

EQUITY

Bias in clinical algorithms, and the federal rule that reaches the practices using them

ABSTRACT

Bias in deployed clinical algorithms has often entered through what a tool was built to predict, how its inputs were measured and what it learned from, rather than through a race variable. Where race was an explicit input, removing it redistributed error; better measurement reduced it. The Section 1557 rule at 45 C.F.R. § 92.210 requires covered entities, including practices paid under Medicare Part B, to make reasonable efforts to identify tools whose inputs measure a protected characteristic and to mitigate the risk. It survived the 2025 vacatur of other provisions and remains in force, while federal enforcement has turned away from disparate-impact liability. What reasonable efforts can mean for a practice that buys its tools is set out.

Type Evidence and regulatory review **References** 28 **Reading time** 15 min

Last reviewed September 2026

1 A tool the practice did not build

A practice that scores a questionnaire, reads a laboratory-reported estimate or accepts a model's suggestion is relying on someone else's choices about what to measure and what to predict. Two questions follow. The empirical one is whether such tools perform or recommend differently by race, sex, age or disability, and why. The legal one arises from 45 C.F.R. § 92.210, which since July 2024 has barred entities covered by Section 1557 of the Affordable Care Act from discriminating through the use of "patient care decision support tools", and since May 2025 has required reasonable efforts to identify and mitigate the risk.^{1,2}

Many of the most widely cited cases of bias entered through the outcome a tool was built to predict, the instrument that measured its inputs, or the data it learned from, not through a race variable (sections 2 to 4). A federally funded systematic review concluded that algorithms can mitigate, perpetuate or exacerbate racial and ethnic disparities "regardless of the explicit use of race and ethnicity".³ Where race was an explicit input, removing it moved the error rather than eliminating it.

Section 92.210 is in force as of September 2026 and was not among the provisions a federal court vacated in October 2025.⁴ Its identification duty, however, is keyed to a tool's input variables, where only some of the documented cases arose, and federal enforcement toward effects-based discrimination changed in 2025 and 2026.

17.7% → 46.5%

Black share of patients auto-selected for extra help, as built and with a cost-trained score's bias remedied

44%

of hospitals using predictive models evaluated them locally for bias

175

samples per group needed, in one illustrative power calculation, to detect 20% vs 50% outcome prevalence

2 The label, not the variable

The best-known case concerns a commercial algorithm used by health systems to select patients for care-management programs. Obermeyer and colleagues showed that at a given risk score Black patients were considerably sicker than White patients, and that remedying the disparity would raise the fraction of Black patients among those automatically identified for the program, at its 97th-percentile threshold, from 17.7% to 46.5%.⁵ The bias arose because the algorithm predicted health care costs rather than illness, and unequal access to care meant less was spent on Black patients. By some measures of predictive accuracy, cost looked like an effective proxy for health; the racial bias arose anyway.

Race was not the mechanism. The original authors state that the algorithm specifically excludes race,⁵ and La Cava and colleagues describe it as relying instead on a racially correlated proxy outcome.⁶ The original authors generalize: a convenient, seemingly effective proxy for ground truth can be an important source of algorithmic bias.⁵

Imaging models show the same pattern without any demographic input.⁶ Seyyed-Kalantari and colleagues examined chest radiograph classifiers across three large datasets and one multi-source dataset and found that they consistently and selectively underdiagnosed under-served populations, labeling patients with disease as healthy.⁷ Female patients, patients under 20, Black patients, Hispanic patients and patients with Medicaid insurance had higher underdiagnosis rates, and intersectional groups such as Hispanic female patients fared worse.

Removing race from the inputs does not remove it from what a model can represent. Gichoya and colleagues trained deep learning models to predict self-reported race from medical images, with areas under the curve of 0.91 to 0.99 on radiographs, 0.87 to 0.96 on chest CT and 0.81 on mammography.⁸ Candidate confounders did poorly alone, body-mass index at 0.55 and breast density at 0.61, and detection persisted in corrupted, cropped and noised images, often where clinical experts could not tell.

