



Rethinking Race and Ethnicity in Biomedical Research

*Committee on the Use of Race and Ethnicity in
Biomedical Research*

OCTOBER 2024

“When we talk about the concept of race, most people believe that they know it when they see it but arrive at nothing short of confusion when pressed to define it.”

Evelyn Higginbotham, 1992

Summary of Statement of Task

Assess the current use of the social constructs of race and ethnicity in biomedical research and provide recommendations to guide the scientific community in the future use of race and ethnicity in biomedical research.

- Document and evaluate how racial and ethnic categories are currently being used in biomedical research
- Review existing guidance for researchers on the use of race as a variable in biomedical research
- Identify circumstances in which it is appropriate to use race and ethnicity in biomedical research and circumstances in which race and ethnicity should not be used

Study Sponsors

THE DORIS DUKE FOUNDATION
THE BURROUGHS WELLCOME FUND

Committee Membership

M. Roy Wilson, *Chair (NAM)*

Wayne State University

Allison E. Aiello, Ph.D.

Columbia University

Efren J. Flores, M.D.

Massachusetts General Hospital

Carmen E. Guerra, M.D., M.S.C.E.

University of Pennsylvania

Elizabeth Heitman, Ph.D.

University of Texas Southwestern

Matthew F. Hudson, Ph.D., M.P.H.

Prisma Health

Husseini K. Manji, M.D. (NAM)

UK Govt Mental Health Mission

Amy Moran-Thomas, Ph.D.

Massachusetts Institute of Technology

Margaret Moss, Ph.D., J.D., RN (NAM)

University of Minnesota, School of Nursing

Elizabeth O. Ofili, M.D., M.P.H. (NAM)

Morehouse School of Medicine

Neil R. Powe, M.D., M.P.H., M.B.A. (NAM)

University of California, San Francisco

Aliya Saperstein, Ph.D.

Stanford University

Roland J. Thorpe, Jr., Ph.D.

Johns Hopkins Bloomberg School of
Public Health

Shyam Visweswaran, M.D., Ph.D.

University of Pittsburgh

Genevieve L. Wojcik, Ph.D.

Johns Hopkins Bloomberg School of
Public Health

Ruqaiijah Yearby, J.D., M.P.H.

The Ohio State University

What Are Race and Ethnicity?

Race and ethnicity are...

- Social and political constructs
- Associated with a person's heritage, "roots," or ancestry
- Aspects of personal identity
- Dynamic social divisions that vary across historical, political, and geographic contexts
- Closely intertwined concepts

Race was conceived to describe and categorize people hierarchically; it has been used to include or exclude individuals and groups

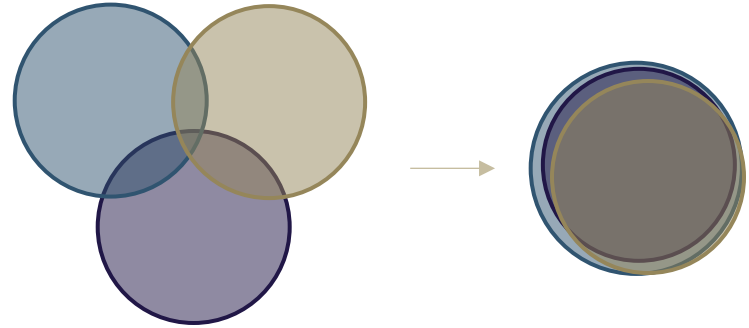
Ethnicity is used to describe people from a similar national or regional background who share common national, cultural, historical, and social experiences

What Race and Ethnicity Are Not

Race and ethnicity *are not*...

- Innate, unchanging characteristics
- Synonymous with skin color
- Based in biology or explained by genetic variation

Research has shown that genetic variation is continuous and overlapping across racial and ethnic groups



Misconception: Human populations form distinct categories with little overlap

New Thinking: Populations mostly overlap

Why Are these Issues Challenging to Solve in Research?

