race and ethnicity fields in EHRs and linked datasets are also crucial, given evidence that misclassification and inconsistent coding undermine both fairness assessment and clinical decision-making [21]. Standardization efforts in race and ethnicity data, combined with transparent reporting of how these fields were collected and used, form a much-needed foundation for race-conscious fairness evaluation [8,21].

Technical fairness measures must be complemented by metrics of trust, consent, and perceived harm stratified by race and ethnicity to capture AI-mediated digital disenfranchisement as it is lived by patients and communities [7,9]. Systematic reviews of AI-driven race-focused observations emphasize that harmfully biased algorithms can erode trust and discourage future engagement with health systems, particularly among poorly resourced communities [9]. Survey-based measures of willingness to share data, consent, and reconsent behaviors, data withdrawal, and refusal to link datasets can reveal patterns of disengagement or strategic withholding that arise in response to repeated harms or lack of transparency around AI [7,8]. Evaluations of health AI that fail to measure these phenomena by race and ethnicity risk missing a vital dimension of disenfranchisement: not only how algorithms behave but how communities respond to them [7,9]. Incorporating psychosocial outcomes, such as perceived stigma, fear, anger, or withdrawal related to AI-mediated health information encounters, can further clarify how technical flaws translate into experiential inequities across racial and ethnic lines [7].

Finally, structural and outcome metrics are necessary to assess whether AI systems ultimately narrow or widen race- or ethnicity-based health outcomes [8,19]. Translational frameworks for clinical fairness call for linking model deployment to changes in concrete outcomes, such as access to care, treatment intensity, or disease control, with analyses that are explicitly stratified by race and ethnicity [8,19]. Recent evaluations of real-world algorithms show that models can exacerbate issues associated with resource allocation even when their overall predictive accuracy is high, which underscores the need to track how AI shifts benefit and burden across racial groups over time [9,10]. Governance metrics like the presence of race-stratified performance documentation, community advisory structures, and formal fairness checkpoints in regulatory review, provide additional signals of whether institutions are treating race and ethnicity as central to AI evaluation [7,8,20]. Together, and combined with a recently proposed comprehensive framework for AI testing, evaluation, verification, and validation for accessibility, these technical, experiential, and structural indicators outline a multidimensional measurement agenda for AI-mediated harm within digitally driven data sharing ecosystems, one in which race and ethnicity are indispensable lenses rather than optional demographic descriptors (Figure 3) [7-9].

XSL•FO
RenderX

Figure 3. Proposed evaluation measures toward the development of inclusive, race-conscious digital health and AI systems.



Conclusions

Digital or online engagement, AI-enabled or otherwise, inevitably involves close, detailed, and unbiased evaluations of its effects on individuals' willingness to participate in data sharing initiatives focused on health. Overall, considering the effects of self-reported race and ethnicity on willingness to share data with researchers, alongside race-stratified documentation of AI performance to capture the effects of AI-driven online engagement in racially motivated contexts, such evaluations are necessary today and will remain so in the foreseeable future. Acting on this necessity means directly supporting initiatives with intentions to build safe and high-quality digital solutions, enterprises, and ecosystems that consider sensitivities associated with real-world race- and ethnicity-related social phenomena.

Here, we provide our perspective as it has been inspired and informed by emerging work to date that should be further scoped in both the literature and in practice given the present rapid pace of AI implementation across health and health research–related settings. We believe our perspective could offer inspiration and inform early thinking among programs seeking to test or implement strategies that account for or wish to be sensitive to individuals' potentially varied exposures to AI-mediated, race-related content online, including but not limited to social media and private online engagement with AI chatbots. Importantly, we believe our perspective could inspire additional, future work exploring how additional demographic and socioeconomic variables or factors intersect or interact with race- and ethnicity-based experiences with the evaluative framing herein in mind.

Funding

The authors declare that no financial support was received for the research and/or publication of this article.

Generative AI Use

The authors declare that generative AI was used in the creation of this manuscript. Generative AI (ChatGPT) was used for visualization or media generation. Generative AI was used to support the creation of Figures 1-3 representing the authors' own ideas, data, and words on the topic. The authors verified all outputs of the tool(s) used in the preparation of this manuscript and take full responsibility for the final content of the manuscript.