Siddique and colleagues' review, funded by the Agency for Healthcare Research and Quality, included 63 studies from January 2011 to September 2023: 51 modeling studies, 4 retrospective, 2 prospective, 5 pre-post and 1 randomized trial.³ Algorithms reduced disparities in some settings, the revised kidney allocation system among them; perpetuated or worsened them in others, such as severity-of-illness scores used to allocate critical care resources; and had no significant effect elsewhere. Most of this evidence is modeling, and the authors caution that the results may be highly context-specific.

3 Where race was an input, removing it moved the error

Vyas, Eisenstein and Jones catalogued algorithms and guidelines across specialties that “correct” their outputs by race or ethnicity; the article has no abstract and no figure is taken from it.⁹ Kidney function estimation shows what removal does.

Inker and colleagues developed estimated glomerular filtration rate (eGFR) equations without race and tested them in 12 studies and 4,050 participants, 14.3% of them Black.¹⁰ The existing creatinine equation, using age, sex and race, overestimated measured GFR in Black participants by a median 3.7 mL/min/1.73 m² (95% CI 1.8 to 5.4) and in non-Black participants by 0.5 (95% CI 0.0 to 0.9). Deleting its race coefficient underestimated GFR in Black participants by 7.1 (95% CI 5.9 to 8.8). The new creatinine equation without race underestimated it in Black participants by 3.6 (95% CI 1.8 to 5.5) and overestimated it in non-Black participants by 3.9 (95% CI 3.4 to 4.4). Equations combining creatinine and cystatin C without race were more accurate and narrowed the gap between groups.

Taking race out did not produce an unbiased equation. It redistributed error across groups, and the improvement came from a better measurement. The joint National Kidney Foundation and American Society of Nephrology task force recommended immediate adoption of the refitted creatinine equation without race, and national efforts to make cystatin C testing routine, particularly to confirm eGFR for clinical decisions.¹¹ An American Thoracic Society statement in 2023 likewise recommended replacing race- and ethnicity-specific pulmonary function reference equations with race-neutral average equations.¹²

Mitigation that reaches backward

Transplant allocation shows mitigation after a biased tool has already shaped decisions. In July 2022 the Organ Procurement and Transplantation Network mandated race-neutral eGFR calculations for kidney transplant candidates; the change also required programs to identify waitlisted candidates who would have qualified earlier under them.¹³ At one center, 60.3% of 126 Black patients evaluated qualified for a modification, gaining a median 570 days; within 6 months, 26 of them (34%) had been transplanted, with no significant difference in graft loss or mortality. It is a single-center retrospective study, but it shows that correcting a tool can require finding the patients the old version affected.

4 When the instrument is the problem

Some bias sits in the measurement a clinician reads rather than in any model. Shi and colleagues’ meta-analysis of pulse oximetry included 32 studies and 6,505 participants.¹⁴ Pulse oximetry probably overestimates oxygen saturation in people with high levels of skin pigmentation, pooled mean bias 1.11% (95% CI 0.29 to 1.93), and in people described as Black or African American, 1.52% (95% CI 0.95 to 2.09), on moderate- and low-certainty evidence.

The finding is narrower than the way it is usually repeated. The review judged mean bias small or negligible in every subgroup and imprecision unacceptably large, with pooled standard deviations above 1%; taken jointly, measurements in all subgroups met an accuracy root-mean-square below 4%.¹⁴ This review does not show that pulse oximeters fail on dark skin. It shows a small average bias in the direction that makes a low saturation read higher, which the authors judge may be small in hospital settings and is unknown in community settings.

FDA's response is still a draft. It announced draft guidance on pulse oximeter performance testing on January 7, 2025, based in part on concerns that skin pigmentation can affect accuracy, after advisory panel meetings in November 2022 and February 2024.¹⁵ As of September 2026 FDA's guidance page still labels it a draft not for implementation, and no notice of a final version was located in the Federal Register.