- Issues of race and ethnicity are complex, nuanced, and personal
- Some uses in research reinforced misconceptions that have proven difficult to root out
 - Racial and ethnic categories were assumed to be a useful shorthand, or proxy, for other variables
 - Unexplained variation is sometimes misattributed to race or ethnicity
- Race and ethnicity are used to track diversity of study participants
 - Researchers in the U.S. are typically required to use the Office of Management and Budget (OMB) race and ethnicity categories to report study enrollment

Although it is not biological, **race continues to be relied upon as a blunt tool, heuristic, or variable** in medicine and biomedical research.

CONCLUSION 5-4

Why Continue to Use Race and Ethnicity?

- Given these challenges, some may think that race should not be used at all, or that it is too difficult and complicated to improve how race and ethnicity are used in research.
- However, the committee found that
 - Race and ethnicity are important parts of how people identify
 - Race and ethnicity data are needed to identify and monitor persistent health disparities
 - Tracking race and ethnicity can improve accountability for diverse enrollment

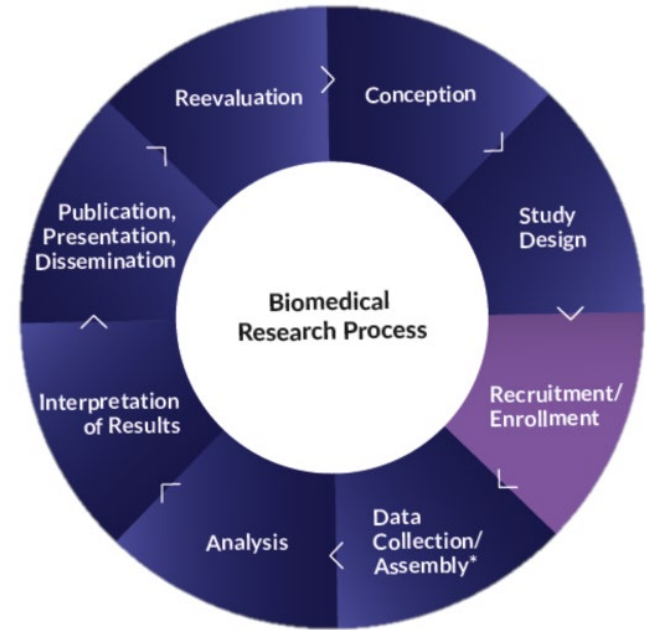
Both deciding to use or deciding to omit race and ethnicity can have advantages and disadvantages in biomedical research. It is important to **evaluate potential implications, benefits, and risks not only of using race and ethnicity but also of forgoing collection of these data entirely.**

CONCLUSION 6-1

The Committee's Approach

Goal: Biomedical researchers and others in the research ecosystem use race and ethnicity data thoughtfully, moving beyond harmful uses that create or perpetuate health inequities.

Approach: Recognizing that biomedical research is varied and multidisciplinary, the committee focused on the research process to develop guidance that would be generalizable to different subfields.



Addressing the use of race and ethnicity at only one stage of a study fails to capture the unique factors and consequences that can emerge at subsequent steps of the process.

CONCLUSION 6-2

Engage in deliberate, ongoing decision-making

At every stage throughout the biomedical research process, researchers should **scrutinize, evaluate, and decide** whether the use of race and ethnicity is appropriate or inappropriate. In their decisions, researchers should consider:

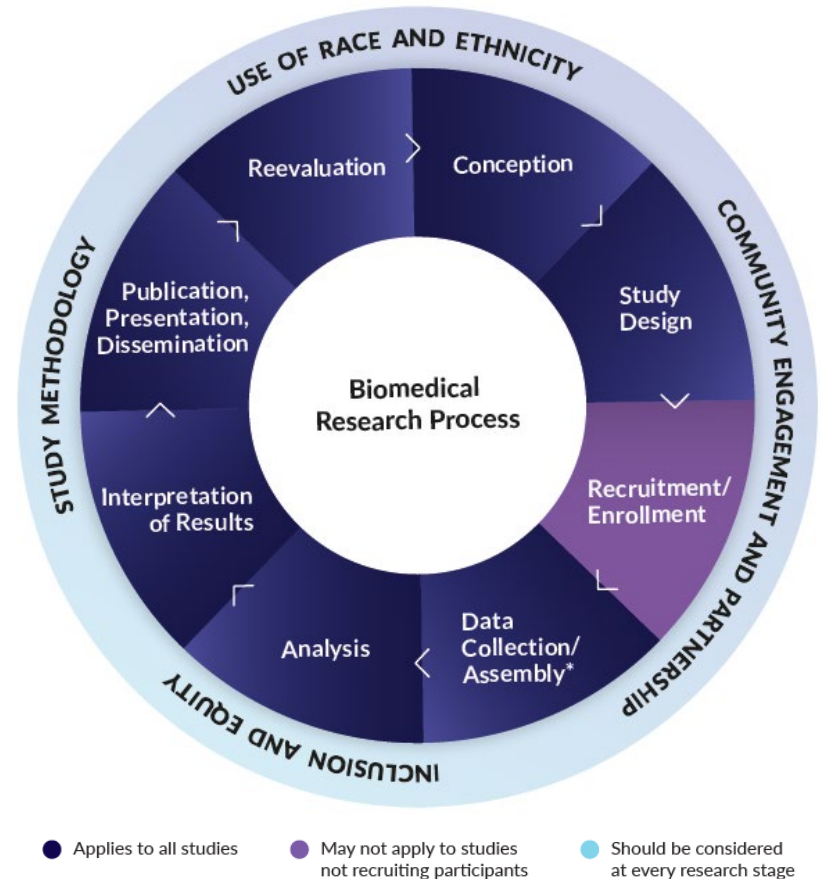
- Historical and social context
- Scientific rationale based on the research question
- Contexts for partnering with specific populations and communities
- Diverse representation and equity
- Whether inferences are supported based on the data and study design
- Potential implications, limitations, benefits, or harms

RECOMMENDATION 1

Overview of the Committee's Approach

The committee identified four considerations to address at every stage of the research process

- 1 Assessing the use of race and ethnicity
- 2 Forming partnerships with communities
- 3 Ensuring inclusion and equity throughout the research process
- 4 Evaluating datasets and study methodology

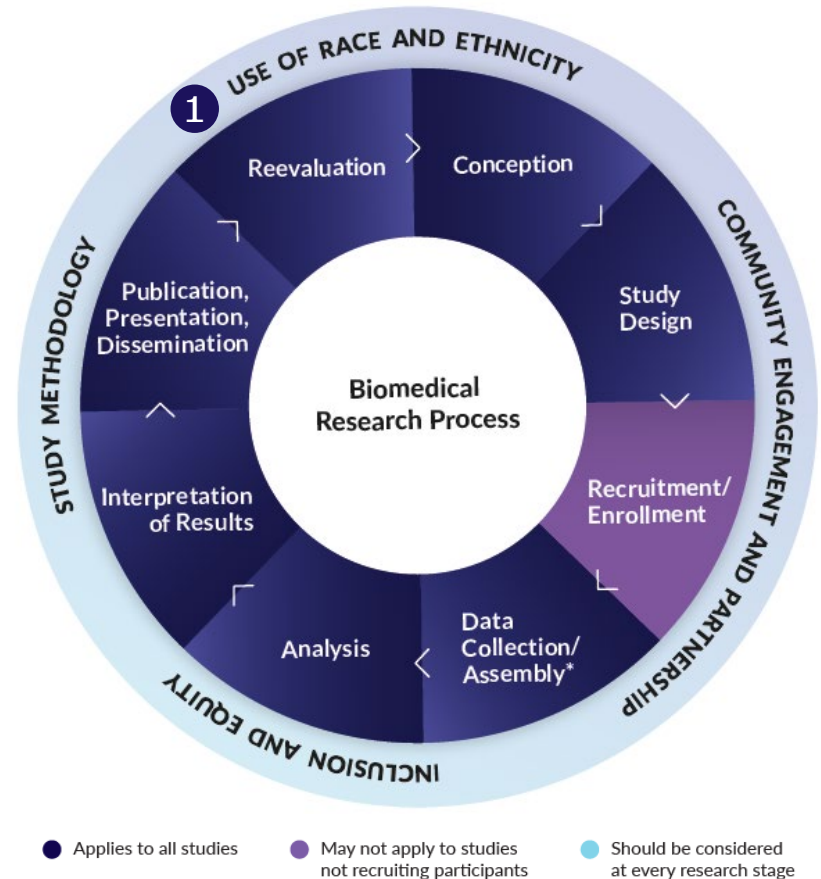


*Researchers conducting observational studies with existing data may enter the cycle here but should consider data provenance and prior stages.

