International Journal of Nursing Studies Advances 9 (2025) 100380

Contents lists available at ScienceDirect

International Journal of Nursing Studies Advances

journal homepage: www.sciencedirect.com/journal/international-journal-of-nursing-studies-advances

# Reconsidering the use of race, sex, and age in clinical algorithms to address bias in practice: A discussion paper

Reanna Panagides [a,*], Jessica Keim-Malpass [b]

[a] University of Virginia, School of Nursing, Charlottesville, VA, USA
[b] University of Virginia, School of Medicine, Charlottesville VA, USA

ARTICLE INFO

*Keywords:*
Clinical decision-making
Bias
Resource allocation
Algorithms
Social Justice
Literacy

ABSTRACT

Clinical algorithms are commonly used as decision-support tools, incorporating patient-specific characteristics to predict health outcomes. Risk calculators are clinical algorithms particularly suited for resource allocation based on risk estimation. Although these calculators typically use physiologic data in estimation, they frequently include demographic variables such as race, sex, and age as well. In recent years, the inclusion of race as an input variable has been scrutinized for being reductive, serving as a poor proxy for biological differences, and contributing to the inequitable distribution of services. Little attention has been given to other demographic features, such as sex and age, and their potential to produce similar consequences. By applying a framework for understanding sources of harm throughout the machine learning life cycle and presenting case studies, this paper aims to examine sources of potential harms (i.e. representational and allocative harm) associated with including sex and age in clinical decision-making algorithms, particularly risk calculators. In doing so, this paper demonstrates how systematic discrimination, reductive measurement practices, and observed differences in risk estimation between demographic groups contribute to representational and allocative harm caused by including sex and age in clinical algorithms used for resource distribution. This paper ultimately, urges clinicians to scrutinize the practice of including reductive demographic features (i.e. race, binary-coded sex, and chronological age) as proxies for underlying biological mechanisms in their risk estimations as it violates the bioethical principles of justice and nonmaleficence. Practicing clinicians, including nurses, must have an underlying model literacy to address potential biases introduced in algorithm development, validation, and clinical practice.

**What is already known**

- Reductive definitions of race contribute to representational harms
- Allocation of healthcare services based on race is unethical and illegal

**What this paper adds**

* Corresponding author at: Claude Moore Nursing Education Building, 225 Jeanette Lancaster Way, Charlottesville, VA 22903.
*E-mail address:* rkp3em@virginia.edu (R. Panagides).

https://doi.org/10.1016/j.ijnsa.2025.100380
Received 24 January 2025; Received in revised form 1 July 2025; Accepted 6 July 2025
Available online 6 July 2025

- Reductive definitions of age and sex also contribute to representational harms
- Use of chronological-age and binary-coded-sex in risk estimation perpetuates existing inequities
- Allocation of healthcare based on binary-coded sex, or chronological-age is also unethical and illegal

## 1. Background

Risk calculators are a type of clinical algorithm designed to help clinicians provide personalized medicine to patients (Margolis, 1983). They do so by leveraging patient-specific characteristics to predict health outcomes thus influencing treatment pathways (Fakhari et al., 2024; Soleimanpour and Bann, 2022). Clinical algorithms typically consist of input variables and an outcome variable. Input variables may include clinically relevant measurements or values such as vital signs, laboratory values, the presence or absence of a current/past medical condition, and often demographic features such as race, sex, and age. These input variables are then used to predict an outcome variable, such as in-hospital mortality, risk of post-surgical complications, risk of organ transplant failure, cancer survival, etc. Risk calculators, specifically, are commonly used to allocate limited resources such as healthcare services or diagnostic procedures (Goodman et al., 2023b). Optimally, clinical algorithms incorporate concepts of precision medicine to improve clinical decision-making by promoting increased efficiency and care standardization (Visweswaran et al., 2025).

One example of a commonly used clinical algorithm is, The Society of Thoracic Surgeons Short-Term Risk Calculator (O'Brien et al., 2018). This calculator predicts a patient's risk of complication and death post-cardiac surgery. It consists of >60 input variables, which include patient-specific characteristics such as operation type, age, sex, race, etc. By identifying a patient's risk based on these personal characteristics, undesirable outcomes are better anticipated and thereby considered when deciding whether surgery may be appropriate or not for that specific patient (Jain, 2002). For example, patients predicted to have a 'high risk' of mortality or post-surgical complications may not be eligible for cardiac surgery due to this stratification. In theory, the use of risk calculators for healthcare resource allocation has the potential to positively influence both clinicians and patients by not only providing services to those who may benefit from them the most but also avoid unnecessary suffering in those who may not (Goodman et al., 2023b).

When including demographic characteristics (i.e., race), this risk calculator systematically stratifies White patients as having inherently lower risk for poor post-surgical outcomes compared to Black patients. For example, when all other variables are held constant, the risk of post-surgical complications and death increases by 20 % if the patient is classified as Black (Vyas et al., 2020). Black patients are then systematically less likely to be eligible for cardiac surgery simply due to subjective physical characteristics observed by their clinician counterpart. An analysis conducted on data between 2000 and 2005 demonstrates the effect of this practice on clinical outcomes (Cram et al., 2009). Black patients with acute myocardial infarction were approximately 25 % less likely to be treated with revascularization compared to White patients with similar insurance policies (Cram et al., 2009; Phillips-Beck et al., 2020). Including race as a predictor in clinical algorithms in this manner promotes race-based medicine and reinforces existing health disparities (Cerdeña et al., 2020).

In recent years, including racial characteristics in algorithms used in clinical settings has sparked significant ethical debate within the medical community (El-Azab and Nong, 2023; Vyas et al., 2020). Historically in medicine, race has been defined as a set of physical characteristics that represent genetic differences observed across geographic ancestry (O'Brien and Clare, 2023). Given the well-documented issues with race-based medicine within clinical research and medical education, many publications and medical organizations have condemned the use of race as a predictor in clinical algorithms, a variable which reflects a history of colonialism and slavery rather than genetic variation and is greatly impacted by measurement error (El-Azab and Nong, 2023; O'Brien and Clare, 2023; Visweswaran et al., 2025). Eliminating the use of race as a predictor combats race-based medicine by rejecting the flawed notion of race as a biological construct and acknowledging that discrimination (i.e., racism) is a key determinant of illness and health (O'Brien and Clare, 2023). The racism that is historically embedded within race-based variables and its potential to perpetuate future inequities outweighs any benefit of attempting to use race as a proxy for genetic geographic variation. This recent realization has resulted in greater scrutiny of including race in medical algorithms, forcing developers creating these algorithms and the clinicians who use them to consider their ethical implications.

Although there has been a recent focus on racial inequities perpetuated by clinical algorithms, these tools also have the potential to discriminate against other demographic characteristics used as input variables, such as sex and age, which may be similarly problematic. Yet, the inclusion of sex and age within algorithms is rarely challenged (Mohottige et al., 2024). Due to the advancement in human genetic research, it is understood that race is not a reliable proxy for genetic difference. Therefore, the inclusion of race in medical algorithms is contradictory, naively embedded, and often criticized (O'Brien and Clare, 2023). Yet, for other demographic groups (i.e., sex and age), observed clinical differences have biological underpinnings recognized within the medical community, therefore, their inclusion in clinical algorithms is not as readily questioned. Similar to race, it is important to acknowledge that binary-coded-sex and chronological age may also be unreliable proxies for biological differences subject to various forms of historical and measurement bias. All demographic variables (race, sex, and age) represent complex interactions between socio-political groups throughout history intertwined with systematic discrimination and complex power dynamics whether they have a biological-foundation or not. Consequently, if the allocation of healthcare resources and services are based on biased and reductive demographic classifications, these risk calculators can inflict systematic discrimination on a large scale and are ultimately subject to legal repercussions (Goodman et al., 2025). The introduction of bias from various sources, the potential harm it may cause, and the lack of attention to these specific forms of bias underscore the need for a nuanced analysis of commonly included demographic features in clinical algorithms.

### 1.1. Scope of discussion

This paper acknowledges that discrimination (sexism and ageism in addition to racism) is a key determinant of illness and health that may partially account for observed population-level differences reflected in risk calculators. In this discussion paper, we argue that while mitigating racial bias in clinical decision-making algorithms has gained substantial research attention, similar mitigation strategies should also be applied to other demographic features routinely used in algorithms, such as binary-coded sex and chronological age. To understand why these techniques can be extended to sex and age, it is crucial to explore two central questions: (1) What do these variables represent? and (2) What are the potential harms? To answer these questions, we first apply a conceptual framework to identify sources of harm introduced within the dataset generation and algorithm deployment stages of the model building and implementation lifecycle that have the potential to perpetuate inequities. Second, we present four case studies evaluating sources of harm within commonly used clinical algorithms, specifically risk calculators, that incorporate these variables and examine how their inclusion can impact the equitable distribution of healthcare services. Finally, we highlight the bioethical and legal significance of addressing this bias, considering the implications for justice and non-maleficence.

## 2. Identifying sources of algorithmic harm – conceptual framework

Adapted from Suresh and Guttag (2021), the framework in Fig. 1 identifies sources of harm in the machine learning life cycle that can be readily applied to clinical algorithms such as risk calculators (Suresh and Guttag, 2021). Before identifying potential sources of harm embedded within clinical risk calculators, it is first important to understand the typical life cycle of a clinical algorithm. First, data is generated about populations from the world in its current historical context. A population is then sampled, and certain aspects characterizing this sample are measured. This data is then cleaned, processed and formulated into a usable dataset. This dataset, which typically consists of input and outcome variables, is then used to train a clinical algorithm (i.e. a risk calculator). Lastly, this algorithm is used to make predictions about future patients and influence the allocation of healthcare services.

Suresh and Guttag (2021) state there are two types of "harm", or "negative consequences" caused by these algorithms: representational harm and allocative harm. Representational harm occurs "when certain people or groups are stigmatized or stereotyped" while allocative harm occurs "when opportunities or resources are withheld from certain people or groups" (Suresh and Guttag, 2021). Representational harms (caused by historical and measurement bias) and allocative harms (caused by deployment bias) are defined and analysed with regard to sex and age in the following sections. To investigate sources of representational harm we conducted a literature review investigating historical bias and measurement bias that contribute to the misrepresentation of complex biological constructs (i.e. sex and age). To investigate sources of allocative harm, a brief search was performed of the reputable and frequently used clinical algorithm database (mdcalc.com) to identify risk calculators that include demographic features as input variables. Risk calculators included for analysis were intentionally chosen to represent a wide range of conditions across the lifespan. Selected algorithms were specifically designed to calculate risk thus maximizing their potential impact on the allocation of healthcare resources among patients. It is also important to note that many other types of bias lead to harm within the algorithmic life cycle specifically during the model building stage (e.g., aggregation bias, learning bias, evaluation bias), but they do not apply to the demographic variables we are investigating and therefore will not be discussed in this paper (Suresh and Guttag, 2021).

## 3. Historical bias – sexism and ageism

Historical bias is a type of algorithmic error that reflects and perpetuates existing societal inequities, with the potential to cause harm. This type of bias arises even if we can perfectly measure data, such as sex and age, because the environment in which the data exists has inflicted harm on certain populations. Suresh and Guttag (2021) state that "historical bias arises…if the world as it is or was

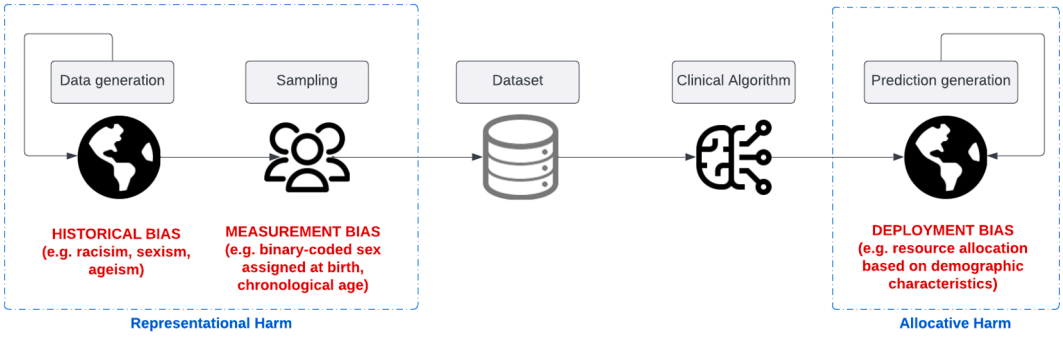


**Fig. 1.** Sources of algorithmic harm

*Note*. Fig. 1 depicts the clinical algorithm life cycle starting with the generation of raw data from the world and ending with the generation of predictions about the world. Historical bias (e.g., racism, sexism, ageism) and measurement bias (e.g. binary-coded sex assigned at birth and chronological age) both can contribute to representational harm. Deployment bias (e.g. the inequitable distribution of healthcare services based on demographic characteristics) may contribute to allocative harm.

leads to a model that produces harmful outcomes." Like patients of different races, people of different sexes and ages also access and receive care differently within the healthcare system, sometimes due to discrimination based on stereotypes and systemic injustices. Homan (2019) outlines our current understanding of how social systems of inequality, rather than true biological differences, influence health: (1) perceived experiences of discrimination or mistreatment, (2) physician bias or discrimination within medical institutions, and (3) structural discrimination that leads to systemic inequality in power and resources(Homan, 2019). It is difficult to distinguish true differences between the biological constructs of sex and age from the influence of historical sexism and ageism. Therefore, the data used to train these clinical algorithms will reflect patterns of racism, sexism, and ageism in our society.

### *3.1. Ageism*

Ageism is defined as discrimination against a person based on age. Although individuals of any age can experience discrimination, in the healthcare system, ageism disproportionately affects older adults. There are largely three ways older adults' health is harmed by discrimination: (1) perception of ageist discrimination, (2) ageism from healthcare professionals in medical institutions, and (3) ableist healthcare structures that disadvantage older adults with limited functioning (i.e., vision, hearing, and mobility limitations). The perception of discrimination based on age can lead older adults to feel less comfortable accessing healthcare services, resulting in delayed care or avoidance of the healthcare system altogether. Additionally, studies have shown that healthcare providers may dismiss the concerns of older adults, attributing symptoms to natural aging rather than exploring treatable conditions such as arthritis or neuropathy(Farrell et al., 2022). This age-based bias can delay diagnoses and deny patients access to necessary treatments and procedures. Lastly, structural systems of ageism, including barriers to care for individuals with declining vision, hearing, or mobility, affect older adults' ability to engage with their own healthcare or adhere to medication regimens. These structural inequities, combined with stereotypes about cognitive decline, influence how older adults make decisions about their care and how providers offer care(Farrell et al., 2022, 2020). These three sources of inequality (perception of discrimination, provider bias, and structural discrimination) affect how patients of different ages interact with the healthcare system, ultimately introducing historical bias into clinical algorithms.

### *3.2. Sexism*

Sexism is defined as discrimination against a person based on sex. Both sexism and gender discrimination have been prevalent throughout history and have impacted how people of different sexes and genders interact with the healthcare system. Although individuals of any gender can experience discrimination, in the healthcare system, sexism disproportionately affects women. There are largely three ways women's health is harmed by discrimination: (1) perception of gendered discrimination or sexual harassment, (2) gender bias from healthcare professionals in medical institutions, and (3) socioeconomic influences of structural sexism and gender discrimination in society (Krieger, 2014). Although a person of any gender can be subject to discrimination, women are more likely to report their symptoms being dismissed or being treated differently because of their gender(Cabral and Dillender, 2021). The perception of discrimination based on gender can lead women to feel less comfortable accessing healthcare services when they need them.

Additionally, studies have shown that women are less likely than men to receive the most advanced diagnostics and treatments available for a variety of health conditions due to physician gender bias(Homan, 2019). Physician gender bias influences what types of treatments women have access to, ultimately influencing their health outcomes as well. Lastly, socioeconomic influences of structural sexism and gender discrimination contribute to poor health outcomes. Structural sexism is defined as the systematic gender inequality in power and resources. Homan (2019) demonstrates how different genders have unequal access to jobs, salaries, and insurance policies, etc. These socioeconomic characteristics influence what treatments patients have access to and who has power over their own healthcare. These 3 social systems of inequality influence observed population-level differences between the sexes that are not attributable to true biological differences but are rather a result of systematic injustice.

## 4. Measurement bias – binary-coded sex and chronological age

Error introduced by historical bias is compounded by the fact that the measurement of sex as a binary variable and chronological age as a continuous/ordinal variable is reductive in nature. Both sex and age are multifaceted biological constructs that are difficult to measure because of their complexity and, therefore, subject to measurement bias. Suresh and Guttag (2021) state that, "measurement bias occurs when choosing, collecting, or computing features and labels to use in a prediction problem. Typically, a feature or label is a proxy (a concrete measurement) chosen to approximate some construct (an idea or concept) that is not directly encoded or observable"(Suresh and Guttag, 2021). Proxies are used when it is not possible, or it is too costly to measure a construct that constitutes many different characteristics. Proxies become problematic when they are a gross oversimplification of a more complex construct(Suresh and Guttag, 2021), resulting in the erasure of nuances within populations. Binary-coded sex and chronological age both serve as oversimplifications of more complex biological constructs.

### *4.1. Binary-coded sex assigned at birth*

In clinical algorithms, sex is typically binary-coded as female or male (Grimstad et al., 2021). This binary representation of sex fails to acknowledge important nuances that may ultimately make binary-coded sex an unreliable proxy. Sex is a multidimensional

biological construct typically defined as the combination of anatomical features, physiology, hormones, and genetics (Bates, 1973). The choice to represent sex as a binary-coded variable not only poorly reflects these characteristics that make up a person's sex but also fails to acknowledge that various combinations of sexual characteristics (i.e. anatomical features, physiology, hormones, etc.) change throughout the lifespan. Surgical, medical, and pharmaceutical histories contribute to diverse combinations of anatomical and physiological sex-based characteristics. When developing these calculators, we must thoroughly scrutinize which specific sexual characteristics are relevant to the outcome variable they attempt to predict rather than rely on naïve assumptions about static sexual characteristics embedded in binary-coded schemas.

Consider the specific sexual characteristics, as defined above, of a post-menopausal woman with a prior hysterectomy. Although this patient would be binary-coded as female, their anatomical characteristics and hormone levels would look much different from a 25-year-old person with an intact uterus. Similarly, consider a male with Kallmann syndrome (Sonne et al., 2025) (a condition by which a genetic mutation causes a deficiency in gonadotropin-releasing hormone, resulting in low sex hormones and the absence of secondary sexual characteristics) who may have undescended testicles and low testosterone levels. Although this patient would be binary-coded as male, their anatomical characteristics and hormone levels would be different from a 25-year-old male without this syndrome. Subsequently, how would these patients' biological characteristics look different if they were on hormone replacement therapies.

Similarly, transgender patients may choose to receive gender-affirming hormonal (Salas-Humara et al., 2019) or surgical therapies (Oles et al., 2022) altering physiological or anatomical characteristics present at birth which may result in various combinations of sexual characteristics difficult to categorize in binary classifications. Binary-classifications of sex contributes to the erasure of gender diverse populations and incorrectly assumes sexual characteristics as being fixed throughout the lifespan. Therefore, it is important to acknowledge that across transgender, cisgender, and intersex patients, sexual characteristics vary by stage of life, surgical history, and use of exogenous hormones (i.e. contraceptives, hormone replacement therapies, gender affirming therapies, anabolic androgenic steroids, etc.)(Alexander et al., 2022; Poteat et al., 2023; Tapper et al., 2019). Most calculators that include binary-coded sex comparing males and females, therefore, fail to acknowledge these sex-based nuances poorly represented in two categories (Poteat et al., 2023).

### *4.2. Chronological age*

Age is typically a ratio-level variable measured in years from birth. Age is a biological phenomenon in which, with time, humans ultimately "succumb to deteriorating function, loss of upkeep and finally ceasing maintenance altogether"(Cho and Na, 2022; Khosla et al., 2022; Kim, 2022). It is a complex interaction between resilience reserve, damage and compensated stress, which ultimately results in frailty and mortality (Ferrucci et al., 2019). Ferrucci et al. (2019) states that "strong correlations have recently been revealed between health dimensions and phenotypes that are typical of aging, especially with autophagy, mitochondrial function, cellular senescence, and DNA methylation." Age measured in years is therefore a poor proxy to capture the complex biological mechanisms that drive the rate of biological aging (Diebel and Rockwood, 2021). Since biological age is impacted by genetic predisposition, multimorbidity and environmental risk, these nuances are not captured in a simple 'year from birth' measurement. Additionally, with how age is currently being measured, it perpetuates age-based discrimination and inequitable allocation of healthcare resources if a person's chronological age is not reflective of their biological age.

Ultimately, binary-coded sex assigned at birth and chronological age are unreliable proxies for multidimensional biological constructs and fail to capture the complexities required for accurate representation. Reductionistic practices of sex and age result in measurement bias and can contribute to representational harm. Since sex and age have the potential to be discriminated against when distributing healthcare services based on historical discrimination and social constructs associated with sex and age-based identities, they should only be included in algorithms if they can be measured accurately to mitigate against unintentional representational harm.

## 5. Deployment bias

This section examines selected clinical algorithms, specifically risk calculators, to demonstrate how including variables such as sex and age can influence the distribution of healthcare services. As we acknowledge the measurement bias inherent in binary-coded sex and chronological age and their potential for harm, it is important to note that these characteristics are frequently used in clinical algorithms. "MDCalc.com" is a frequently used online medical calculator database that publishes risk calculators commonly used by clinicians. For this analysis, MDCalc was searched for risk calculators that include sex and/or age within their models, and these calculators were investigated for potential biases in fairness or inclusivity. Interestingly, MDCalc recently provided a statement on the inclusion of race in medical calculators and risk estimates, warning clinicians about the dangers of algorithmic bias. MDCalc states, "Every segment of our society is taking a hard look at race and racism, and medicine is no exception. Just like all institutions, there are numerous examples of racism in medicine in the past and today, and we must continue to challenge, question, and change them." However, they have yet to comment on the inclusion of sex and age for potential sources of bias that may reinforce sexism and ageism, paralleling the structural racism acknowledged in their statement.

Four risk calculators were selected from MDCalc as case studies for this paper. Although this is not an exhaustive list, the algorithms discussed below serve as examples of how sex and age are included as input variables and the unintentional harm that can result from their inclusion. The selected algorithms represent a wide variety of use cases across the lifespan and are readily used within the general population to show the pervasive nature of these input variables throughout the field of medicine. It is important to note that risk calculators were intentionally selected for this analysis due to their primary role in resource allocation. This gives them the greatest

potential to contribute to the inappropriate distribution of healthcare resources and, in turn, to violate bioethical principles and legal obligations. Refer to Table 1 for additional information about the risk calculators analysed in this section.

## 6. Case studies

### 6.1. Pediatric appendicitis risk calculator

The Pediatric Appendicitis Risk Calculator (Kharbanda et al., 2018) predicts appendicitis risk in pediatric patients presenting with abdominal pain. It uses predictors such as sex, pain duration, pain migration, nausea, and fever. When all other variables are held constant (and are set to be consistent with symptoms of appendicitis), the calculated risk for a female is 46 %. In contrast, for a male, it is 69 %. This difference implies that males may be perceived as having a higher risk of appendicitis solely due to their sex, which can impact clinical decision-making. Consequently, patients categorized as female may be less likely to undergo further diagnostic imaging or surgery because of their estimated risk. The underlying biological mechanisms contributing to observed differences in the pathogenesis of appendicitis across sex have not been thoroughly studied, making it difficult to determine the degree to which these differences are driven by societal versus biological factors (Kollias et al., 2024).

**Table 1**
Examples of risk calculators that include sex and age as predictors.

| Tool | Clinical Utility | Input variables | Equity Concern |
|---|---|---|---|
| Pediatric Appendicitis Risk Calculator (Kharbanda et al., 2018) | Predicts appendicitis risk in pediatric patients with abdominal pain | -Setting (e.g. community, pediatric emergency department)<br>-Sex (e.g. female, male)<br>-Duration of pain, hrs (e.g. <24, 24–48, 48–96 etc.)<br>-White blood cell count<br>-Neutrophil<br>-Presence of pain with walking<br>-Maximal tenderness in right lower quadrant<br>-Abdominal guarding<br>-History of migration of pain to right lower quadrant | When all other variables are held constant (and are set to be consistent with symptoms of appendicitis), there is a 23 % difference in risk between males and females. |
| Kidney Failure Risk Calculator ( Tangri et al., 2011) | Predicts progression to kidney failure in patients with chronic kidney disease | -Estimated glomerular filtration rate<br>-Sex (e.g. female, male)<br>-Urine albumin-to-creatinine ratio<br>-Age, yrs<br>-Serum albumin<br>-Serum phosphorus<br>-Serum bicarbonate<br>-Serum calcium | When all other variables are held constant (and are set to be consistent with chronic kidney disease, there is a 7 % difference in calculated 5-year risk of kidney failure between males and females.<br>When all other variables are held constant (and are set to be consistent with chronic kidney disease, there is a 7.3 % difference in calculated 5-year risk of kidney failure between patients aged 50–59 years and 60–69 years. |
| Atherosclerotic Cardiovascular Disease 2013 Risk Calculator from AHA/ACC (Goff et al., 2014) (without race as a variable) | Predicts 10-year risk of heart disease or stroke to guide management strategies | -Age, yrs<br>-Diabetes<br>-Sex (e.g. female, male)<br>-Smoker<br>-Total Cholesterol<br>-High density lipoprotein<br>-Systolic blood pressure<br>-Treatment for hypertension | When all other variables are held constant (and are consistent with cardiovascular disease), there is a 9.4 % difference in calculated 10-year risk between males and females.<br>When all other variables are held constant (and are consistent with cardiovascular disease), there is a 14.9 % difference in calculated 10-year risk between a patient aged 60 years and a patient aged 70 years. |
| Obesity Surgery Mortality Risk Score (OS-MRS)(DeMaria et al., 2007) | Predicts mortality risk in patients undergoing gastric bypass surgery | -Body mass index ≥50 kg/m^2<br>-Male gender (e.g. no, yes)<br>-Hypertension<br>-Risk of pulmonary embolism<br>-Age ≥45 (e.g. no, yes) | When all other variables are held constant (and are set to factors consistent with increased mortality risk), there is a 5.66 % difference in calculated mortality risk between males and females.<br>For women, when all other variables are held constant (and are set to factors consistent with increased mortality risk), there is a 5.66 % difference in calculated mortality risk between women greater than and <45 years old. |

*Note*. The table above highlights components and equity concerns of four age and sex adjusted clinical algorithms. Differences in estimated risk between age and sex groupings were observed throughout.

### 6.2. Kidney failure risk calculator

The Kidney Failure Risk Calculator (Tangri et al., 2011) predicts the 5-year risk of kidney failure in patients with chronic kidney disease. It uses predictors such as eGFR (estimated glomerular filtration rate), sex, urine albumin-to-creatinine ratio, age, and serum albumin levels, phosphorus, bicarbonate, and calcium. When all other variables are held constant (and are set to be consistent with chronic kidney disease, the calculated 5-year risk of kidney failure for a female is 56.3 %, while a male's risk is calculated to be 63.3 %. This difference suggests that males may receive more aggressive treatments due to their higher calculated risk. Additionally, age also impacts this calculated risk. A person aged 50–59 years has a 70.9 % 5-year risk, while a person aged 60–69 years is calculated to have a 63.6 % risk. Interestingly, this inverse relationship between age and risk implies that younger patients might receive more aggressive treatments, potentially excluding older adults from receiving treatment simply due to their age. The underlying biological mechanisms mediating these observed differences across sex and age have not been thoroughly studied, making it difficult to determine the degree to which these differences are driven by societal versus biological factors as well (Neugarten and Golestaneh, 2019).

### 6.3. Atherosclerotic cardiovascular disease 2013 risk calculator

The Atherosclerotic Cardiovascular Disease Risk Calculator (Goff et al., 2014) predicts a 10-year risk of heart disease or stroke to guide management strategies, including high-intensity statin prescriptions. It uses predictors such as age, diabetes status, sex, smoking status, total cholesterol, high-density lipoproteins, systolic blood pressure, and treatment for hypertension. When all other variables are held constant (and are consistent with cardiovascular disease), the calculated 10-year risk for a female is ~42.3 %, while a male's risk is calculated to be ~51.7 %. This difference suggests that women may be systematically less likely to receive high-intensity statin prescriptions due to their perceived lower risk. Additionally, age also impacts the calculated risk. A person aged 60 years is calculated to have a ~36.9 % risk, while a person aged 70 years is calculated to have a ~51.75 % risk. This difference implies that younger patients are automatically at lower risk because of their age, potentially excluding them from receiving preventative care when all other risk factors indicate they may benefit from a statin prescription.

The biological mechanisms driving observed differences in atherosclerotic cardiovascular disease rates across age and sex have been thoroughly studied. Age-related factors associated with atherosclerosis include changes in immune function, which reduces the body's ability to respond to new antigens and increases the likelihood of autoimmune responses (Poznyak et al., 2023). Similarly, sex-hormones like estrogen and testosterone play a role in the pathogenesis of atherosclerosis contributing to inflammation, plaque morphology, and vascular remodelling (Poznyak et al., 2023). Although observed differences can be partially explained by endogenous sex-hormones' role in atherosclerotic pathogenesis, hormone levels across sexes are influenced by a range of external factors that are inaccurately represented in binary classifications. Sex-hormone levels (which mediate inflammatory biomarkers such as C-reactive protein) are influenced by pharmacologic mechanisms such as oral hormone replacement and gender affirming therapies (Poznyak et al., 2023). Patients on these therapies have been shown to have statistically different levels of sex-hormones compared to those not on these therapies (Harper et al., 2021). Accounting for this hormonal variability through direct feature measurements should be used to support more inclusive and accurate risk estimations. A similar conceptual model can be applied to the use of chronological age as a proxy for biological aging.

### 6.4. Obesity surgery mortality risk score

The Obesity Surgery Mortality Risk Score (DeMaria et al., 2007) predicts mortality risk in gastric bypass surgery patients. It uses predictors such as body mass index ($\geq$50 kg/m$^2$), male sex (non-male or male), hypertension, risk of pulmonary embolism, and age ($\geq$45 years). However, including sex and age as variables introduces potential biases in treatment eligibility, as discussed previously. When all other variables are held constant (and are set to be consistent with increased mortality risk), the calculated risk for a "non-male" is 1.9 %. In comparison, a male patient's risk is calculated to be 7.56 %. Additionally, age influences this calculated risk, particularly in non-male patients. For instance, a "non-male" under 45 years has a calculated risk of 1.9 %, whereas a non-male above 45 years has a calculated risk of 7.56 %. A male who is $<$45 years old has a risk of 7.56 %, while a male who is greater than 45 years of age has a risk of 7.56 %. This suggests there is an intersection between age and sex. This intersection between age and sex suggests that male patients may be systematically less likely to receive gastric bypass surgery based on their higher calculated risk of mortality simply due to their sex. At the same time, older females might also face limited access due to increased perceived risk compared to younger females. Lastly, the underlying biological mechanisms contributing to observed differences in successful recovery from gastric bypass surgery across age and sex have not been thoroughly studied, making it difficult to determine the degree to which these differences are driven by societal versus biological factors (Benotti et al., 2014).

## 7. Bioethical and legal implications

Clinicians are ethically and legally obligated to abide by bioethical principles and antidiscrimination laws when providing patient care (Goodman et al., 2025; Varkey, 2020). Because of ease and utility, risk calculators are widely used in inpatient and outpatient settings. As demonstrated in the examples previously stated, risk calculators that include demographic characteristics such as race, sex, and age violate the bioethical principles of justice and non-maleficence. Additionally, if the allocation of healthcare resources and services are based on biased and reductive demographic classifications, these risk calculators can inflict systematic discrimination on a large scale and are ultimately subject to legal repercussions (Goodman et al., 2025). Ensuring algorithms that influence clinical

decision-making are unbiased and fair is essential for engaging in ethical medical practice. In this section we will evaluate the bioethical and legal implications of including demographic variables within clinical algorithms.

### 7.1. Justice

Clinicians guided by the principles of justice are to treat all patients fairly and equitably (Varkey, 2020). Distributive justice refers to the fair and equitable distribution of services (Varkey, 2020). Farrell et al. (2022) states that, "a just healthcare system recognizes that membership in groups, whether classified by age, race, gender, socioeconomic status, or other descriptors, should not affect the quality of the healthcare that is delivered or who is trained to deliver that care"(Farrell et al., 2022). But, the various representational errors that arise when measuring demographic characteristics and the attribution error associating these observed differences with having biological origins may result in incorrect diagnoses or inadequate treatment of some groups over others. The principle of justice is then violated when individuals are treated differently based on their race, sex, or age instead of the biological mechanisms that these variables represent (Basu, 2023; Farrell et al., 2020; Visweswaran et al., 2025). This is especially salient since risk calculators do not necessarily need these variables to have high predictive performance (Garg et al., 2017; Ghosh et al., 2024). A study published in 2024 from The Journal of the American Medical Association demonstrated that replacing race-based predictors with social determinants of health resulted in statistically similarly performing predictive models, attesting to this claim (Ghosh et al., 2024). Using demographic characteristics as proxies for biological differences introduces considerable error, making their inclusion in algorithms that have the potential for such widespread impact ultimately unethical.

### 7.2. Non-maleficence

Clinical practices guided by non-maleficence are to avoid or minimize patient harm (Varkey, 2020). Varkey (2020) states that, "the application of non-maleficence is for the physician to weigh the benefits against the burdens of all interventions and treatments, to eschew those that are inappropriately burdensome, and to choose the best course of action for the patient" (Varkey, 2020). Theoretically speaking, this aligns with the ultimate purpose of risk calculators. They attempt to choose the best course of action for the patient based on personal health characteristics. But, by including demographic characteristics embedded with societal biases rather than variables that simply capture biological characteristics, representational and allocative harms are introduced, ultimately, threatening non-maleficient practice.. By excluding demographic characteristics in risk calculators, clinicians and algorithm developers mitigate harms stemming from historical and measurement biases. This practice also encourages the use of more accurate and inclusive variables that represent the underlying condition being captured better aligning with the principle of non-maleficence.

### 7.3. Legal implications

Lastly, clinicians are legally obligated to adhere to antidiscrimination laws protecting patients who have been historically subjected to unfair treatment. These traits, which are legally protected from discrimination in the United States, include race, colour, national origin, sex, age, and disability (Goodman et al., 2025, 2023a). In the United States, antidiscrimination protections in healthcare are enforced by the Department of Health and Human Services, which recently extended these provisions to clinical decision support tools (i.e., clinical algorithms) (Goodman et al., 2025). Clinicians and health care organizations routinely using clinical algorithms that include predictors naively promoting discriminatory practices could violate these protections and be subject to federal disciplinary measures.

Racial discrimination lawsuits have already been raised citing Section 1557 of the Affordable Care Act for the use of race-adjusted glomerular filtration rate (GFR) (Goodman et al., 2025). In these cases, the prosecutors are Black patients who are denied placement on the kidney-transplant waiting list because of an eGFR above the eligibility cutoff. While the Health and Human Services department recognizes race-adjusted eGFR as being unlawful because it relies on "discriminatory assumptions and thresholds" essentially using race as a proxy for muscle mass, they have yet to acknowledge issues with these assumptions as they relate to other "protected traits" (i.e., sex, age). Observed differences in average muscle mass across sex and age groups could influence accurate GFR calculations thus impacting an individual's transplant-list placement, similarly to race (Nankivell et al., 2020). As Goodman et al. (2025) argues, "a male patient who is denied placement on the kidney-transplant list because of an eGFR above the eligibility cutoff could sue, alleging sex-based discrimination… (e.g., he could argue that his lower muscle mass or shorter stature render him more similar to an average woman than to an average man)". Similarly, a 30-year-old patient with sarcopenia, as a result of another primary condition (i.e., spinal cord injury, malignancy, metabolic syndrome, physical inactivity, etc.)(Jung et al., 2023), who is denied placement on the kidney transplant list because of an eGFR above the eligibility cutoff could sue, alleging age-based discrimination (e.g., they could argue that their lower muscle mass render them more similar to an average older adult). This highlights the main issues with using demographic features to represent biological mechanisms (i.e., muscle mass): they are often poor proxies prone to measurement and representation error. In an attempt to avoid unintentional race-, sex-, and age-based discrimination in healthcare settings, it is best to simply measure the direct physiological characteristics or biological mechanisms these variables aim to represent (i.e., muscle mass). Legal protections that apply to race may also be applied to sex and age in the future thus more studies must be done to validate these findings and to promote explicit legal language preventing sex-based and age-based discrimination as well.

## 8. Discussion

While risk calculators enhance clinical decision-making by leveraging patient-specific characteristics to predict outcomes and guide clinical decision-making, they also have great potential to do harm if developed without rigorous methodological evaluation. The inclusion of race, sex, and age as predictors within risk calculators has legal and ethical implications requiring clinicians and algorithm developers to scrutinize their theoretical underpinnings. While the use of race as a variable within risk calculators is commonly rejected due to representational biases, this paper underscores the equally critical need to evaluate sex and age for potential sources of bias and the associated clinical, ethical, and legal implications of their inclusion.

The inclusion of binary-coded sex assigned at birth and chronological age in risk calculators has long been accepted as being biologically relevant and uncontroversial. However, this paper demonstrates that these variables, like race, are subject to similar historical biases and can perpetuate inequities. Sources of representational harm stem from a history of structural sexism and ageism embedded in demographic-based variables, combined with reductive measurement practices that eliminate important nuances within these constructs. Sources of allocative harm stem from the inappropriate distribution of resources driven by faulty assumptions that observed group differences are the result of true biological processes, rather than manifestations of societal discriminatory practices. Four risk calculators from a clinical algorithm database were included for the investigation of deployment bias as a source of allocative harm. All selected risk calculators observed differences in risk estimation across demographic categories. These differences were found to be as large as 20 % in some cases (see Table 1). The observed deployment bias of healthcare services reinforces persisting inequities across demographic populations. Harms caused by the inclusion of demographic variables within risk calculators undermines core bioethical principles that guide good clinical practice putting clinicians at risk for serious legal repercussions.

### *8.1. Mitigation strategies*

While the inclusion of race as a predictor is generally condemned, a preliminary framework for evaluating the inclusion of sex in clinical algorithms was recently developed and proposed in January of 2025, called the Framework for Algorithm Inclusion Review for Sex (FAIRS) (Goodman et al., 2025). This framework calls for clinicians and algorithm developers to ask a series of questions determining whether the inclusion of sex is clinically or ethically appropriate. The following questions include: (1) Is sex prognostically necessary? (2) Why is sex prognostically informative? and (3) Would the algorithm's anticipated use "penalize" the disadvantaged sex from Question 2? The ultimate goal of this framework is to ensure that the inclusion of sex is consistent with existing antidiscrimination laws to protect clinicians from unintended discriminatory practices.

While this framework is a great starting point for mitigating against legal consequences of including demographic variables in clinical algorithms, it fails to address the various sources of representational bias that influences observed sex-based differences as discussed in this paper. This framework does not address and mitigate against the effects of historical bias and measurement bias when representing binary-coded sex assigned at birth (or other demographic variables) as reductive proxies for more complex constructs. The assumptions of this framework is that, "there is a strong legal argument for including sex in clinical algorithms when sex-based biologic mechanisms largely mediate sex's association with an outcome"(Goodman et al., 2025). But, current binary-coded

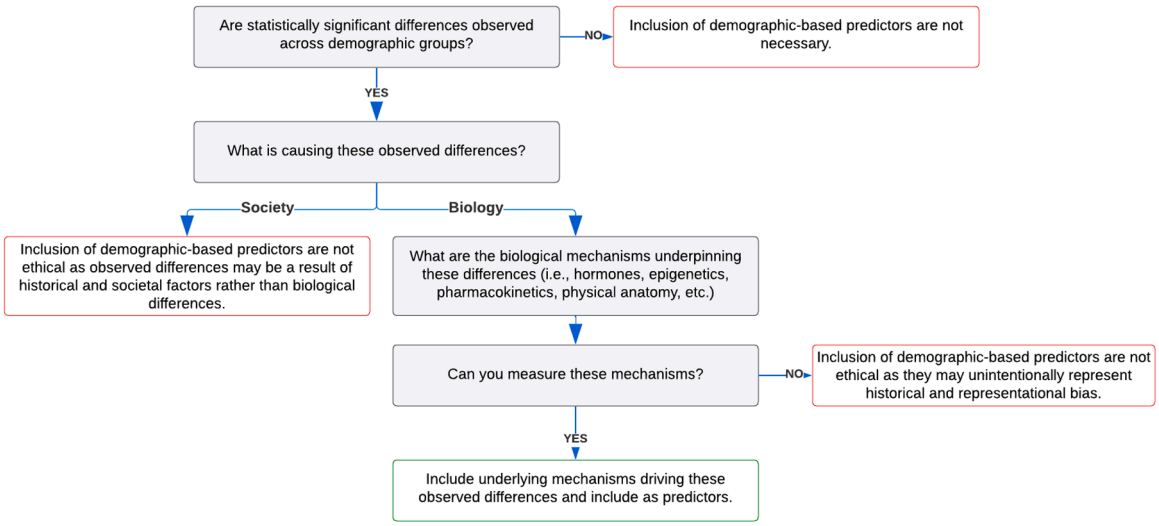


**Fig. 2.** Framework for Evaluating Demographic Variables in Clinical Algorithms
*Note*. Fig. 2 should be used as a conceptual framework for evaluating the ethical inclusion of demographic variables within clinical algorithms. It is important to note the shortcomings of this framework. Historical bias may be represented in the underlying biological mechanisms which may contribute to existing observed differences in the outcome variable. Future research should investigate how to control for historical bias to make more ethical and accurate predictions.

schemas for classifying sex contribute to the erasure of our gender-diverse population and fails to acknowledge that cisgender, transgender, and intersex patients all participate in pharmaceutical and surgical therapies throughout their lifetime impacting their hormones, muscle mass, and anatomy. Directly measuring functional biological processes rather than representing them using proxies is presented as an inclusive way to capture the spectrum of sexual characteristics across the lifespan. Thus, there remains a need for a conceptual framework to guide algorithm developers in thoughtfully approaching the use of demographic variables within clinical algorithms.

The revised framework in Fig. 2, builds on the foundation laid by Goodman et al. (2025). The purpose of this conceptual model is to guide algorithm developers through stages of critical thinking, enabling them to reach their own conclusions about the ethical implications of including a specific demographic variable. This framework may be applied to sex, age, and/or race. Ultimately, this framework prohibits clinicians and developers from using race, sex, or age as proxies for biological mechanisms, such as genetic geographic variation, pharmacokinetics, physiology, and metabolism, etc. and instead forces them to consider what aspects of these variables are important in predicting the desired outcome.

Proxies are often used in the first place because they are inexpensive and easy to measure. Although proxies may be practical, convenience-driven complacency should not guide responsible clinical decision making. For example, race is used to represent genetic geographic ancestry.It is expensive to analyse everyone's DNA so race is used as a more cost-efficient and convenient proxy. Yet, this shortcut is recognized as having major flaws. Similarly, sex is used as a proxy for hormone levels, but it is impractical to measure hormone levels on everyone so sex is used as a cost-efficient proxy. We want to acknowledge that while measuring these exact biological processes or clinical characteristics are costly, this exactness is important in mitigating against unintended harmful effects. If these processes are unable to be reasonably measured, reductive measurements of race, binary-sex assigned at birth, and chronological age from birth should not be used as replacements.

Lastly, it is extremely important to acknowledge that regardless of whether binary sex-based biologic mechanisms largely mediate sex's association with an outcome, these mechanisms are associated with binary sex schemas that remain entangled with historical discrimination and stereotypes reflected in these outcomes. While directly measuring biological mechanisms helps mitigate measurement bias that contributes to inaccurate predictions, disentangling historical bias from this feature is undoubtedly more challenging. That being said, accurately identifying and measuring historical bias should be recognized as an important area of future research assisting the ethical integration and deployment of precision medicine.

### *8.2. Implications for nurse education*

The responsibility of promoting ethically conscious use of clinical algorithms falls on clinicians and clinical education systems. Nurses and nurse practitioners are integral in linking the use of clinical algorithms to appropriate clinical actions following dynamic decision-making processes (Keim-Malpass et al., 2018). This paper serves as a call on clinicians to adopt a more critical and analytical approach when using risk calculators in practice and urges algorithm developers to rigorously scrutinize the inclusion of variables in their models. To do this, clinicians – including nurses – must have a working level of model literacy to critically evaluate model development to assess the potential for bias and representativeness. Competency-based curriculums have been proposed in the medical education literature, and equivalent competencies have not been proposed for nursing (Lee et al., 2024; Russell et al., 2023) education. Nurse scientists should be empowered to reconsider the use of demographic features in their own model development (Keim-Malpass and Moorman, 2021).

## 9. Conclusion

Racism, sexism, and ageism have plagued the healthcare system in various ways throughout history and are therefore represented in the data used to train clinical algorithms used as decision-support tools. Because of this, algorithm developers, including nurse scientists, must ensure transparency in their decision-making processes, clearly articulating the rationale for variable inclusion to avoid unintentional harm caused by ignorance and complacency. By acknowledging and addressing these harms, clinicians and developers can mitigate them and uphold the ethical principles of justice, and non-maleficence in their medical and nursing practice. Lastly, more studies must be done to explore the impact of structural discrimination on observed differences between various demographic groupings to mitigate against this attribution error.

## CRediT authorship contribution statement

**Reanna Panagides:** Writing – review & editing, Writing – original draft, Visualization, Project administration, Investigation, Conceptualization. **Jessica Keim-Malpass:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

## Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Jessica Keim-Malpass reports financial support was provided by National Institutes of Health Agency for Healthcare Research and Quality. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Funding

This work was supported by the National Institutes of Health, Agency for Healthcare Research and Quality [grant number R01HS028803].

## References

Alexander, S.E., Pollock, A.C., Lamon, S., 2022. The effect of sex hormones on skeletal muscle adaptation in females. Eur. J. Sport Sci. 22, 1035–1045. https://doi.org/10.1080/17461391.2021.1921854.

Basu, A., 2023. Use of race in clinical algorithms. Sci. Adv. 9, eadd2704. https://doi.org/10.1126/sciadv.add2704.

Bates, G.W., 1973. Diagnostic applications of ultrasound in obstetrics: a review. J. Tenn. Med. Assoc. 66, 329–332.

Benotti, P., Wood, G.C., Winegar, D.A., Petrick, A.T., Still, C.D., Argyropoulos, G., Gerhard, G.S., 2014. Risk factors associated with mortality after roux-en-Y gastric bypass surgery. Ann. Surg. 259, 123–130. https://doi.org/10.1097/SLA.0b013e31828a0ee4.

Cabral, M., Dillender, M., 2021. Disparities in health care and medical evaluations by gender: a review of evidence and mechanisms. AEA Pap. Proc. 111, 159–163. https://doi.org/10.1257/pandp.20211016.

Cerdeña, J.P., Plaisime, M.V., Tsai, J., 2020. From race-based to race-conscious medicine: how anti-racist uprisings call us to act. Lancet North Am. Ed. 396, 1125–1128. https://doi.org/10.1016/S0140-6736(20)32076-6.

Cho, S.T., Na, H.R., 2022. Urology and geriatrics in Korea: present status and future directions. Int. Neurourol. J. 26, 20–25. https://doi.org/10.5213/inj.2142046.023.

Cram, P., Bayman, L., Popescu, I., Vaughan-Sarrazin, M.S., 2009. Racial disparities in revascularization rates among patients with similar insurance coverage. J. Natl. Med. Assoc. 101, 1132–1139. https://doi.org/10.1016/s0027-9684(15)31109-3.

DeMaria, E.J., Portenier, D., Wolfe, L., 2007. Obesity surgery mortality risk score: proposal for a clinically useful score to predict mortality risk in patients undergoing gastric bypass. Surg. Obes. Relat. Dis. 3, 134–140. https://doi.org/10.1016/j.soard.2007.01.005.

Diebel, L.W.M., Rockwood, K., 2021. Determination of biological age: geriatric assessment vs biological biomarkers. Curr. Oncol. Rep. 23, 104. https://doi.org/10.1007/s11912-021-01097-9.

El-Azab, S., Nong, P., 2023. Clinical algorithms, racism, and "fairness" in healthcare: a case of bounded justice. Big. Data. Soc. 10, 20539517231213820. https://doi.org/10.1177/20539517231213820.

Fakhari, H., Scherr, C.L., Moe, S., Hoell, C., Smith, M.E., Rasmussen-Torvik, L.J., Chisholm, R.L., McNally, E.M., 2024. From calculation to communication: using risk score calculators to inform clinical decision making and facilitate patient engagement. Med. Decis. Making 44, 900–913. https://doi.org/10.1177/0272989X241285036.

Farrell, T.W., Francis, L., Brown, T., Ferrante, L.E., Widera, E., Rhodes, R., Rosen, T., Hwang, U., Witt, L.J., Thothala, N., Liu, S.W., Vitale, C.A., Braun, U.K., Stephens, C., Saliba, D., 2020. Rationing limited healthcare resources in the COVID-19 era and beyond: ethical considerations regarding older adults. J. Am. Geriatr. Soc. 68, 1143–1149. https://doi.org/10.1111/jgs.16539.

Farrell, T.W., Hung, W.W., Unroe, K.T., Brown, T.R., Furman, C.D., Jih, J., Karani, R., Mulhausen, P., Nápoles, A.M., Nnodim, J.O., Upchurch, G., Whittaker, C.F., Kim, A., Lundebjerg, N.E., Rhodes, R.L., 2022. Exploring the intersection of structural racism and ageism in healthcare. J. Am. Geriatr. Soc. 70, 3366–3377. https://doi.org/10.1111/jgs.18105.

Ferrucci, L., Gonzalez-Freire, M., Fabbri, E., Simonsick, E., Tanaka, T., Moore, Z., Salimi, S., Sierra, F., Cabo, R.de, 2019. Measuring biological aging in humans: a quest. Aging Cell 19, e13080. https://doi.org/10.1111/acel.13080.

Garg, N., Muduli, S.K., Kapoor, A., Tewari, S., Kumar, S., Khanna, R., Goel, P.K., 2017. Comparison of different cardiovascular risk score calculators for cardiovascular risk prediction and guideline recommended statin uses. Indian Heart J. 69, 458–463. https://doi.org/10.1016/j.ihj.2017.01.015.

Ghosh, A.K., Venkatraman, S., Nanna, M.G., Safford, M.M., Colantonio, L.D., Brown, T.M., Pinheiro, L.C., Peterson, E.D., Navar, A.M., Sterling, M.R., Soroka, O., Nahid, M., Banerjee, S., Goyal, P., 2024. Risk prediction for atherosclerotic cardiovascular disease with and without race stratification. JAMA Cardiol. 9, 55–62. https://doi.org/10.1001/jamacardio.2023.4520.

Goff, D.C., Lloyd-Jones, D.M., Bennett, G., Coady, S., D'Agostino, R.B., Gibbons, R., Greenland, P., Lackland, D.T., Levy, D., O'Donnell, C.J., Robinson, J.G., Schwartz, J.S., Shero, S.T., Smith, S.C., Sorlie, P., Stone, N.J., Wilson, P.W.F., Jordan, H.S., Nevo, L., Wnek, J., Anderson, J.L., Halperin, J.L., Albert, N.M., Bozkurt, B., Brindis, R.G., Curtis, L.H., DeMets, D., Hochman, J.S., Kovacs, R.J., Ohman, E.M., Pressler, S.J., Sellke, F.W., Shen, W.-K., Smith, S.C., Tomaselli, G.F., American College of Cardiology/American Heart Association Task Force on Practice Guidelines, 2014. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Circulation 129, S49–S73. https://doi.org/10.1161/01.cir.0000437741.48606.98.

Goodman, K.E., Blumenthal-Barby, J., Redberg, R.F., Hoffmann, D.E., 2025. FAIRS — A framework for evaluating the inclusion of sex in clinical algorithms. N. Engl. J. Med. 392, 404–411. https://doi.org/10.1056/NEJMms2411331.

Goodman, K.E., Morgan, D.J., Hoffmann, D.E., 2023a. Clinical algorithms, antidiscrimination laws, and medical device regulation. JAMA 329, 285–286. https://doi.org/10.1001/jama.2022.23870.

Goodman, K.E., Rodman, A.M., Morgan, D.J., 2023b. Preparing physicians for the clinical algorithm era. N. Engl. J. Med. 389, 483–487. https://doi.org/10.1056/NEJMp2304839.

Grimstad, F., Kremen, J., Streed, C.G., Dalke, K.B., 2021. The health care of adults with differences in sex development or intersex traits is changing: time to prepare clinicians and health systems. LGBT Health 8, 439–443. https://doi.org/10.1089/lgbt.2021.0018.

Harper, J., O'Donnell, E., Khorashad, B.S., McDermott, H., Witcomb, G.L., 2021. How does hormone transition in transgender women change body composition, muscle strength and haemoglobin? Systematic review with a focus on the implications for sport participation. https://doi.org/10.1136/bjsports-2020-103106.

Homan, P., 2019. Structural sexism and health in the United States: a new perspective on health inequality and the gender system. Am. Sociol. Rev. 84, 486–516. https://doi.org/10.1177/0003122419848723.

Jain, K.K., 2002. Personalized medicine. Curr. Opin. Mol. Ther. 4, 548–558.

Jung, H.N., Jung, C.H., Hwang, Y.-C., 2023. Sarcopenia in youth. Metabolism 144, 155557. https://doi.org/10.1016/j.metabol.2023.155557.

Keim-Malpass, J., Kitzmiller, R.R., Skeeles-Worley, A., Lindberg, C., Clark, M.T., Tai, R., Calland, J.F., Sullivan, K., Randall Moorman, J., Anderson, R.A., 2018. Advancing continuous predictive analytics monitoring: moving from implementation to clinical action in a learning health system. Critical Care Nurs. Clinics of North Am., Human Factors and Technol. in the ICU 30, 273–287. https://doi.org/10.1016/j.cnc.2018.02.009.

Keim-Malpass, J., Moorman, L.P., 2021. Nursing and precision predictive analytics monitoring in the acute and intensive care setting: an emerging role for responding to COVID-19 and beyond. Int. J. Nurs. Studies Adv. 3, 100019. https://doi.org/10.1016/j.ijnsa.2021.100019.

Kharbanda, A.B., Vazquez-Benitez, G., Ballard, D.W., Vinson, D.R., Chettipally, U.K., Kene, M.V., Dehmer, S.P., Bachur, R.G., Dayan, P.S., Kuppermann, N., O'Connor, P.J., Kharbanda, E.O., 2018. Development and validation of a novel pediatric appendicitis risk calculator (pARC). Pediatrics 141, e20172699. https://doi.org/10.1542/peds.2017-2699.

Khosla, L., Gong, S., Weiss, J.P., Birder, L.A., 2022. Oxidative stress biomarkers in age-related lower urinary tract disorders: a systematic review. Int. Neurourol. J. 26, 3. https://doi.org/10.5213/inj.2142188.094.

Kim, J.W., 2022. Biomarkers of aging: the proxy of time. Int. Neurourol. J. 26, 1. https://doi.org/10.5213/inj.2222edi01.

Kollias, T.F., Gallagher, C.P., Albaashiki, A., Burle, V.S., Slouha, E., 2024. Sex differences in appendicitis: a systematic review. Cureus 16, e60055. https://doi.org/10.7759/cureus.60055.

Krieger, N., 2014. Discrimination and health inequities [WWW Document]. URL. https://journals.sagepub.com/doi/abs/10.2190/hs.44.4.b (accessed 11.20.24).

Lee, Y.-M., Kim, S., Lee, Y.-H., Kim, H.-S., Seo, S.W., Kim, H., Kim, K.J., 2024. Defining medical AI competencies for medical school graduates: outcomes of a Delphi survey and medical student/educator questionnaire of South Korean medical schools. Acad. Med. 99, 524. https://doi.org/10.1097/ACM.0000000000005618.
Margolis, C.Z., 1983. Uses of clinical algorithms. JAMA 249, 627–632. https://doi.org/10.1001/jama.1983.03330290049028.
Mohottige, D., Farouk, S., Poteat, T., Radix, A., Witchel, S.F., 2024. Considerations of sex as a binary variable in clinical algorithms. Nat. Rev. Nephrol. 20, 347–348. https://doi.org/10.1038/s41581-024-00840-2.
Nankivell, B.J., Nankivell, L.F.J., Elder, G.J., Gruenewald, S.M., 2020. How unmeasured muscle mass affects estimated GFR and diagnostic inaccuracy. eClinical Med. 29. https://doi.org/10.1016/j.eclinm.2020.100662.
Neugarten, J., Golestaneh, L., 2019. Influence of sex on the progression of chronic kidney disease. Mayo Clin. Proc. 94, 1339–1356. https://doi.org/10.1016/j.mayocp.2018.12.024.
O'Brien, J., Clare, C., 2023. Race-based versus Race-conscious medicine in obstetrics and Gynecology [WWW Document]. URL. https://scholar.google.com/scholar_lookup?title=Race-based+versus+race-conscious+medicine+in+obstetrics+and+gynecology&journal=Clin.+Obstet.+Gynecol.&volume=66&issue=1&pages=95-106&publication_year=2023&author=O %E2 %80 %99Brien %2CJ&author=Clare %2CCA&inst=13751854044476728232 (accessed 4.30.25).
O'Brien, S.M., Feng, L., He, X., Xian, Y., Jacobs, J.P., Badhwar, V., Kurlansky, P.A., Furnary, A.P., Cleveland, J.C., Lobdell, K.W., Vassileva, C., Wyler von Ballmoos, M. C., Thourani, V.H., Rankin, J.S., Edgerton, J.R., D'Agostino, R.S., Desai, N.D., Edwards, F.H., Shahian, D.M., 2018. The Society of Thoracic Surgeons 2018 adult Cardiac Surgery Risk Models: part 2-statistical methods and results. Ann. Thorac. Surg. 105, 1419–1428. https://doi.org/10.1016/j.athoracsur.2018.03.003.
Oles, N., Darrach, H., Landford, W., Garza, M., Twose, C., Park, C.S., Tran, P., Schechter, L.S., Lau, B., Coon, D., 2022. Gender Affirming Surgery: a comprehensive, systematic review of all peer-reviewed literature and methods of assessing patient-centered outcomes (Part 1: breast/Chest, Face, and Voice). Ann. Surg. 275, e52–e66. https://doi.org/10.1097/SLA.0000000000004728.
Phillips-Beck, W., Eni, R., Lavoie, J.G., Avery Kinew, K., Kyoon Achan, G., Katz, A., 2020. Confronting racism within the Canadian Healthcare System: systemic exclusion of First Nations from quality and consistent care. Int. J. Environ. Res. Public Health 17, 8343. https://doi.org/10.3390/ijerph17228343.
Poteat, T., Lett, E., Rich, A.J., Jiang, H., Wirtz, A.L., Radix, A., Reisner, S.L., Harris, A.B., Malone, J., La Cava, W.G., Lesko, C.R., Mayer, K.H., Streed, C.G., 2023. Effects of race and gender classifications on atherosclerotic cardiovascular disease risk estimates for clinical decision-making in a cohort of black transgender women. Health Equity 7, 803–808. https://doi.org/10.1089/heq.2023.0066.
Poznyak, A.V., Sukhorukov, V.N., Guo, S., Postnov, A.Y., Orekhov, A.N., 2023. Sex differences define the vulnerability to atherosclerosis. Clin. Med. Insights. Cardiol. 17, 11795468231189044. https://doi.org/10.1177/11795468231189044.
Russell, R.G., Lovett Novak, L., Patel, M., Garvey, K.V., Craig, K.J.T., Jackson, G.P., Moore, D., Miller, B.M., 2023. Competencies for the use of artificial intelligence–Based tools by health care professionals. Acad. Med. 98, 348. https://doi.org/10.1097/ACM.0000000000004963.
Salas-Humara, C., Sequeira, G.M., Rossi, W., Dhar, C.P., 2019. Gender affirming medical care of transgender youth. Curr. Probl. Pediatr. Adolesc. Health Care 49, 100683. https://doi.org/10.1016/j.cppeds.2019.100683.
Soleimanpour, N., Bann, M., 2022. Clinical risk calculators informing the decision to admit: a methodologic evaluation and assessment of applicability. PLoS One 17, e0279294. https://doi.org/10.1371/journal.pone.0279294.
Sonne, J., Leslie, S.W., Lopez-Ojeda, W., 2025. Kallmann Syndrome. StatPearls. StatPearls Publishing, Treasure Island (FL).
Suresh, H., Guttag, J.V., 2021. A framework for understanding sources of harm throughout the machine learning life cycle, in: equity and access in algorithms, mechanisms, and optimization. pp. 1–9. https://doi.org/10.1145/3465416.3483305.
Tangri, N., Stevens, L.A., Griffith, J., Tighiouart, H., Djurdjev, O., Naimark, D., Levin, A., Levey, A.S., 2011. A predictive model for progression of chronic kidney disease to kidney failure. JAMA 305, 1553–1559. https://doi.org/10.1001/jama.2011.451.
Tapper, J., Huang, G., Pencina, K.M., Li, Z., Arver, S., Martling, A., Blomqvist, L., Buchli, C., Travison, T.G., Storer, T.W., Bhasin, S., Basaria, S., 2019. The effects of testosterone administration on muscle areas of the trunk and pelvic floor in hysterectomized women with low testosterone levels: proof-of-concept study. Menopause 26, 1405–1414. https://doi.org/10.1097/GME.0000000000001410.
Varkey, B., 2020. Principles of clinical ethics and their application to practice. Med. Princ. Pract. 30, 17. https://doi.org/10.1159/000509119.
Visweswaran, S., Sadhu, E.M., Morris, M.M., Vis, A.R., Samayamuthu, M.J., 2025. Online database of clinical algorithms with race and ethnicity. Sci. Rep. 15, 10913. https://doi.org/10.1038/s41598-025-94152-5.
Vyas, D.A., Eisenstein, L.G., Jones, D.S., 2020. Hidden in plain sight — Reconsidering the use of race correction in clinical algorithms. N. Engl. J. Med. 383, 874–882. https://doi.org/10.1056/NEJMms2004740.