The reference standard for pain

Pierson and colleagues trained a model on knee radiographs to predict the pain patients reported, rather than the severity radiologists graded.¹⁶ Radiologist-graded severity accounted for 9% (95% CI 3% to 16%) of racial disparities in pain; the algorithmic measure accounted for 43%, 4.7 times more (95% CI 3.2 to 11.8), with similar results for lower-income and less-educated patients. The authors note that severity measures influence decisions such as arthroplasty. Here the shortfall lay in the human grading standard, which the authors found captured less of underserved patients' pain, and a grading scale used to support clinical decisions reads as a "method" within the rule's definition of a decision support tool.¹⁷

General-purpose language models reproduce the older errors. Omiye and colleagues put 9 questions to 4 commercial models, 5 times each, and found race-based medicine in responses from every model; asked how to calculate eGFR, both ChatGPT-3.5 and GPT-4 produced runs promoting the use of race.¹⁸

5 The pain clinic's own instrument

Opioid risk screening is where the rule meets pain practice most directly. The Opioid Risk Tool was validated in 185 consecutive new patients at one pain clinic, followed for aberrant behaviors for 12 months.¹⁹ Its items cover personal and family history of substance abuse, age, history of preadolescent sexual abuse and certain psychological diseases, and it reported c statistics of 0.82 for its male and 0.85 for its female prognostic models.

Separate male and female models and an age item mean the tool employs input variables that measure sex and age, the trigger in § 92.210(b) on its face, whether it is scored on paper or in software.^{2,17} The rule does not say that using sex or age is discrimination. It says the covered entity must identify the use and make reasonable efforts to mitigate the risk.

A sex-specific item can come out without loss, at least in one sample. In Cheatle and colleagues’ development study, the original tool discriminated between patients with and without opioid use disorder (OR 1.624; 95% CI 1.539 to 1.715); a weighted version eliminating the gender-specific preadolescent sexual abuse item performed comparably (OR 1.648); and a revised unweighted version without it, the ORT-ODU, performed notably better (OR 3.085; 95% CI 2.725 to 3.493).²⁰ That is a derivation result, not an external validation.

Klimas and colleagues’ systematic review found that assessment tools combining patient characteristics and risk factors were not useful for identifying risk of prescription opioid addiction; only the absence of a mood disorder appeared useful for identifying lower risk.²¹ A tool with weak discrimination overall offers little basis for judging its fairness between groups.

WHAT THIS SECTION DOES NOT RECOMMEND

This is not clinical guidance. The paper takes no position on whether or how any opioid risk instrument should be used; it reports how the instruments are built, what their studies found, and how the rule’s trigger applies to them.

TABLE 1

Four routes by which bias entered documented tools, and whether the input-variable trigger in § 92.210(b) would flag the tool. The last column is this paper’s reading of the regulatory text, not an agency determination.

Route	Documented example	Flagged by § 92.210(b)?
Explicit protected input	Race-coefficient eGFR ¹⁰ ; race-specific pulmonary function equations ¹² ; Opioid Risk Tool, sex-specific models and age ¹⁹	Yes
Label chosen as a proxy	Care-management score trained on cost ^{5,6}	No, as described: race excluded, bias through the label
Measurement or reference standard	Pulse oximetry by skin pigmentation ¹⁴ ; radiologist grading of knee osteoarthritis ¹⁶	No protected input
Learned from the data	Chest radiograph classifiers ⁷ ; race recoverable from images ⁸	No protected input ⁶

6 What § 92.210 requires

The instrument is the HHS final rule *Nondiscrimination in Health Programs and Activities*, 89 Fed. Reg. 37522–37703 (May 6, 2024), effective July 5, 2024.¹ Paragraph (a) of § 92.210 states that “A covered entity must not discriminate on the basis of race, color, national origin, sex, age, or disability in its health

programs or activities through the use of patient care decision support tools.” Paragraph (b) imposes an “ongoing duty to make reasonable efforts to identify uses” of tools “that employ input variables or factors that measure race, color, national origin, sex, age, or disability.” Paragraph (c) requires, for each tool so identified, reasonable efforts “to mitigate the risk of discrimination resulting from the tool’s use.”²