CONSIDERATIONS TO ADDRESS AT EVERY STAGE OF THE RESEARCH PROCESS

1 Assessing the use of race and ethnicity

- 2 Forming partnerships with communities
- 3 Ensuring inclusion and equity throughout the research process
- 4 Evaluating datasets and study methodology



*Researchers conducting observational studies with existing data may enter the cycle here but should consider data provenance and prior stages.

Assessing the Use of Race and Ethnicity

- Race and ethnicity are often defined and used differently across biomedical research, contributing to confusion and misunderstanding
- Race and ethnicity categories frequently serve as a proxy, or stand-in, for other concepts, variables, and unknowns
- Various concepts are often collapsed into a single “race or ethnicity” descriptor or variable



Race and ethnicity conflate many concepts and collapse multidimensional information about people's experience and identity. There is a **need for disaggregation of related concepts** and for increased granularity in the data collected to better capture the information for which race has been a proxy.

CONCLUSION 6-4

Identify applicable concepts related to race and ethnicity

Researchers should strive to identify which concepts often conflated with race or ethnicity are relevant to their study—that is, **environmental, economic, behavioral, and social factors, including those related to racism.**

Based on those concepts, researchers should select applicable measures, using multiple to the greatest extent possible.

RECOMMENDATION 4

Assessing the Use of Race and Ethnicity

- Contemplate key questions:
 - If race and ethnicity are to be included, what is the purpose for including them? How do these constructs relate to factors that influence health?
 - What associated concepts could be measured to better inform or complement the analysis? Can multiple measures be included, analyzed and compared?
- Identify relevant concepts and specific measures based on the research questions of interest. For example:

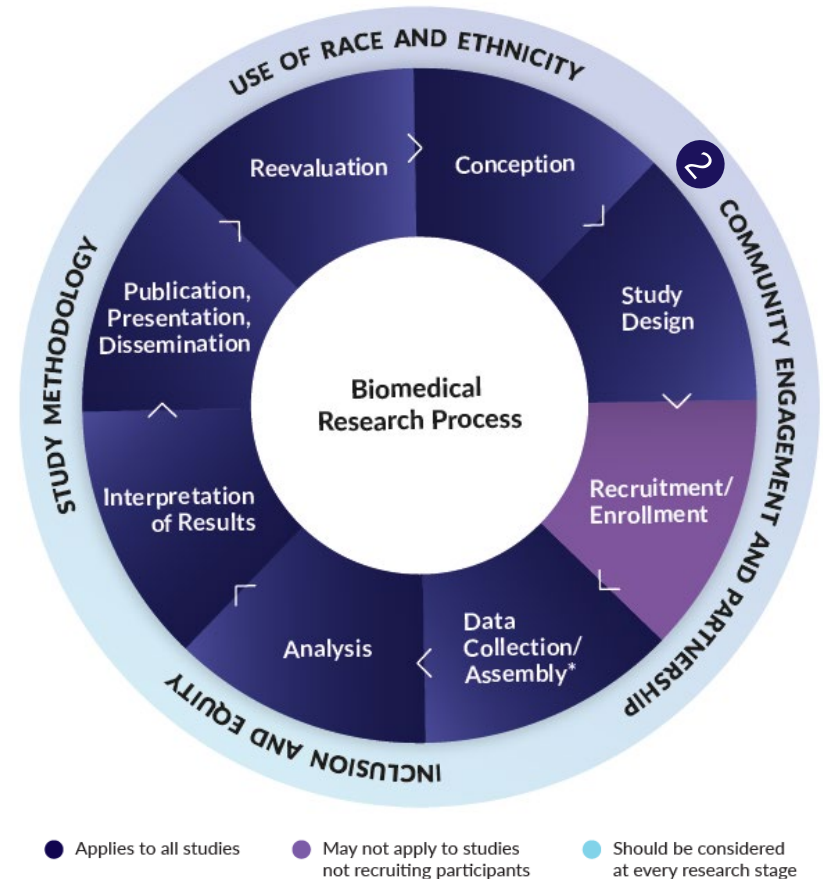
– Relational aspects of race – Structural racism – Social determinants of health – Ethnicity (e.g., language, religion) – Immigration status	– Indigeneity – Skin color or pigmentation – Known ancestry – Genetic markers – Biomarkers and other biological indicators
--	--
- Analyze data to tease apart underlying mechanisms