Authors' Contributions

Conceptualization: RMH-S, GW
Data curation: RMH-S, GW
Formal analysis: RMH-S, GW
Investigation: RMH-S; GW
Methodology: RMH-S, GW
Project administration: RMH-S, GW
Resources: RMH-S, GW
Validation: RMH-S, GW

XSL•FO
RenderX

Visualization: RMH-S
Supervision: RMH-S
Writing—original draft: RMH-S, GW
Writing—review and editing, RMH-S, GW

Conflicts of Interest

RMH-S is the Lead Editor of the *Journal of Medical Internet Research* at the time of this publication. RMH-S had no involvement in the editorial review and processing of this manuscript.

References

1. A refutation. Theater of War Productions. URL: https://theaterofwar.com/projects/a-refutation [accessed 2026-01-12]
2. Van Bortel T, Lombardo C, Guo L, Solomon S, Martin S, Hughes K, et al. The mental health experiences of ethnic minorities in the UK during the coronavirus pandemic: a qualitative exploration. Front Public Health. Oct 06, 2022;10:875198. [FREE Full text] [doi: 10.3389/fpubh.2022.875198] [Medline: 36276403]
3. Cambodia: fight discrimination amid pandemic. Human Rights Watch. Mar 30, 2020. URL: https://www.hrw.org/news/2020/03/30/cambodia-fight-discrimination-amid-pandemic [accessed 2026-01-12]
4. Marsh J, Deng S, Gan N. Africans in Guangzhou are on edge, after many are left homeless amid rising xenophobia as China fights a second wave of coronavirus. CNN World. Apr 13, 2020. URL: https://www.cnn.com/2020/04/10/china/africans-guangzhou-china-coronavirus-hnk-intl [accessed 2026-01-12]
5. Dyer O. Meta is sued by 41 US states for driving children's addiction to social media and hiding the harms. BMJ. Oct 30, 2023;383:2518. [doi: 10.1136/bmj.p2518] [Medline: 37903513]
6. Kang C, Mac R, Tan E. Meta and YouTube found negligent in landmark social media addiction case. The New York Times. Mar 25, 2026. URL: https://www.nytimes.com/2026/03/25/technology/social-media-trial-verdict.html [accessed 2026-03-31]
7. Chinta SV, Wang Z, Palikhe A, Zhang X, Kashif A, Smith MA, et al. AI-driven healthcare: fairness in AI healthcare: a survey. PLOS Digit Health. May 20, 2025;4(5):e0000864. [FREE Full text] [doi: 10.1371/journal.pdig.0000864] [Medline: 40392801]
8. Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring fairness in machine learning to advance health equity. Ann Intern Med. Dec 4, 2018;169(12):866-872. [doi: 10.7326/m18-1990]
9. Haider SA, Borna S, Gomez-Cabello CA, Pressman SM, Haider CR, Forte AJ. The algorithmic divide: a systematic review on AI-driven racial disparities in healthcare. J Racial Ethn Health Disparities. Feb 2026;13(1):188-217. [doi: 10.1007/s40615-024-02237-0] [Medline: 39695057]
10. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. Oct 25, 2019;366(6464):447-453. [FREE Full text] [doi: 10.1126/science.aax2342] [Medline: 31649194]
11. Hendricks-Sturrup RM, Lu CY. A survey of United States adult privacy perspectives and willingness to share real-world data. J Clin Transl Sci. Feb 01, 2023;7(1):e64. [FREE Full text] [doi: 10.1017/cts.2023.4] [Medline: 37008612]
12. Hendricks-Sturrup RM, Zhang F, Lu CY. A survey of research participants' privacy-related experiences and willingness to share real-world data with researchers. J Pers Med. Nov 17, 2022;12(11):1922. [FREE Full text] [doi: 10.3390/jpm12111922] [Medline: 36422098]
13. Hendricks-Sturrup RM, Lu CY. A survey of individuals' willingness to share real-world data postmortem with researchers. J Clin Trans Sci. Nov 11, 2024;8:1-8. [doi: 10.1017/cts.2024.652]
14. Marrie RA, Sharma M, Cutter GR, Fox RJ, Salter A. Perspectives of people with multiple sclerosis regarding data linkage and sharing. Neurology. May 13, 2025;104(9):e213587. [doi: 10.1212/wnl.0000000000213587]
15. Schooley B, Feldman SS. User perceptions about sharing exposure notification information for communicable diseases. Front Digit Health. Jul 28, 2022;4:926683. [FREE Full text] [doi: 10.3389/fdgth.2022.926683] [Medline: 35966143]
16. Schooley B, Feldman SS. Corrigendum: User perceptions about sharing exposure notification information for communicable diseases. Front Digit Health. 2023;5:1157699. [FREE Full text] [doi: 10.3389/fdgth.2023.1157699] [Medline: 37497186]
17. Chavarria-Soley G, Francis-Cartin F, Jimenez-Gonzalez F, Ávila-Aguirre A, Castro-Gomez MJ, Robarts L, et al. Attitudes of Costa Rican individuals towards donation of personal genetic data for research. Per Med. Mar 2021;18(2):141-152. [FREE Full text] [doi: 10.2217/pme-2020-0113] [Medline: 33576268]
18. Pratap A, Allred R, Duffy J, Rivera D, Lee HS, Renn BN, et al. Contemporary views of research participant willingness to participate and share digital data in biomedical research. JAMA Netw Open. Nov 01, 2019;2(11):e1915717. [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.15717] [Medline: 31747031]
19. Mbakwe AB, Lourentzou I, Celi LA, Wu JT. Fairness metrics for health AI: we have a long way to go. EBioMedicine. Apr 2023;90:104525. [FREE Full text] [doi: 10.1016/j.ebiom.2023.104525] [Medline: 36924621]
20. Chen RJ, Wang JJ, Williamson DF, Chen TY, Lipkova J, Lu MY, et al. Algorithmic fairness in artificial intelligence for medicine and healthcare. Nat Biomed Eng. Jun 2023;7(6):719-742. [FREE Full text] [doi: 10.1038/s41551-023-01056-8] [Medline: 37380750]

