

Collecting Race and Ethnicity Data

Guidelines for Community
Supervision Agencies

About APPR

Advancing Pretrial Policy and Research (APPR) launched as an initiative in 2019 with support from Arnold Ventures. APPR seeks to achieve fair, just, effective, and equitable pretrial practices that positively impact people, systems, and communities. Its mission is to support and engage pretrial professionals and community members in the meaningful delivery of data-informed pretrial justice.

The **Center for Effective Public Policy** (cepp.com) leads all implementation and technical assistance activities for APPR. CEPP is a national nonprofit that helps practitioners, policymakers, and communities reimagine a justice system that works for all. CEPP works on projects across the spectrum of the criminal justice system—from pretrial to sentencing, as well as corrections, probation, parole, and reintegration.



Center for Effective Public Policy

Contents

Introduction	1
Frequently Asked Questions	2
1. Why should community supervision agencies collect race and ethnicity data?	2
2. What standards exist to collect race and ethnicity data?	2
3. How do disparities manifest in community supervision?	5
4. How can collecting and analyzing race and ethnicity data impact racial disparities in community supervision?	6
5. How can community supervision agencies manage data collection challenges?	7
Developing a Race and Ethnicity Data Protocol: Questions to Consider	8

Introduction

Data collection and analysis rightly play a major role in the operations of criminal legal systems. Local, state, and federal justice agencies, as well as policy and research institutions, all rely on data collection and analysis to guide decisions on policy and practice. In the last 40 years, the community supervision field (pretrial, probation, and parole) has focused on using data to develop evidence-based strategies and to study factors, such as supervision conditions, technical violations, and revocations, that impact community supervision outcomes.¹

*Data analysis is essential to evaluating the current state of systems, agencies, and institutions and where they are heading.*²

A deeper look beyond aggregate data shows that Black, Indigenous, and people of color (BIPOC) populations—in particular, Black people—are overrepresented in community supervision. As of 2022, people who identify as Black made up 14.4 percent³ of the population in the United States but 22 percent and 31 percent of the probation and parole supervision populations, respectively, and 43 percent of people detained pretrial.⁴ Racial and ethnic disparities can have broad negative impacts on BIPOC populations within community supervision and beyond. For example, inequities in community supervision revocations and violations can lead to disparities in incarceration and increase inequality in housing, education, income, and health.⁵

Analyzing data by race and ethnicity helps community supervision agencies understand if their practices, policies, and outcomes for people under supervision are equitable for all. As highlighted by the Urban Institute in 2021⁶ and the National Conference of State Legislatures in 2022,⁷ efforts to address racial disparities can be misguided, ineffective, or counterproductive without accurate data.

This resource answers frequently asked questions about collecting and analyzing race and ethnicity data for community supervision agencies and offers guidelines for this important task.

-
- 1 National Institute of Corrections. (n.d.) *Guidelines for developing a Criminal Justice Coordinating Council: Data Collection and Analysis*. <https://info.nicic.gov/cjcc/node/50>
 - 2 National Institute of Corrections, n.d.
 - 3 Moslimani, M., Tamir, C., Budiman, A., Noe-Bustamante, L., & Mora, L. (2024). *Facts about the U.S. Black population*. Pew Research Center. <https://www.pewresearch.org/social-trends/fact-sheet/facts-about-the-us-black-population/#:~:text=The%20Black%20population%20of%20the,14.4%25%20of%20the%20country's%20population>
 - 4 Kaeble, D. (2024). *Probation and parole in the United States, 2022* (NCJ 308575). Bureau of Justice Statistics. <https://bjs.ojp.gov/document/ppus22.pdf>. Prison Policy Initiative. (2024). New data and visualizations spotlight states' reliance on excessive jailing. https://www.prisonpolicy.org/blog/2024/04/15/jails_update/
 - 5 Bradner, K., Schiraldi, V., Mejia, N., & Lopoo, E. (2020). *More work to do: Analysis of probation and parole in the United States, 2017–2018*. Columbia University Justice Lab. <https://justicelab.columbia.edu/sites/default/files/content/More%20Work%20to%20Do.pdf>
 - 6 Nembhard, S., & Robin, L. (2021). *Racial and ethnic disparities throughout the criminal legal system: A result of racist policies and discretionary practices*. Urban Institute. <https://www.urban.org/sites/default/files/publication/104687/racial-and-ethnic-disparities-throughout-the-criminal-legal-system.pdf>
 - 7 National Conference of State Legislatures. (2022). *Racial and ethnic disparities in the justice system*. https://documents.ncsl.org/wwwncsl/Criminal-Justice/Racial-and-Ethnic-Disparities-in-the-Justice-System_v03.pdf

Frequently Asked Questions

1. Why should community supervision agencies collect race and ethnicity data?

- **It promotes fairness and justice.** Supervision agencies can ensure that people on pretrial, probation, or parole supervision receive the support, access, and resources they need to succeed.
- **It supports evidence-based decision making.** With comprehensive data, decisions can be made using facts rather than assumptions. Supervision agency leaders can identify patterns, correlations, and root causes.
- **It leads to equity in policies, practices, and outcomes.** Disaggregated data helps supervision agencies understand how policies and practices impact different populations based on race and ethnicity and can help agency leaders advance equitable outcomes.
- **It facilitates the customization of solutions.** Specific, localized data allows for solutions that cater to each community or region's unique needs and challenges (e.g., appointment notifications in different languages, etc.).

2. What standards exist to collect race and ethnicity data?

Racial and ethnic data collection standards vary among federal, state, and local agencies. Government and community entities often draw guidance from different sources depending on reporting mandates, available funding, statutory requirements, and local practices. Each source sets minimum standards to collect race and ethnicity data. Jurisdictions, agencies, or organizations are encouraged to expand categories based on the demographics and specific needs of their communities and populations.

- The Office of Management and Budget (OMB) issued a new directive in 2024 that set forth a minimum set of categories that all federal agencies must use when collecting information on race and ethnicity (figure 1).⁸ As of March 28, 2024, race and ethnicity data is collected in one combined question, where respondents can select multiple options relevant to how they identify.⁹ Additionally, “Middle Eastern or North African” is now one of the minimum categories for race and ethnicity.

8 Office of Management and Budget. (2024). *Revisions to OMB's Statistical Policy Directive No. 15: Standards for maintaining, collecting, and presenting federal data on race and ethnicity*. <https://www.federalregister.gov/documents/2024/03/29/2024-06469/revisions-to-ombs-statistical-policy-directive-no-15-standards-for-maintaining-collecting-and>

9 Office of Management and Budget, 2024

Figure 1. Race and Ethnicity Questions with Minimum Categories.

What is your race and/or ethnicity?
Select all that apply.

American Indian or Alaska Native
For example, Navajo Nation, Blackfeet Tribe of the Blackfeet Indian Reservation of Montana, Native Village of Barrow Inupiat Traditional Government, Nome Eskimo Community, Aztec, Maya, etc.

Asian
For example, Chinese, Asian Indian, Filipino, Vietnamese, Korean, Japanese, etc.

Black or African American
For example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Somali, etc.

Hispanic or Latino
For example, Mexican, Puerto Rican, Salvadoran, Cuban, Dominican, Guatemalan, etc.

Middle Eastern or North African
For example, Lebanese, Iranian, Egyptian, Syrian, Iraqi, Israeli, etc.

Native Hawaiian or Pacific Islander
For example, Native Hawaiian, Samoan, Chamorro, Tongan, Fijian, Marshallese, etc.

White
For example, English, German, Irish, Italian, Polish, Scottish, etc.

- The updated directive issued by OMB also set forth standards for the collection, compilation, and presentation of race and ethnicity data:¹⁰
 - **Categories and Definitions:** Detailed information on characteristics, specific affiliations, and place of origin is included in the description of each racial and ethnic group to support identification.
 - **Method of Data Collection:** Collecting information that allows self-identification (information provided directly by an individual in written or verbal form) of race and ethnicity separately is preferred. When self-reporting is not feasible, the new format recommends the collection of race and ethnicity data in one question and the selection of all applicable responses.
 - **Presentation of Data:** Sharing detailed information on race and ethnicity is encouraged unless doing so violates policies on data quality, privacy, or confidentiality. Additionally, the following data presentation approaches are recommended to reflect the selection of multiple race and ethnicity categories: 1) **alone or in combination** (i.e., all people who belong to a particular racial or ethnic group,

Data Presentation Approaches

The following example illustrates the categories into which someone who reports being both Asian and Black or African American would fall:

- **Alone or in combination:** “Asian alone or in combination” and “Black or African American alone or in combination”
- **Most frequent multiple responses:** “Asian alone,” “Black or African American alone,” and “Asian and Black or African American”
- **Combined multiracial and/or multiethnic responses:** “Multiracial and/or multiethnic”

¹⁰ Office of Management and Budget, 2024

whether alone or in combination with another), 2) **most frequent multiple responses** (i.e., the seven minimum race and ethnicity categories alone and all reported race and ethnicity combinations), and 3) **combined multiracial and/or multiethnic responses** (i.e., the seven minimum race and ethnicity categories alone and a “multiracial or multiethnic” category for people who report multiple race and/or ethnicity categories).¹¹

- The Conference of State Court Administrators and the National Center for State Courts have developed court data standards through the National Open Court Data Standards (NODS) project (figure 2). Race and ethnicity data are collected in all case types, and designations are broader than those included in the OMB directive. Collecting this level of information allows courts to identify issues with:
 - Access and fairness
 - Operational needs, such as interpreters
 - Equal representation in court programs that meet community needs (e.g., pretrial supervision, diversion programs, specialty courts)¹²

Figure 2. Racial Equity Data Elements from NODS Data Element Spreadsheet.

	Data Element	Definition	Values
6	Race <i>(Select all that apply)</i>	Party's identification with one or more social groups	White Black or African American American Indian or Alaska Native Asian Native Hawaiian or other Pacific Islander Other
6a	Race source	The source or agency where the race data was collected	Court (direct inquiry) Driver's license Law enforcement Jail Corrections/Probation Prosecutor Petitioner Another state agency Unknown
6b	Race self-identified or observed	An indicator for whether the race source relies upon self-identification by the party or an observation from someone else (e.g., law enforcement)	Self-identified Observed or perceived
7	Ethnicity	Party's identification with an ethnicity	Hispanic/Latinx Non-Hispanic/Latinx
7a	Ethnicity source	The source or agency where the ethnicity data was collected	Court (direct inquiry) Driver's license Law enforcement Jail Corrections/Probation Prosecutor Petitioner Another state agency Unknown
7b	Ethnicity self-identified or observed	An indicator for whether the ethnicity source relies upon self-identification by the party or an observation from someone else (e.g., law enforcement)	Self-identified Observed or perceived

11 Please refer to OMB's updated directive (Office of Management and Budget, 2024) for additional information and examples on presenting multiple race and ethnicity data.

12 National Center for State Courts. (2023). *National Open Court Data Standards (NODS)*. <https://www.ncsc.org/consulting-and-research/areas-of-expertise/data/national-open-court-data-standards-nods>

- States can also issue guidance or mandates to collect race and ethnicity data. For example, Washington State sets forth standards on how agencies should collect race and ethnicity data on employees and the populations they serve. A memorandum by the state’s Office of Financial Management outlines minimum standards for collecting race information, including the ability to select more than one race.¹³ The state also guides how to report data when multiple races are selected and requires that a question identifying ethnicity be placed before the question identifying race.

3. How do disparities manifest in community supervision?

There is evidence of racial disparities at various points in the criminal justice system, such as arrest, prosecution, pretrial detention, sentencing, and incarceration. BIPOC populations are overrepresented in pretrial, probation, and post-release (parole) in comparison to white people.

- Often, the percentage of BIPOC people placed on community supervision exceeds the rate at which they are represented in the U.S. population. Based on a review of data collected by the Bureau of Justice Statistics (BJS), the Columbia Justice Lab concluded that white people were underrepresented in parole and probation, “...while Black, Latinx, American Indian, and Alaska Native people were overrepresented in at least one form of community supervision.”¹⁴ The data revealed that Black people were twice as likely to be on probation, four times as likely to be on parole, and three times as likely to be detained pretrial compared to white people.¹⁵
- While the available research on racial disparities in supervision violation charges and outcomes is limited, the information available shows that disparities exist in these areas. For example, based again on the BJS data analyzed by the Columbia Justice Lab, Black people are more likely to return to prison for a parole violation, which in turn increases disparities in incarceration.¹⁶

A report by the Council of State Governments Justice Center found that, in 2020, twenty states had an increase in parole revocation rates for Black people when compared to white people.¹⁷

¹³ Opitz, W. (2022). *Standards for data collection and reporting on race/ethnicity* [Memorandum]. Office of Financial Management. https://ofm.wa.gov/sites/default/files/public/legacy/pop/asr/ofm_standards_race_ethnicity_data.pdf

¹⁴ Bradner et al., 2020, p. 6

¹⁵ Bradner et al., 2020. Prison Policy Initiative, 2024

¹⁶ Bradner et al., 2020

¹⁷ CSG Justice Center. (2024). *Supervision violations and their impact on incarceration*. <https://projects.csjusticecenter.org/supervision-violations-impact-on-incarceration/key-findings/>

- Academic literature on pretrial populations shows that racial disparities also persist on the front end of the system. BIPOC people are at least 10–25 percent more likely than white people to be detained pretrial. They also receive higher money bond amounts and are less likely to be able to afford them.¹⁸
- Social determinant factors contributing to or exacerbating racial disparities in community supervision include access to housing, healthcare, and education opportunities; income inequality; and behavioral health challenges. Understanding the impact of these factors on people under pretrial, probation, and parole supervision is an important step in reducing disparities.

4. How can collecting and analyzing race and ethnicity data impact racial disparities in community supervision?

Community supervision leaders and staff can use race and ethnicity data to identify racial disparities across agency operations and programs and then work to address them. They can also use race and ethnicity data to apply a racial equity lens to their work.

- Agencies that collect programmatic information on referrals, enrollments, active caseloads, supervision levels, compliance with conditions of release, violations, revocations, length of stay, completions, and so on can overlay race and ethnicity data across these metrics. This type of analysis identifies whether people on community supervision have equitable experiences relative to each other based on race and ethnicity. For example, if BIPOC people have higher technical violations leading to revocation and incarceration, this type of analysis can answer whether programmatic or operational factors contribute to these outcomes.

Analyzing existing performance measure metrics by race and ethnicity allows leaders to identify disparities and disproportionalities.

- Supervision agencies can use race and ethnicity data to understand the impacts of policies and practices on racial disparities and to shape policy changes to ensure equitable outcomes for all, regardless of race. For example, the Columbia Justice Lab notes that staying on supervision longer disproportionately impacts BIPOC people, making them more likely to be charged with violations that lead to harsher punishment.¹⁹ Agencies can analyze the length of supervision disaggregated by race and ethnicity to identify any policies or practices contributing to disparate outcomes for BIPOC

¹⁸ Sawyer, S. (2019). *How race impacts who is detained pretrial*. Prison Policy Initiative. https://www.prisonpolicy.org/blog/2019/10/09/pretrial_race/#:~:text=Across%20the%20country%2C%20Black%20and,detained%20pretrial%20than%20white%20defendants

¹⁹ Bradner et al., 2020

populations, if any exist. They can then implement a policy to reduce the length of supervision based on compliant behavior.

- Collecting race and ethnicity data informs operational needs by identifying internal and external resource gaps. For example, jurisdictions that serve large non-English-speaking populations may need access to translation services or bilingual staff. Additionally, collecting this level of data can help identify potential cost efficiencies. For example, agencies serving large non-English-speaking populations may discover that it is more cost-effective or efficient to hire bilingual staff rather than to use language lines or interpreter services. Collecting race and ethnicity data can also enhance agency reports for important constituencies like community, civic, and legislative stakeholders.

Reducing or eliminating racial disparities can help with cost-saving efforts, improve program and organizational performance, and enhance relationships with communities of color that are disproportionately impacted by community supervision.

5. How can community supervision agencies manage data collection challenges?

Collecting data can prove challenging at times; factors like accuracy, consistency, analysis, and interpretation influence data. Differences in the definition of data categories and metrics are common within an organization and across systems or partners. Understanding the purpose of collecting race and ethnicity data and how the agency plans to use the data can help leaders anticipate barriers and manage challenges when they arise. The following practices can help agencies mitigate data collection challenges:

- Understand when and how race and ethnicity data is collected. If the agency or organization is interested in collecting self-identified or self-reported data, or if data is being collected based on observation or by a third party, identify at which point this information should be captured. In addition, when data is collected based on observation or from a third party, determine which staff is best suited to ensure accuracy and consistency.

Potential Challenges When Using Technology or Manual Systems to Collect Data

- Where information is collected in existing data systems, the ability to create or update fields to incorporate race and ethnicity could be limited.
- When existing race and ethnicity fields are not required, they can be bypassed or forgotten.
- Agencies or organizations relying on manual data collection can grapple with accuracy and consistency problems.
- Adding race and ethnicity data fields to manual collection protocols can have a considerable impact on staff time.

- When possible, make investments in technology to improve access to and use of automated systems for data collection, and develop quality assurance protocols to facilitate manual data collection.
- To increase staff buy-in, include in staff trainings information on the prevalence of racial disparities across all decision points in the criminal legal system, the purpose of collecting race and ethnicity data, and the use of such data. In addition, incorporating motivational interviewing techniques and cultural competence into training can help staff feel more at ease when asking race and ethnicity questions, and it can assist in eliciting accurate information.
- Share race and ethnicity data-collecting policies and practices with partners to help them understand how, when, and at what level of specificity that information is captured. This can help partner agencies that capture less information identify how their data collection and analysis for the impacted program or populations need to be supplemented or how racial data analysis needs to be modified.

Developing a Race and Ethnicity Data Protocol: Questions to Consider

As you work on updating or creating a race and ethnicity data protocol for your agency or organization, consider the following questions:

- Is the agency or organization currently collecting race and ethnicity data? If yes, does it include the minimum categories identified in the updated OMB directive?
- If data is collected, is it self-identified, observed, or collected from a third-party source (e.g., data exchange)?
- Are barriers or limitations (e.g., staffing, training, technology, funding, etc.) impacting the accurate and consistent collection of self-identified or observational data?
- How is the data currently collected being used?
- If data is not currently collected, what is the purpose for beginning to collect it, and how will it be used?
- What champions or partners do you have for this task?
- What factors incentivize collecting race and ethnicity data (e.g., funding opportunities, reporting requirements, etc.)?

Collecting and analyzing accurate and consistent race and ethnicity data is the first step to ensuring fair, just, and equitable outcomes for people in the criminal justice system broadly and under community supervision specifically. Acquiring a better understanding of the impact of policies and practices on racial disparities will help community supervision agencies address the underlying factors and issues that exacerbate disparities for BIPOC populations.