A patient care decision support tool is “any automated or non-automated tool, mechanism, method, technology, or combination thereof used by a covered entity to support clinical decision-making in its health programs or activities.” A covered entity includes any recipient of federal financial assistance. Paragraph (a) applied from July 5, 2024; paragraphs (b) and (c) carried a compliance date 300 days later, May 1, 2025.¹⁷

Who is covered

HHS announced with the rule that, for the first time, it would treat Medicare Part B payments as federal financial assistance for the civil rights laws it enforces, so that providers and suppliers receiving Part B funds are prohibited from discriminating on these bases.^{1,22} On that reading, a physician practice billing Part B is a covered entity. The preamble applies § 92.210 regardless of size, while stating that “the size and resources of the covered entity will factor into the reasonableness of their mitigation efforts and their compliance with § 92.210.”¹

What the preamble says reasonable efforts involve

The regulation does not define “reasonable efforts”. For identification, the preamble lists factors OCR may weigh: size and resources, since “a large hospital with an IT department and a health equity officer would likely be expected to make greater efforts to identify tools than a smaller provider without such resources”; whether the tool was used as its developer intended or was adapted; whether the developer disclosed the potential for discrimination or the entity learned that the inputs include a protected characteristic; and whether the entity has a method for evaluating tools, such as asking the developer, reviewing the literature, drawing on medical associations or analyzing complaints.¹ For mitigation it repeats examples from the proposed rule: written policies and governance for how tools are used, monitoring impacts and handling complaints, and training staff.

The input-variable trigger was deliberate. OCR wrote that it “recognizes the challenges in identifying the discriminatory potential of every use of each patient care decision support tool, and therefore § 92.210(b) requires covered entities to make reasonable efforts to identify tools that employ input variables based on a protected basis.”¹ La Cava and colleagues note the consequence: ongoing attention is required only for tools using protected attributes as inputs, while many widely cited examples of bias involve tools that do not.⁶ The general prohibition in paragraph (a) still reaches every tool; the duties to look and to mitigate do not.

The scope stops at clinical decisions. Where unrelated to clinical decision-making affecting patient care, § 92.210 does not apply to administrative, billing, coding, scheduling, supply-chain or staffing tools, among others.¹ Part 92’s general prohibition on discrimination in a health program is not confined to clinical

tools.¹⁷

The preamble also counted on the certification program, expecting the ONC decision support interventions criterion, 45 C.F.R. § 170.315(b)(11), to “work in tandem with § 92.210” by letting a covered entity learn from a developer whether an intervention relies on attributes that measure a protected characteristic.¹ The HTI-5 proposed rule of December 29, 2025 would reduce the scope of that criterion “to fully remove the artificial intelligence (AI) ‘model card’ requirements”; comments closed February 27, 2026, and no final rule had appeared in the Federal Register as of late September 2026.²³

TABLE 2

What § 92.210 requires of a covered entity, and what it does not. Status as of September 2026.

Provision	Requires	Does not	Date
§ 92.210(a) ²	No discrimination on six bases through the use of any patient care decision support tool	State an intent or effects standard in its text	July 5, 2024
§ 92.210(b) ²	Ongoing reasonable efforts to identify tools whose inputs measure a protected characteristic	Require identifying tools whose bias runs through labels, measurement or training data	May 1, 2025
§ 92.210(c) ²	Reasonable efforts to mitigate the risk for each identified tool	Require removing the tool or the variable, or set any test, metric or threshold	May 1, 2025

7 Status in September 2026

The litigation over the rule concerned gender identity, not algorithms. In *Tennessee v. Kennedy*, No. 1:24-cv-161-LG-BWR, the U.S. District Court for the Southern District of Mississippi preliminarily enjoined the gender-identity provisions on July 3, 2024 and on October 22, 2025 entered final judgment vacating §§ 92.101(a)(2)(iv), 92.206(b)(1)–(4), 92.207(b)(3)–(5), 92.8(b)(1), 92.10(a)(1)(i) and 92.208, with related CMS provisions, to the extent they expand Title IX’s definition of sex discrimination to include gender identity.⁴ HHS’s notice of the vacatur, published June 2, 2026, states that the rule’s other provisions remain in force. Section 92.210 is not among those vacated.

OCR said it would continue to enforce the rule’s protections against discrimination based on race, color, national origin, age, disability and the aspects of sex discrimination the order did not affect.²⁴ The electronic Code of Federal Regulations, current to September 24, 2026, shows § 92.210 as published in 2024.² A search of Federal Register documents affecting part 92 found no proposed rule published since the 2024 final rule.

The enforcement environment moved

Executive Order 14281, signed April 23, 2025, declared a policy “to eliminate the use of disparate-impact liability in all contexts to the maximum degree possible” and directed agencies to deprioritize enforcement of statutes and regulations to the extent they include disparate-impact liability.²⁵ On July 24,

2026 HHS removed the disparate-impact provisions of its Title VI regulations, including 45 C.F.R. § 80.3(b)(2), in a final rule effective on publication and issued without prior public comment, stating that Title VI prohibits only intentional discrimination.²⁶

The interaction with Section 1557 is unresolved in the documents. Section 92.101(b)(1)(i) requires recipients to comply with the specific prohibitions in HHS's Title VI regulations at part 80,¹⁷ yet the July 2026 rule does not mention Section 1557, part 92 or § 92.210.²⁶ Section 92.210(a) states no intent requirement. The mechanisms in sections 2 to 4, a neutral label or instrument producing a racial effect, are in structure cases of effect without intent, and no document located says how OCR would treat a § 92.210 complaint of that kind. This review located no published OCR enforcement action or resolution agreement citing § 92.210.

WHAT THIS PAPER WILL NOT CLAIM

It does not claim that § 92.210 reaches, or does not reach, disparate impact; no court or agency document located decides that. An executive order directs agency enforcement and does not amend § 92.210, and enforcement priorities are not repeal. Nothing here is legal advice. Each status statement is dated to September 2026 and should be rechecked.

8 Reasonable efforts for a practice that buys its tools

The preamble's factors describe inquiry more than statistics,¹ which suits a small practice, because the statistical route is largely closed to it. La Cava and colleagues illustrate the arithmetic with 198,823 emergency department admissions at one academic medical center from 2011 to 2019, grouped by sex, race, age and language preference, the last a proxy for national origin, which yields 270 groupings.⁶ Detecting an outcome prevalence of 20% against 50% needs at least 175 samples per group, available in every single-attribute group but only about 65% of four-way intersections; detecting 40% against 50% is possible in only 20% of them. An independent practice sees a fraction of that volume, so local testing mostly cannot separate bias from noise, and evidence of fairness would have to come largely from the developer or external validation.

Hospitals are not far ahead. Using the 2023 American Hospital Association Annual Survey Information Technology Supplement, Nong and colleagues found that 65% of US hospitals used predictive models, 79% of those from their EHR developer; 61% of users evaluated models locally for accuracy but only 44% for bias.²⁷

Properties any system would need

1. **An inventory that includes paper.** The definition covers non-automated methods, so a questionnaire, a grading scale or a laboratory-reported equation belongs on the list with software.¹⁷
2. **Inputs, labels and measurements recorded for each tool.** The inputs answer the question § 92.210(b) asks;² what the tool predicts and which instrument supplies its inputs answer the questions the evidence says matter more.^{5,14}
3. **Developer disclosures kept.** The preamble treats developer information as a factor and relied on certified source attributes to supply it, a supply now proposed for removal.^{1,23}
4. **A recorded mitigation for each identified tool.** Written policy, staff training and a complaint route are the preamble's own examples, and each should leave evidence that it happened.¹
5. **The ability to look back.** When a tool changes, as eGFR did, the patients whose decisions rested on the old version have to be found,¹³ which requires knowing which version of which tool produced each value in the record.
6. **Monitoring stated with its limits.** Subgroup outcome monitoring is worth doing where volume allows and should say where it does not.⁶ An expert panel convened by AHRQ and the National Institute on Minority Health and Health Disparities adds transparency, explicit fairness trade-offs and accountability for outcomes.²⁸

9 What survives

Five statements are supported. Many of the most widely cited cases of bias in clinical algorithms arose through the label, the measurement or the training data, and "race was not an input" does not establish its absence.^{3,5,7,8} Where race was an input, removing it redistributed error and a better measurement reduced it.^{10,11} Section 92.210 is in force, survived the 2025 vacatur, reaches practices paid under Medicare Part B on HHS's 2024 interpretation, and has required identification and mitigation since May 1, 2025.^{2,4,17,22} Its trigger catches tools with sex or age inputs, including a commonly used opioid risk tool, and misses a cost-trained score whose bias ran through its label.^{6,19,20} And federal policy on effects-based discrimination has moved without saying what that means for § 92.210.^{25,26}

Three things would settle what remains open: a statement from OCR on how it reads § 92.210 after the July 2026 Title VI rule; a final decision on HTI-5's model-card provisions; and outcome studies of mitigation in practice, since most evidence that algorithms widen or narrow disparities is still modeling.³ Until then, the rule's text and the evidence together point to knowing which tools are in use, what they take in, what they predict and how their inputs are measured, and to recording what was done about each.

REFERENCES

Entries 5, 7, 8, 10, 14, 16 and 27 are the empirical studies carrying the argument, read against the systematic review at entry 3; entries 1, 2, 4, 15, 17 and 22 to 26 are regulations, agency notices, press releases and an executive order, cited for what they require or announce, with entry 15 a draft guidance and entry 23 a proposed rule. Entry 6 is a legal-policy analysis. Entry 9 published no abstract and no figure is taken from it; entry 13 is a single-center retrospective study, from which the 2022 transplant-network policy is also described, and entry 20 is a development study comparing patients who did and did not develop opioid use disorder, not an external validation.

1. U.S. Department of Health and Human Services, Office for Civil Rights, and Centers for Medicare & Medicaid Services. *Nondiscrimination in Health Programs and Activities*. Final rule. 89 Fed. Reg. 37522–37703 (May 6, 2024); Doc. No. 2024-08711; RIN 0945-AA17; effective July 5, 2024. [federalregister.gov](https://www.federalregister.gov) REGULATION
2. U.S. Department of Health and Human Services. *Nondiscrimination in the use of patient care decision support tools*. 45 C.F.R. § 92.210; 89 Fed. Reg. 37692 (May 6, 2024); eCFR current to Sept. 24, 2026. [ecfr.gov](https://www.ecfr.gov) REGULATION
3. Siddique SM, Tipton K, Leas B, et al. *The Impact of Health Care Algorithms on Racial and Ethnic Disparities*. *Annals of Internal Medicine*. 2024;177(4):484–496. [doi:10.7326/M23-2960](https://doi.org/10.7326/M23-2960) SYSTEMATIC REVIEW
4. U.S. Department of Health and Human Services, Office for Civil Rights, and Centers for Medicare & Medicaid Services. *Notice of Vacatur Regarding Certain Provisions of the 2024 Nondiscrimination in Health Programs and Activities Final Rule*. 91 Fed. Reg. 32887–32888 (June 2, 2026); Doc. No. 2026-11015; reporting *Tennessee v. Kennedy*, No. 1:24-cv-161-LG-BWR (S.D. Miss. Oct. 22, 2025). [federalregister.gov](https://www.federalregister.gov) REGULATION
5. Obermeyer Z, Powers B, Vogeli C, et al. *Dissecting racial bias in an algorithm used to manage the health of populations*. *Science*. 2019;366(6464):447–453. [doi:10.1126/science.aax2342](https://doi.org/10.1126/science.aax2342) RETROSPECTIVE COHORT
6. La Cava WG, Cohen IG, Aysola J. *The future of algorithmic nondiscrimination compliance in the affordable care act*. *npj Digital Medicine*. 9:49; published online December 10, 2025. [doi:10.1038/s41746-025-02224-7](https://doi.org/10.1038/s41746-025-02224-7) POLICY ANALYSIS
7. Seyyed-Kalantari L, Zhang H, McDermott MBA, et al. *Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations*. *Nature Medicine*. 2021;27(12):2176–2182. [doi:10.1038/s41591-021-01595-0](https://doi.org/10.1038/s41591-021-01595-0) BENCHMARK

8. Gichoya JW, Banerjee I, Bhimireddy AR, et al. *AI recognition of patient race in medical imaging: a modelling study*. The Lancet Digital Health. 2022;4(6):e406–e414. [doi:10.1016/S2589-7500\(22\)00063-2](https://doi.org/10.1016/S2589-7500(22)00063-2) BENCHMARK
9. Vyas DA, Eisenstein LG, Jones DS. *Hidden in Plain Sight — Reconsidering the Use of Race Correction in Clinical Algorithms*. New England Journal of Medicine. 2020;383(9):874–882. [doi:10.1056/NEJMms2004740](https://doi.org/10.1056/NEJMms2004740) REVIEW
10. Inker LA, Eneanya ND, Coresh J, et al. *New Creatinine- and Cystatin C–Based Equations to Estimate GFR without Race*. New England Journal of Medicine. 2021;385(19):1737–1749. [doi:10.1056/NEJMoa2102953](https://doi.org/10.1056/NEJMoa2102953) CROSS-SECTIONAL
11. Delgado C, Baweja M, Crews DC, et al. *A Unifying Approach for GFR Estimation: Recommendations of the NKF-ASN Task Force on Reassessing the Inclusion of Race in Diagnosing Kidney Disease*. Journal of the American Society of Nephrology. 2021;32(12):2994–3015. [doi:10.1681/ASN.2021070988](https://doi.org/10.1681/ASN.2021070988) GUIDELINE
12. Bhakta NR, Bime C, Kaminsky DA, et al. *Race and Ethnicity in Pulmonary Function Test Interpretation: An Official American Thoracic Society Statement*. American Journal of Respiratory and Critical Care Medicine. 2023;207(8):978–995. [doi:10.1164/rccm.202302-0310ST](https://doi.org/10.1164/rccm.202302-0310ST) POSITION PAPER
13. Ebadinejad A, Cobar JP, Cyr-Long PL, et al. *Appraisal of Impact of Race-Neutral Estimated Glomerular Filtration Rate Waiting Time Modification on Transplant Wait Time and Outcomes for Black Kidney Candidates: Importance of Transplant Readiness on the Waitlist*. Journal of the American College of Surgeons. 2025;240(6):859–866. [doi:10.1097/XCS.0000000000001349](https://doi.org/10.1097/XCS.0000000000001349) RETROSPECTIVE COHORT
14. Shi C, Goodall M, Dumville J, et al. *The accuracy of pulse oximetry in measuring oxygen saturation by levels of skin pigmentation: a systematic review and meta-analysis*. BMC Medicine. 2022;20:267. [doi:10.1186/s12916-022-02452-8](https://doi.org/10.1186/s12916-022-02452-8) META-ANALYSIS
15. U.S. Food and Drug Administration. *Pulse Oximeters for Medical Purposes—Non-Clinical and Clinical Performance Testing, Labeling, and Premarket Submission Recommendations; Draft Guidance for Industry and Food and Drug Administration Staff; Availability*. 90 Fed. Reg. 1150 (Jan. 7, 2025); Doc. No. 2024-31540; Docket No. FDA-2023-N-4976. [federalregister.gov](https://www.federalregister.gov); draft status: [fda.gov](https://www.fda.gov) GUIDANCE
16. Pierson E, Cutler DM, Leskovec J, et al. *An algorithmic approach to reducing unexplained pain disparities in underserved populations*. Nature Medicine. 2021;27(1):136–140. [doi:10.1038/s41591-020-01192-7](https://doi.org/10.1038/s41591-020-01192-7) COHORT

17. U.S. Department of Health and Human Services. *Nondiscrimination in Health Programs or Activities*. 45 C.F.R. part 92, §§ 92.1 (purpose and effective date), 92.4 (definitions) and 92.101; 89 Fed. Reg. 37692 (May 6, 2024); eCFR current to Sept. 25, 2026. [ecfr.gov](https://www.ecfr.gov) REGULATION
18. Omiye JA, Lester JC, Spichak S, et al. *Large language models propagate race-based medicine*. npj Digital Medicine. 2023;6:195. doi:10.1038/s41746-023-00939-z BENCHMARK
19. Webster LR, Webster RM. *Predicting Aberrant Behaviors in Opioid-Treated Patients: Preliminary Validation of the Opioid Risk Tool*. Pain Medicine. 2005;6(6):432–442. doi:10.1111/j.1526-4637.2005.00072.x COHORT
20. Cheattle MD, Compton PA, Dhingra L, et al. *Development of the Revised Opioid Risk Tool to Predict Opioid Use Disorder in Patients with Chronic Nonmalignant Pain*. The Journal of Pain. 2019;20(7):842–851. doi:10.1016/j.jpain.2019.01.011 CASE-CONTROL
21. Klimas J, Gorfinkel L, Fairbairn N, et al. *Strategies to Identify Patient Risks of Prescription Opioid Addiction When Initiating Opioids for Pain*. JAMA Network Open. 2019;2(5):e193365. doi:10.1001/jamanetworkopen.2019.3365 SYSTEMATIC REVIEW
22. U.S. Department of Health and Human Services. *HHS Issues New Rule to Strengthen Nondiscrimination Protections and Advance Civil Rights in Health Care*. Press release, April 26, 2024. [hhs.gov](https://www.hhs.gov) GOVERNMENT REPORT
23. Assistant Secretary for Technology Policy / Office of the National Coordinator for Health Information Technology. *Health Data, Technology, and Interoperability: ASTP/ONC Deregulatory Actions To Unleash Prosperity*. Proposed rule. 90 Fed. Reg. 60970 (Dec. 29, 2025); Doc. No. 2025-23896; RIN 0955-AA09; comments closed Feb. 27, 2026. [federalregister.gov](https://www.federalregister.gov) PROPOSED RULE
24. U.S. Department of Health and Human Services. *HHS Informs Covered Entities of Partial Vacatur of 2024 ACA Nondiscrimination Final Rule; Core Protections Remain in Effect post Tennessee v. Kennedy*. Press release, June 1, 2026. [hhs.gov](https://www.hhs.gov) GOVERNMENT REPORT
25. Executive Office of the President. *Restoring Equality of Opportunity and Meritocracy*. Executive Order 14281 of April 23, 2025. 90 Fed. Reg. 17537–17539 (Apr. 28, 2025); Doc. No. 2025-07378. [federalregister.gov](https://www.federalregister.gov) REGULATION
26. U.S. Department of Health and Human Services, Office of the Secretary. *Rescinding Portions of the U.S. Department of Health and Human Services Title VI Regulations To Align With the Statutory Text and Conform to Executive Order 14281*. Final rule. 91 Fed. Reg. 46746–46756 (July 24, 2026); Doc. No. 2026-15000; RIN 0945-AA29; effective July 24, 2026. [federalregister.gov](https://www.federalregister.gov) REGULATION

27. Nong P, Adler-Milstein J, Apathy NC, et al. *Current Use And Evaluation Of Artificial Intelligence And Predictive Models In US Hospitals*. Health Affairs. 2025;44(1):90–98. [doi:10.1377/hlthaff.2024.00842](https://doi.org/10.1377/hlthaff.2024.00842) NATIONAL SURVEY
28. Chin MH, Afsar-Manesh N, Bierman AS, et al. *Guiding Principles to Address the Impact of Algorithm Bias on Racial and Ethnic Disparities in Health and Health Care*. JAMA Network Open. 2023;6(12):e2345050. [doi:10.1001/jamanetworkopen.2023.45050](https://doi.org/10.1001/jamanetworkopen.2023.45050) POSITION PAPER

RELATED PAPERS

Regulating AI in certified EHR technology

The (b)(11) source attributes the 1557 preamble relied on, now proposed for removal

Evaluating clinical language models

Benchmarks that never ask the questions where race-based answers appear

After go-live: dataset shift and monitoring

What watching a deployed model over time actually requires