XSL•FO
RenderX

21. Tsalidis A, Bharadwaj L, Shen FX. Standardization and accuracy of race and ethnicity data: equity implications for medical AI. PLOS Digit Health. May 29, 2025;4(5):e0000807. [FREE Full text] [doi: 10.1371/journal.pdig.0000807] [Medline: 40440407]
22. Xiao L, He Y, Wu H, Xu X, Gao F, Hu R, et al. Willingness of the Chinese public to participate in medical research during a pandemic: a cross-sectional study. Medicine (Baltimore). Nov 21, 2025;104(47):e45998. [FREE Full text] [doi: 10.1097/MD.0000000000045998] [Medline: 41305789]
23. Smirnoff M, Wilets I, Ragin DF, Adams R, Holohan J, Rhodes R, et al. A paradigm for understanding trust and mistrust in medical research: the Community VOICES study. AJOB Empir Bioeth. 2018;9(1):39-47. [FREE Full text] [doi: 10.1080/23294515.2018.1432718] [Medline: 29368998]
24. Garza MA, Quinn SC, Li Y, Assini-Meytin L, Casper ET, Fryer CS, et al. The influence of race and ethnicity on becoming a human subject: factors associated with participation in research. Contemp Clin Trials Commun. Sep 2017;7:57-63. [FREE Full text] [doi: 10.1016/j.conctc.2017.05.009] [Medline: 29226266]
25. Anderson JW, Visweswaran S. Algorithmic individual fairness and healthcare: a scoping review. JAMIA Open. Dec 30, 2025;8(1):ooae149. [FREE Full text] [doi: 10.1093/jamiaopen/ooae149] [Medline: 39737346]

Abbreviations

EHR: electronic medical record
OR: odds ratio

Edited by G Nicora; submitted 28.Jun.2026; peer-reviewed by M Rozier, D Verran; comments to author 31.Aug.2026; revised version received 15.Oct.2026; accepted 23.Oct.2026; published 10.Nov.2026

Please cite as:
Hendricks-Sturrup RM, Waters G
Understanding Race as a Key Analytic Dimension in AI-Enabled, Health-Related Social Engagement Online and Its Role and Impact in Health Research
JMIR AI 2026;5:e105750
URL: https://ai.jmir.org/2026/1/e105750
doi: 10.2196/105750
PMID:

©Rachele M Hendricks-Sturrup, Gabriella Waters. Originally published in JMIR AI (https://ai.jmir.org), 10.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR AI, is properly cited. The complete bibliographic information, a link to the original publication on https://www.ai.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX