# DEI in Science and Medicine: Missing Metrics and Measures

ROGER B. COHEN AND AMY L. WAX

The process for developing and approving new drugs and treatments for cancer is rigorous. As overseen by the federal Food and Drug Administration, it involves multiple rounds of testing and systematic and precise analysis of patient outcomes, including comparisons to how patients fare with standard therapies. A key part of clinical testing is to define the metrics and measures that determine success or failure. Either a new therapy effectively treats the cancer, improves symptoms and signs, or prolongs life, or it doesn't. Those effects must be demonstrated by substantial amounts of reliable data. And, of course, assessments of new drugs also include careful monitoring for side effects, which can be life-shortening in themselves or otherwise intolerable and can defeat any positive effects. Most experimental drugs for cancer, often hailed with great fanfare and with millions of dollars invested, fail the tests. For those that do pass and are approved for licensing and marketing, data collection and monitoring continue to ensure that the initial claims of benefits are accurate and that the drug is performing as expected in the real world and outside of the clinical trials process.

The clinical trial process in medicine is meticulous and unforgiving. Nothing is taken for granted, and all assumptions are designed to be falsifiable. Impressionistic observations and wishful thoughts and feelings don't count. The supporting data must be detailed, statistically valid, and demonstrably reliable.

This article is written by a specialist in cancer clinical trials (Roger Cohen) and a law professor (Amy Wax, his wife). In this essay, the authors argue that the gold standard for clinical trials should apply to all interventions and changes in health-care fields that are planned, proposed, or implemented. In particular, the changes proposed for medicine in the name of so-called "diversity, equity, and inclusion" should be subject to rigorous evaluation and scrutiny. Proposed DEI initiatives range widely and affect all aspects of the medical system, from how new doctors and other personnel are selected and trained to the design and selection of treatment methods and protocols and to research design and funding. The injection of DEI principles into medicine warrants no less scrutiny than any other intervention in the health-care field. Before implementing any proposed program or modification of existing health-care practices, key questions must be considered and defined. What are the goals? What are the metrics, measures, and "deliverables"? What precise improvements or benefits do we hope to generate, and how will they be demonstrated and monitored? Once the intervention has occurred, what plans are in place to assess whether the goals were accomplished and whether health outcomes are improved or degraded? What kind of data will be gathered, and how? Are the methods for analyzing the data rigorous and unbiased? Are there plans to replicate any initially promising results before widespread acceptance and implementation occur?

Currently, these questions are rarely being asked, pursued, or analyzed in a sustained and systematic way. To the extent that DEI is being evaluated at all, the efforts are scattershot and fall well below any scientifically defensible standard. In other words, DEI initiatives are being implemented in medicine on every front with little sustained effort to subject the DEI-inspired changes to the gold standard of scrutiny and assessment that prevails in every other area of modern medicine. Examples abound. Last June, for instance, a document was released to the community by the leadership of the main teaching hospital of the Perelman School of Medicine at the University of Pennsylvania, one of the best medical schools in the country. The document described how Penn is signing up with enthusiasm for a new Joint Commission health equity certification. The Joint Commission inspects hospitals once every three years to determine whether they meet defined standards of sound medical practice. The additional health equity accreditation by the Joint Commission, a new part of the certification process, is still technically voluntary. But it is understood that declining to seek such certification would not only be unwise, but downright unthinkable. A hospital's ability to operate legally and bill for its services rests on private accreditation by the Joint Commission. Proving "equitable" bona fides—which means adopting "diversity, equity, and inclusion" principles, priorities, and practices in all hospital activities—is therefore effectively mandatory. No hospital, even the most prestigious and venerable, can afford to refuse.

The Joint Commission has gone full woke: "*This new, voluntary advanced certification for [a] health care equity program provides the structure to guide your organization's journey to achieving health care equity. It guides forward movement in imbedding health care equity in all aspects of care, treatment, and service delivery.*" The Joint Commission's ukase on DEI does not even try to explain or justify the new equity certification in the traditional terms of better health outcomes for patients, which is the Joint Commission's traditional ambit. Will "health care equity" make the hospital safer and more efficient, or the patients better off? Will fewer patients die or develop complications if this new certification occurs? Will it improve the services that hospitals exist to deliver? Radio silence on these issues.

We are now about fifteen years into the grand medical DEI experiment. Its grip on the field only grows stronger. The infiltration of DEI into every aspect of medicine, starting with undergraduate (medical school) and graduate (residency and specialty fellowships) admissions, education, and training, and now extending into hospital management, hiring and personnel management, human resource policies, medical and treatment resource allocation, and medical research priorities and funding, rests on a series of unproven and often dubious assertions. Each step into the DEI abyss either rests on no credible evidence at all or on a small number of oft-cited but questionable and poorly designed studies that rely on crude and unexamined assumptions. These studies have never been subjected to the painstaking and searching scrutiny that has long been the standard at the top of the American medical establishment. This wholly inadequate evidence, which would never be tolerated in any other scientific field, is used to justify sweeping changes in every aspect of our profession. Validation, replication, control variables and groups, critical analysis and re-analysis, the application of well-established statistical standards, and diligent vigilance against the elementary error of conflating correlation with causation are conspicuously missing from the literature that is driving these changes. Vague jargon and buzzwords (the ubiquitous parlance of "diversity, equity, and inclusion," "systemic racism," "structural racism," and "social determinants of health") are routinely deployed without ever being precisely defined. These egregious methodological flaws, and specifically the absence of well-defined outcome metrics and their systematic, careful measurement, are the focus of this essay.

The shortcomings described here are grievous and ought to concern everyone and not just people in the field. Modern medicine, which has matured over centuries to rest on a sound scientific foundation, is based on constant, unsparing examination and scrutiny. Like any biological or physical science, the search for truth and the process that leads to progress requires constant testing and debate, often leading to the outright rejection of treatments and approaches that don't work or fail to meet well-defined metrics and expectations. Nothing like this is happening today in the newly hatched field of "health equity." Two centuries of methodological and scientific progress in medicine have been jettisoned, or at best ignored, in the name of enshrining the DEI woke transformation of a proudly precise and disciplined field.

The return to rigor and standards should start with medical education. One of us has written previously about the dramatic shifts that have recently occurred in how we select and train future doctors and has urged that we approach such reforms with great caution.

Deciding to place a major and sometimes primary emphasis in medical admissions and training on social justice and diversity, equity, and inclusion, rather than on candidates *demonstrated and measured* (not, as DEI acolytes would have it, so-called latent) ability, is nothing short of reckless.

Before making such changes, we should, at the very least, subject departures from longstanding practices to objective and unbiased analysis of their effectiveness. Without such an assessment, we imperil the overarching goal of achieving the very best outcomes for our patients based on the highest quality medical care. And we also jeopardize future advances in basic and medical science that depend critically on the quality of the next generation of scientists. Future therapeutic advances will necessarily be sacrificed if diversity initiatives take priority over finding, admitting, and training the best and the brightest people who have demonstrated the intellectual attributes and abilities for scientific achievement at the highest levels. Once students are selected, what they are taught is also crucial. Medical school and physician training is brief. More focus on social justice and health equity means less time and attention to the technical, demanding, and vastly time-consuming efforts to acquire medical knowledge, including learning scientific methods and techniques and performing scientific research. Advocates for curricular changes and a reorientation of medical training to prioritize social justice and equity rather than the scientific and medical fundamentals rarely (never, really) claim that this radical reform project will promote scientific excellence and sustain the scientific innovation that has long characterized American medicine. Rather, they speak vaguely of a "new excellence," which is never actually defined.

Currently, empirical support for the goal of "*embedding health care equity in all aspects of care, treatment, and service delivery*" is, when properly examined, practically nonexistent. Apart from a few studies, repeatedly cited but easily faulted, the main justifications for DEI in health care are rhetorical and based on politically popular, feel-good assertions that are far from proven but that people in the medical field are very reluctant to challenge. It is worth pointing out that the diversity in DEI has nothing to do with diversity of thought or political point of view, including any ideas that challenge DEI paradigms and the basic assumptions behind a social justice focus. The diversities valued by the AAMC (American Association of Medical Colleges), as customarily presented in public diversity talks, are socioeconomic status, race, ethnicity, language, nationality, sex, gender identity, sexual orientation, religion, geography, disability, and age. Conspicuously missing are the categories of diversity of ideas, thought, political position, and point of view that would encourage, if not demand, that DEI initiatives stand on more solid ground before they are allowed to transform long-standing admissions, training, and treatment practices. In fact, the lack of proper skepticism towards the unscientific, irrational, empirically ungrounded nature of the changes presently proposed and implemented under the banner of DEI is demonstrated by the attitude towards nonbelievers. In most health-care institutions and medical schools, even the most prestigious, questioning the tenets of DEI and the "bias narrative" at its heart or positing any alternative explanation other than racism or prejudice for minority group ills is a form of punishable heresy. The mere existence of such heresy indicates the need to double down. It is taken as proof positive that even more DEI is needed.

What kinds of assertions and evidence are currently adduced to justify the DEI-based transformation of health-care practices?

> Numerous articles and presentations in the medical literature assert: "Diversity, equity, and inclusion (DEI) are essential principles that physicians must embrace to enhance rapport with patients, increase patient compliance with treatments, improve patient outcomes, increase patient satisfaction, and build trust."

This set of claims is one variation of an oft-stated assertion: DEI efforts produce better health outcomes, and especially for minority patients. How exactly do DEI efforts produce this result? The focus is on increasing the number of underrepresented and minority physicians and health-care workers. Patient outcomes will improve if their doctor is the same race, it is claimed, because patients are more likely to adhere to and trust treatment recommendations, and minority providers will be more effective and astute caretakers of their minority patients. The so-called "racial concordance" thesis, which posits positive effects from doctors and patients sharing a common group identity, has also been extended selectively to other categories, such as gay people—but not to whites or South Asians. The bald and sweeping assertion in the statement quoted above is typical of the genre. But it is backed up by little or no solid evidence. What evidence should we be looking for? The physicist Wolfgang Pauli famously stated that a hypothesis cannot be established as true unless and until an experiment can be devised to prove it wrong. It must be "falsifiable." If it is impossible to conduct a test to disprove a claim, then that claim is not one that belongs to science but rather to the realm of ideology, fantasy, dogma, and wishful thinking. The racial concordance fad reveals the current state of the practice of DEI in health care. No studies have been proposed, or carried out, in any attempt to disprove the hypothesis.

Instead, this idea continues to exert a pervasive influence despite weak and limited evidence to support it. The so-called "Oakland study" (Alsan et al. 2019), purports to show that black patients with black doctors have better health outcomes. Despite being repeatedly cited, this study is riddled with basic methodological flaws that effectively render it useless. These include the lack of an adequate control group, the projection of facially preposterous lifetime mortality benefits from a single point-in-time observation, and conclusions based only on patients' expressed willingness to engage in preventative care without any quantitative evidence of increased use of preventative services or of actual health benefits, from whatever source, in the short- or long-term.

Another widely touted and reported study claims that newborn black babies are more likely to survive if they are cared for by black doctors (Greenwood et al. 2020, 117 (35)). The single study that is the basis for this claim is also fatally defective. For one thing, it uses an administrative database of the general racial composition of caretakers in the studied health-care facilities as a substitute for determining the actual race of the doctors caring for individual babies—a methodology that is entirely inadequate and unreliable by any standard of sound medical social science. It also fails to control for the birthweight and medical condition of newborns as an important determinant of the doctors who end up caring for them. (Borjas and VerBruggen 2024).

These shoddy articles are repeatedly and uncritically cited to justify dramatic overhauls in policies and practices in health care delivery and training. What is remarkable is how the two studies just described have achieved such hallowed status, with eight hundred-plus citations in the literature for the Alsan et al. study and approximately three hundred for the Greenwood et al. study. The few academic critiques that do exist are generally ignored. Among the DEI bureaucrats, proponents, and so-called "experts" who now abound in medical schools and other medical establishments, attempts to replicate, disprove, or even systematically critique such "research" almost never occur. Instead, presentations of the "concordance" idea are routinely accompanied by vague references to "many studies" or "studies show" without specific citations. Because DEI initiatives in health care are motivated overwhelmingly by political priorities rather than actual evidence, there is no reason to believe that research that comes to the "wrong result" will ever get reported or published. "Another widely cited DEI pillar is the assertion that 'teams with diverse perspectives achieve higher scientific and economic impact.'"

In 2015, 2018, and 2020, the venerable consulting firm McKinsey released three reports purporting to show that greater racial and gender diversity in large public companies' executive ranks results in higher profits. Reliance on these studies has become a staple of DEI presentations, in which speakers regularly draw a straight causal line from more personnel diversity to higher profits. The problem here, of course, is that correlation is not causation. Maybe the causal arrow, in fact, runs the other way, with more profitable firms choosing to prioritize hiring a more diverse workforce or being better able to afford to do so.

There are other reasons to doubt the bona fides of the McKinsey study claims. A paper published earlier this year by economists Jeremiah Green and John Hand (*Econ Journal Watch* 2024) tried to replicate the McKinsey data. They failed. They found no link between racial and ethnic diversity and the financial performance of S&P 500 firms. To be sure, there are differences between the Green and Hand study and the original McKinsey reports, including the types of firms examined and their location, which could explain the discrepant results. Unfortunately, McKinsey's refusal to disclose the raw data behind their study makes a more detailed comparison impossible. And there are additional metanalytic and quasi-meta-analytic papers in the literature that find insignificant, negative, or mixed effects on productivity and profits based on workforce age, gender, and cultural background.

There are also other sources on business performance that could be analyzed to test the McKinsey reports' conclusions. For instance, at least two prominent mutual funds that invest in "diverse" businesses have been created: iShares Refinitiv Inclusion and Diversity UCITS ETF and SPDR® MSCI USA Gender Diversity ETF. The performance of these funds can be readily examined; they grossly underperform their peers. The experience with these so-called "diverse" investment funds is in tension with the conclusions of the McKinsey studies. These facts should be front and center in any discussion of the relationship of personnel and leadership diversity to firm profits. Yet they are not.

The story for scientific team performance is just as weak. There are summary analyses in the literature, referred to above, that have shown that the link between team diversity and aspects of science team performance is decidedly mixed and sometimes negative. But the literature on this topic—which is admittedly sparse—is short on specific metrics and measures of scientific success, quality, and contributions. Some possible parameters to examine might include publications, patents, citations, H-indices (which reflect how often papers are cited), intradepartmental collaborations, interdepartmental collaborations, and number of invitations and oral presentations at major meetings. The literature does not fully explore the available data.

One metric of outcomes that is currently useless, unfortunately, is the number of NIH or other prestigious government grants awarded to an investigator or team. That is because affirmative action is now rampant in federal science funding. The NIH FIRST Grants Program is a typical and lavishly financed example of this phenomenon. Scathing exposés of questionable practices in this program can be found on several sites[1]. The stated goal of NIH FIRST and programs like it is to achieve more "diversity" in science, with an emphasis on race and skin color. In describing and justifying this objective, the relationship of diversity to desirable outcomes and "success" is both assumed and touted in vague, jargon-laden terms that have nothing to do with well-established, concrete measures of the outcomes that should matter: greater human health and longer life. An emblematic statement can be found in the 2023–2027 NIH-Wide Strategic Plan for DEIA: *"To achieve institutional and research excellence, NIH must foster and sustain an inclusive and equitable culture that embraces DEIA, both in the workplace and in the pursuit of biomedical and behavioral science. The true measure of success for cultural change is belonging—the feeling and knowledge of being included in the NIH mission."* Another passage in the 2024 NIH government grants guide[2] states: *"Research shows [no references given] that diverse teams working together and capitalizing on innovative ideas and distinct perspectives outperform homogeneous teams. Scientists and trainees from diverse backgrounds and life experiences bring different perspectives, creativity, and individual enterprise to address complex scientific problems."* How those "different perspectives" result in specific health improvements is neither addressed nor elaborated. Once again, the standards of basic scientific rigor that inform other areas of medical science are conveniently forgotten.

DEI principles have now become a scientifically unproven priority for government science funding. This distorts which scientists and projects are supported. But DEI considerations also now dictate how the research is performed! One egregious example can be found in research on HPV-related head and neck cancer. In recent years, there has been a pronounced uptick in this disease in the USA, Europe, and Australia. This potentially fatal type of head and neck cancer is not an equal opportunity affliction. Whites are affected almost ten times more often than blacks, and 80 percent of patients are men. Epidemiology studies have established that the incidence patterns are in large part related to sexual behaviors that differ among groups, including the self-reported performance of male-on-female oral sex.

For a recent clinical study in HPV-related head and neck cancer conducted at the University of Pennsylvania, the National Institutes of Health (NIH) program officer, reviewing the study's progress, asked the investigators why there were no blacks in the study. In response, the investigators pointed out that blacks seldom get this disease and that this well-known fact had been noted and discussed in the original grant proposal. That answer did not satisfy the DEI commissar: "That answer is non-responsive and unacceptable, find and enroll them or we will stop funding your grant." How the cessation in funding would help in finding better treatments for this type of head and neck cancer was not addressed.

To illustrate the absurdity and destructiveness of this reflexive bureaucratic reaction, consider an analogy to sickle cell anemia. Yes, some whites are afflicted with sickle cell disease, but they are few. The ratio of blacks to whites is around fifty to one. A demand that all research on sickle cell disease involve some "adequate number" of white patients is not "equitable." It is irrational. And such a demand would mean that important lifesaving research on sickle cell anemia, which mostly affects blacks, would be seriously impeded.

*A core DEI claim is that adopting DEI principles and practices is the only way to effectively address health disparities.*

Obamacare dramatically improved patient access to health care after 2011. More people acquired personal physicians and visited their doctors, including minorities who were previously underserved. But better access has not improved racial and other group disparities in health-care outcomes. As one national DEI proponent, Dr. Consuelo H. Wilkins from Vanderbilt University Medical Center, recently pointed out, *"[t]he absence of progress [on the impacts of racism and inequities in health care] compared with the survey from nearly 3.5 years ago is striking.... Both the extent of the disparities and the perception of how patients are impacted is virtually unchanged. This is despite all the discussions we have had and the new programs and increased funding for equity initiatives."* See NEJM Catalyst in 20243. One possible explanation for these disappointing results might be that the diagnosis and prescription are wrong. DEI proponents routinely point to "structural" or "systemic" racism as the main cause of intractable group disparities, and especially for poor health outcomes for blacks. The proposed solution is "anti-racism," a concept that is never precisely defined. Alternative explanations of health disparities (including choices and behaviors of the patients themselves) are rarely proposed, let alone investigated. Indeed, suggestions of any personal responsibility for inferior health outcomes are quickly shut down under the angry rubric of victim-blaming. And the suggested solutions are always the same: more resources, money, programs, and "accountability." That is obviously an inadequate and dogma-driven approach to health disparities. Perhaps, for example, weak and unstable families and a relative lack of family support might contribute to the failure to see a doctor, comply with treatment, or make lifestyle adjustments conducive to positive health outcomes. At the very least, the claim that DEI in health care might mitigate or remedy the effects of the lack of familial comfort and support should be studied and verified. Specific interventions to compensate for family failure should be proposed, and their effects scrutinized and measured. The data must be made available for evaluation. None of this is currently happening. Such tunnel vision would never be tolerated in any other area of medicine. The routine failure to consider alternative hypotheses and causal factors, even glaringly obvious ones, would be regarded as scientific and empirical dereliction.

> DEI proponents are ardent supporters of funds for disparities research, especially between racial groups. What is the evidence that health disparities research has resulted in better health outcomes?

The increase since 1990 in government funding for research on so-called "health disparities" is staggering. As of 2020, a third of all National Science Foundation (NSF) grant awards included one of the following terms: "equity," "diversity," "inclusion," "gender," "marginalize," "underrepresented," or "disparity," up from 3 percent in 1990 (data summarized by the Center for Partisanship and Ideology 2021). Young investigators not surprisingly follow the money, so the number of grant requests to support disparities research has exploded. In the same vein, applications for prestigious residency or fellowship training programs in medicine, even from MD/PhD candidates being trained to do basic research, increasingly mention and highlight the applicant's interest in disparities research. Since 2015, the National Institutes for Health (NIH) has required the inclusion of "score driving" (NIH's own minatory words) diversity priorities for all projects seeking funding within its seventy-two National Cancer Centers, including research on preclinical models of basic biological processes in organisms such as worms and zebrafish! The money diverted into disparities research, or used to contort studies where diversity priorities have no place and add nothing, is money that cannot be used for other projects. Resources are scarce, so disparities research and priorities should at least be scrutinized, tested and evaluated, and called upon to demonstrate efficacy. Nothing like that is happening now. Indeed, the standards for disparities research are embarrassingly low and would not pass muster in any other area of medical investigation. Most disparities research is highly descriptive, and the vocabulary is vague, ideologically informed, and stereotypical. The terms structural racism, systemic racism, and anti-racism provide little guidance on how to address exceedingly complex, multifactorial societal problems, let alone scientific ones.

> DEI supporters assert that diversity in medicine benefits everyone, not just racial and ethnic minorities. The claim is that white physicians in racially diverse medical schools are more culturally responsive and report feeling more comfortable treating diverse patient populations.

Even if, as often claimed, physicians and patients trained in DEI "feel more comfortable around diverse people," there is effectively no evidence that these feelings lead to better health outcomes. It should be obvious that feeling better is not the same as making people better. Cultural sensitivity and cultural humility are different from competence; they may or may not go together. The goal of creating more diverse medical schools and promoting a "sense of community" has generated a cottage industry of trainers, handlers, and consultants and an endless plethora of training sessions, workshops, and presentations on microaggressions, implicit bias, unconscious discrimination, and similar topics. But there is no reliable evidence that these activities help us get along better, let alone that they improve health care or medical results. In fact, some troubling evidence has emerged that these training courses may have negative effects, promoting cynicism and resentment rather than insight and cooperation.

One especially favored training tool is the implicit association (IAT) test, also called the implicit bias test, which claims to help people discover potential prejudices that lurk beneath their awareness—and correct them! The flaws of this test are well established. It does not even measure what its aficionados say it does (racism), is not replicable (based on test-retest data), and its well-respected academic inventors have publicly decried its routine use in institutional anti-racism training. Yet, it has become the centerpiece of a profitable industry. Health-care institutions such as Penn Medicine now routinely subject their employees to regular IAT training and testing and boast of exceptionally high compliance rates.

Fifteen years in and with increasing recognition of its pernicious effects on institutions and scientific excellence, the time has come to recognize that ideologically driven DEI initiatives have no place in institutions of higher learning, especially in medical science, unless they are validated by hard evidence and data. Metrics and outcome measures for all diversity initiatives must be defined, and the results of all interventions must be precisely measured and demonstrated. Some universities have as many as one hundred staff in the DEI office. In business parlance, there must be "accountability" and "deliverables." Medical science is adept at measuring—that is its stock-in-trade. Its tried-and-true practices should be extended to investigating the value of DEI.

There are some simple, specific steps that could be taken. The weak studies underpinning many sweeping diversity initiatives need to be sunsetted, starting with the Oakland adults and Florida newborns studies. Neither article is worthy of respect even under the basic standards of social science. In science, mediocre and flawed papers get replaced by better papers. (See, for example, the most recent PNAS study quoted earlier on black infant mortality.) Older treatment paradigms in medicine are regularly abandoned in favor of better treatments. Drugs are retired or have their FDA approval rescinded. There is nothing wrong with this. Quite the contrary—it is essential to medical quality and progress. Without new data and disruptive thinking, we would still be bloodletting. These insights should be applied to DEI, and as soon as possible.