Full Details

- Recommendation 4
- Table 6-1 Race, Ethnicity and Associated Concepts
- Background: Chapter 5 Reconceptualizing the Use of Race and Ethnicity in Biomedical Research

CONSIDERATIONS TO ADDRESS AT EVERY STAGE OF THE RESEARCH PROCESS

- 1 Assessing the use of race and ethnicity
- 2 Forming partnerships with communities**
- 3 Ensuring inclusion and equity throughout the research process
- 4 Evaluating datasets and study methodology



*Researchers conducting observational studies with existing data may enter the cycle here but should consider data provenance and prior stages.

Forming Partnerships with Communities

- Community leaders and members can provide valuable input throughout the research process by:
 - Identifying and prioritizing research questions
 - Helping with recruitment and data collection
 - Designing protocols, including collection of race and ethnicity information
 - Disseminating results in their communities
- Some populations, such as Tribal Nations, may have unique needs, preferences, or requirements
- Challenges remain, including:
 - Some engagement efforts are short-lived or perfunctory
 - Funding models and reporting timelines do not always support in-depth, sustained community partnerships



Sustained community engagement requires investment

Researchers collecting and using race and ethnicity data in biomedical research ... should collaborate with community engagement experts and, to the greatest extent possible, partner directly with community members to **optimize authentic, continuous, and sustained researcher-community member engagement undergirded by mutual trust.**

RECOMMENDATION 7

Funders should **provide resources and timelines that encourage researchers to build and sustain collaborations.**

Research institutions, medical centers, and other biomedical research organizations should develop and support lasting, equitable relationships with community partners.

RECOMMENDATION 9

Forming Partnerships with Communities

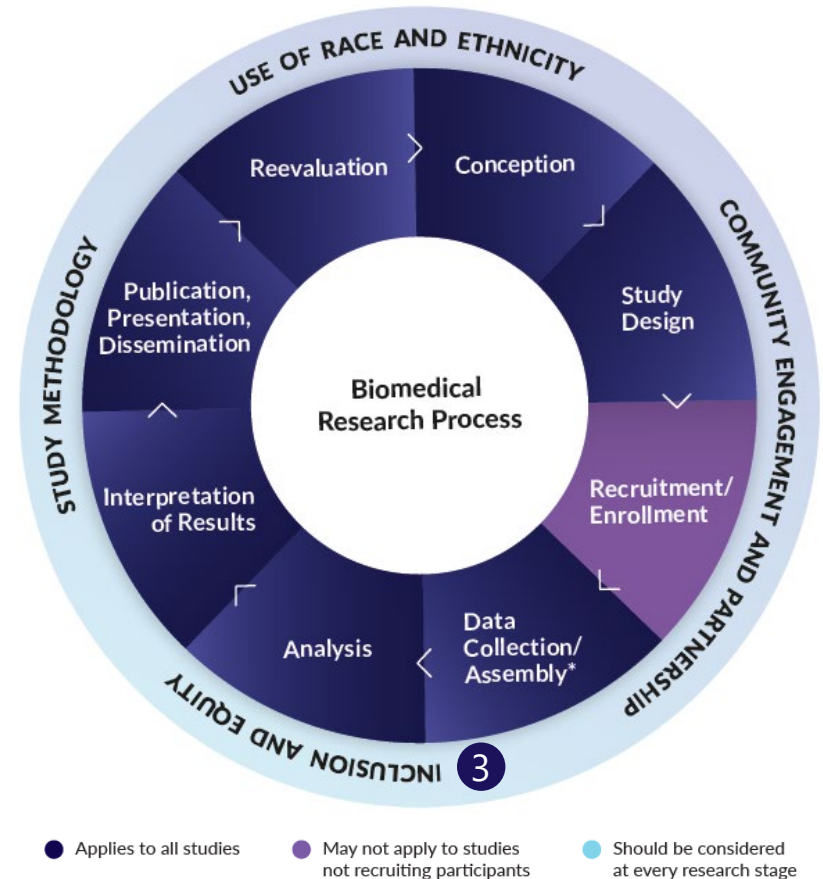
- First steps will depend on the research context and the resources and networks available via the research institution
- Contemplate key questions:
 - How would community partnerships be effective for accomplishing the study's specific aims and for understanding community views on race and ethnicity? What type of engagement would meet these needs?
 - What efforts have been made to understand the steps and account for the time required for successful community outreach?
 - How can we maintain community relationships after this study has concluded?
- There are key roles for funders, sponsors, research institutions, and medical systems to ensure lasting, meaningful engagement

Full Details

- Recommendation 7
- Recommendation 9
- Figure 6-2 Community Engagement and Translational Stages of Biomedical Research
- Chapter 6, section “Building Community Partnerships”
- Background: Chapter 4, section “Guidance for Community and Participant Engagement” including Table 4-1 Community Engagement Continuum

CONSIDERATIONS TO ADDRESS AT EVERY STAGE OF THE RESEARCH PROCESS

- 1 Assessing the use of race and ethnicity
- 2 Forming partnerships with communities
- 3 Ensuring inclusion and equity throughout the research process**
- 4 Evaluating datasets and study methodology



*Researchers conducting observational studies with existing data may enter the cycle here but should consider data provenance and prior stages.

Ensuring Inclusion and Equity

- Participants are sometimes left out of research analyses due to ...
 - Missing data
 - Small group sizes
 - A lack of categories that fit their identity
 - Their selection of multiple race and ethnicity categories
- Multiracial identification is increasing in the U.S.
- There is no standard way to account for members of small populations or people who are multiracial
- More research is needed to identify best practices



Ensure inclusion throughout the research process

At each stage of the research process, **all racial or ethnic category inclusions and exclusions should be based on a clear scientific rationale** motivated by the research question.

Researchers should explain analytic decisions such as the choice of reference population or the reason for aggregating categories.

Researchers should not aggregate participants into nonspecific categories like “Other.”

RECOMMENDATION 5

Researchers should consider the inclusion and analysis of multiracial and multiethnic participants at each stage of the research process.

Researchers should **select a classification scheme for including multiracial and multiethnic participants in analysis**, based on the research question or context.

RECOMMENDATION 6

Ensuring Inclusion and Equity

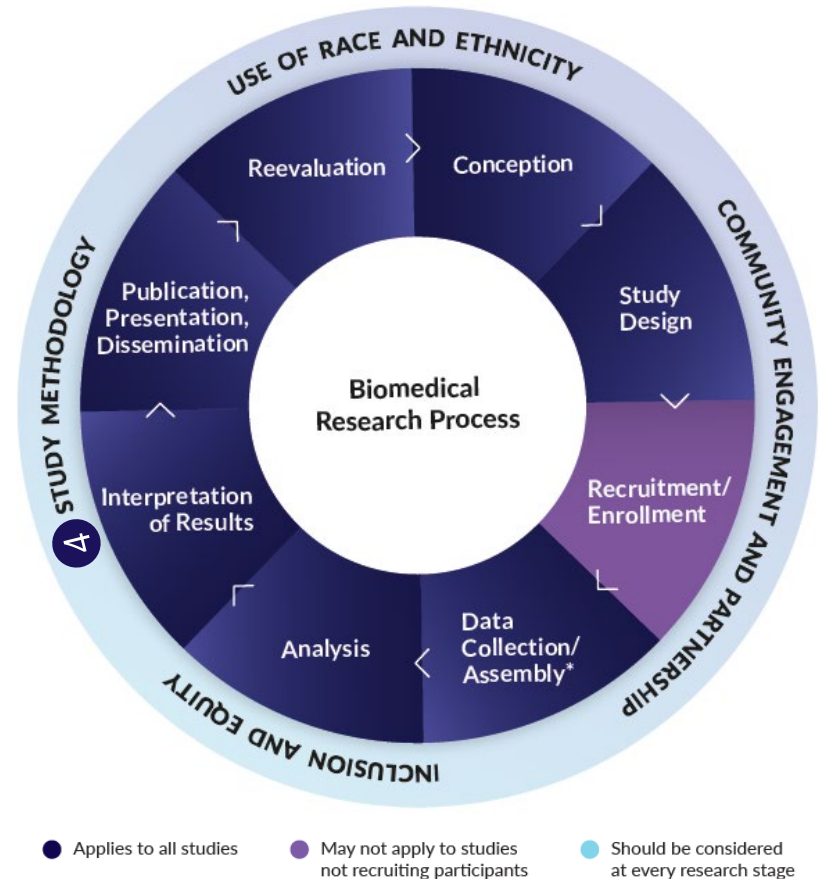
- Design studies from the beginning with small populations and multiracial people in mind
- Contemplate key questions:
 - What methods will be used to recruit and include a diverse sample of participants? Are oversamples needed to include smaller populations in analysis?
 - How is multiracial ancestry or multiracial identity relevant to the research context? What mechanism may be driving the outcome of interest?
 - What is known about this context for individuals with multiracial ancestry or identity? What does existing evidence suggest about aspects of multiracial identity that are most relevant?
- Consider comparing results using alternate approaches for multiracial analysis

Full Details

- Recommendation 5
- Recommendation 6
- Chapter 6, section “Incorporating People Who Are Members of Small Populations and Who Identify as Multiracial or Multiethnic”
- Background: Chapter 5, section “Approaches to Analyzing Data from Multiracial Individuals” including Table 5-1 Multiracial Categorization Schemes Currently in Use in the Literature

CONSIDERATIONS TO ADDRESS AT EVERY STAGE OF THE RESEARCH PROCESS

- 1 Assessing the use of race and ethnicity
- 2 Forming partnerships with communities
- 3 Ensuring inclusion and equity throughout the research process
- 4 Evaluating datasets and study methodology**



*Researchers conducting observational studies with existing data may enter the cycle here but should consider data provenance and prior stages.

Evaluating Datasets and Study Methodology

- Biomedical research is often conducted with previously collected datasets, known as *secondary data*, such as legacy data from prior studies or electronic health records
- Limitations can include underrepresentation, bias, and incomplete or inaccurate race and ethnicity data
- Increasing use of AI could exacerbate problems of bias in existing datasets



Some legacy race and ethnicity data should no longer be used. Since knowledge and reporting have changed over time, **combining legacy with current datasets is problematic**. Particularly common issues in legacy data are missing race and ethnicity data and aggregating data into an “Other” category.

CONCLUSION 6-3

Characterize datasets and disclose limitations

Whether conducting primary research or secondary data analysis, biomedical researchers should **provide an operational definition of race and ethnicity**, if used, in all grant applications, manuscripts, and related products.

Within these products, researchers should **explain their rationale and the limitations of their approach** as well as describe attributes of data provenance.

RECOMMENDATION 2

Researchers should operate with transparency at every stage in the development, application, and evaluation of biomedical technology that may influence health (e.g., clinical algorithms, AI models and tools, medical devices).

Researchers should **assess and report the performance of biomedical technology across a range of racial and ethnic groups**.

RECOMMENDATION 3

Evaluating Datasets and Study Methodology

- Explicitly define how race and ethnicity data were collected, assembled, and used
- Contemplate key questions:
 - What is the source of the data: primary/original data collection, secondary data (e.g., legacy study data, EHR-derived, or financial claims)? What are the associated limitations?
 - What is the provenance of secondary datasets? Is there bias? What efforts have been made to acknowledge or reconcile the limitations of the data?
 - How will race and ethnicity be used in analysis, if at all? How could context (e.g., social, historical, environmental) influence analysis and interpretation of results?
- Even if race is not an input, conduct and report subgroup analyses to assess fairness and bias in performance

Full Details

- Recommendation 2
- Recommendation 3
- Box 6-1 Questions for Researchers to Consider
- Background: Chapter 3, section “Race and Ethnicity in Secondary Data Analysis”; Chapter 4 Existing Guidance on Using Race and Ethnicity in Biomedical Research

Key Takeaways

Appropriate use of race and ethnicity is **context dependent, nuanced**, and requires **ongoing evaluation**.

Researchers should **evaluate potential uses and articulate choices** and decisions throughout the research process, not only at the beginning or end.

Race and ethnicity are not biological variables.

Racial and ethnic categories can still be useful for some research purposes like recruitment but are not always fit-for-purpose.

Researchers should **examine other concepts and factors** (biological, social, environmental) that have more direct effects on health, instead of relying on race and ethnicity as a shorthand.



*Researchers conducting observational studies with existing data may enter the cycle here but should consider data provenance and prior stages.

Change Requires Coordination and Mutual Accountability

Funders, sponsors, publishers, and editors of biomedical research should **provide consistent guidelines** and should periodically **evaluate their policies on the use of race and ethnicity**.

RECOMMENDATION 8

- Researchers need support, including resources, from the broader biomedical research community
- Appropriate use of race and ethnicity is context-dependent, and conversations about adopting the report's recommendations are needed across subfields of biomedicine

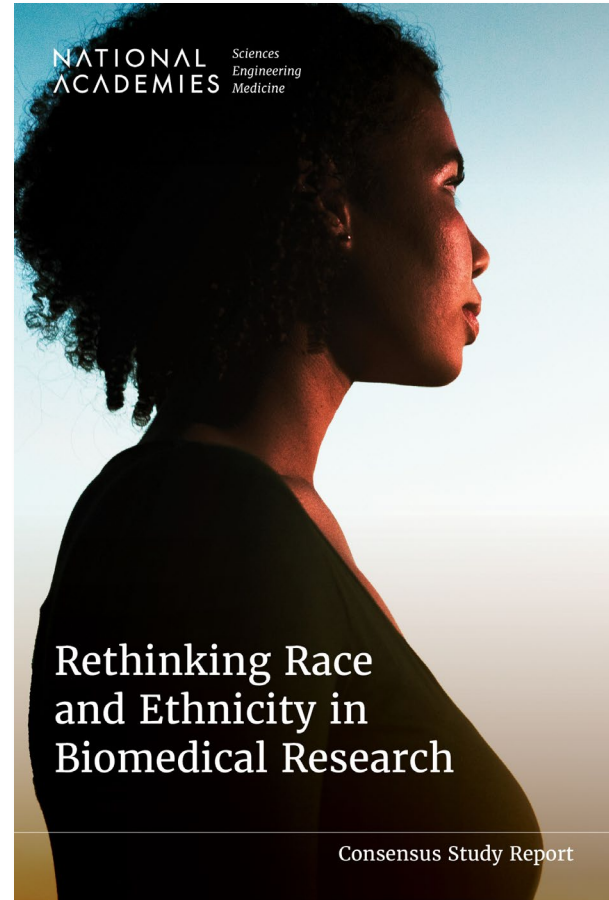


Parting Thoughts

The emphasis on **race has overshadowed other concepts** such as racism and discrimination and their effects on health.

The changes outlined in this report require **rethinking ingrained assumptions** and methods as well as investment in **rebuilding the evidence base** to advance science by improving its rigor and accuracy.

It is an exciting time for the biomedical research community to chart a path forward to improve the use of race and ethnicity for **better science and better health**.



Report Resources

- A free PDF of the report is available to read and download at <http://www.nap.edu/27913>
- Related materials are available through the study page, including:



Scan the QR code to access

- Report highlights
- Interactive webpage
- Action guides
- Educational 1-pagers
- FAQs

Thank You!

