

# The Importance of Context in Biorepository Research

*How to reduce the possibility of harm in repository-enabled research by building context into each step*

---

Most scientists don't mean to harm communities with their research, but research harms communities surprisingly often.<sup>1</sup> These harms can be social, like reinforcing stereotypes.<sup>2</sup> They can also be practical, like when research leads to fewer resources for certain groups.

For example, Obermeyer and colleagues found that a popular healthcare algorithm in the U.S. gave Black patients lower risk scores than White patients with similar needs. The algorithm used healthcare spending to measure health. But because less money is spent on Black patients due to unequal access to care, it wrongly judged them as healthier. This lowered the number of Black patients identified for treatment by more than half. This means that providers who used the algorithm were less likely to give black people the treatment they needed.<sup>3</sup>

To prevent this kind of harm, researchers must include relevant context—the conditions that surround the research that may or may not influence the outcomes—in their work. Research doesn't happen in isolation. It is shaped by many factors, such as:

- How and where the data were collected
- The social and environmental circumstances of participants
- The researcher's own perspective

Group harm often happens when these contexts are ignored or forgotten. The healthcare algorithm was biased *because it ignored the context* that Black patients have unequal access to care. Researchers must think about context from the very beginning to the very end of a project.

# Know Your Dataset

Understanding your data is the first step in being mindful of context in your work. In repository research, researchers often use data that was collected for another reason. This can introduce bias if researchers don't fully understand how this outside context is reflected in the dataset(s).

Researchers can ask themselves questions like:

- What question was this data originally collected to answer?
- Who is included in the data, and who is left out?
- Is there a control or comparison group?
- Does this sample represent the population being studied?
- Is the sample big enough?
- What are the characteristics of any outliers? Are they true outliers or are there just too few of them in the dataset?
- Do I see any bias in the dataset?

After looking closely at the data, researchers might realize that their question doesn't fit the dataset. They may need to change their question or find new data to use.

# Study Design and Analysis Approach

The choices researchers make as they design their study and choose their analysis approach can be critical to including or excluding context. Researchers should design the study keeping in mind what they know about the dataset(s) and any potential bias.

## Accounting for Bias

Although we know that some bias can be corrected for with statistical modeling, our responsibility does not end there. In addition to documenting the bias and approach to

correction, researchers must think critically about any residual bias. For example, many studies "adjust for race" in their analysis. However, adjustment of race may mask racial disparities and perpetuate structural racism.<sup>4</sup>

If you don't have expertise with statistics or study design, it's a good idea to work with someone who does.

## **Clearly Defining Variables**

Another common cause of error is misclassification, when data are put in the wrong category. Researchers need to make sure every variable is clearly defined and measured in the same way.

They can ask questions like:

- How were these variables collected?
- Is there a clear definition for each variable?
- Do these definitions match what might be expected, or what has been used in other studies?
- If variables are being used to "lump" together or "split" apart groups of participants, are they clearly defined enough to do that accurately?

## **Comparison Groups**

It's also important to understand what comparisons are being made. When a study compares groups, it is usually comparing a group with a condition or risk factor to a group without it. These groups should be as similar as possible in all other ways—like age, sex, health, environment, socioeconomic makeup, etc. If they aren't, the results can be misleading.

For example, a Swedish study by Dhejne et al. looked at death rates and crime among transgender people. (You might remember this example from our reading "Cases of Group Harm in Biorepository Research" [link]). The researchers wanted to study the effects of

gender-affirming surgery. They compared transgender people who had the surgery with *cisgender people*. They found higher rates of suicide among the transgender group.<sup>5</sup>

But this result ignored the fact that transgender people face much more discrimination and stress.<sup>6</sup> The study *should* have compared transgender people who had this surgery to *transgender people who had not*. These results have been taken out of context and weaponized in politics and media to argue against trans rights.<sup>7</sup>

## Sharing Findings

After analysis, how researchers share their findings also matters. Researchers can't fully control how others interpret their work, but they can share it in ways that reduce confusion or harm. Key to this is making sure findings are delivered with the relevant context. This is extra important when the research is about vulnerable groups.

A 2024 study by Song and Zhang used UK Biobank data to look at male bisexual behavior, risk-taking, and number of offspring. The authors said that they found a genetic link between male bisexual behavior and risk-taking. They claimed that this explains men in this group having more offspring.<sup>8</sup>

Critics point out that the study had a number of problems from the start.<sup>9-11</sup> For instance, it used people's answers about past sexual experiences to measure bisexuality. The data was also collected in 2006, and only from older adults. "Risk-taking behavior" was measured using only one survey question.

Critics list many ways that these factors likely affected the findings. For example, past sexual experiences do not reveal sexual orientation. Also, fear of discrimination may have influenced the answers people gave. These issues are even more relevant because of the participants' older ages. Many likely lived at times and places when sex between men was illegal.<sup>12</sup> This is relevant social context that suggests non-genetic reasons for the study's results.

But the authors ignored these limitations and made bold claims about genetics. These claims give fuel to harmful stereotypes about bisexual people.<sup>13</sup> The paper doesn't directly say that bisexual men are more promiscuous, but the way it is written implies this. News articles about the study have spread this idea.<sup>14-17</sup> If the authors had put their findings in the relevant social context, it would have sent a different message. But they presented their findings in a way that encouraged harmful takeaways.

## Reflecting on Your Own Context

Researchers bring their own point of view to their work. Everyone has a "lens" shaped by their identity, background, and lived experience. It's important to think about how these things might shape how researchers design a study or understand the results.

After finishing an analysis, researchers can ask:

- Did I have any guesses about how this would turn out?
- Could those guesses have changed how I looked at the results?
- Do the findings really support my idea, or could I have made choices that influenced the outcome?
- Could I be repeating stereotypes without meaning to?

## Exploratory Research

This kind of self-reflection is extra important in exploratory, or hypothesis-generating, research. This is when researchers don't start with a research question, but instead look for patterns. This kind of work can lead to new discoveries. It also has a higher risk for problems like cherry picking, false positives, and confirmation bias.

To lower these risks, researchers should:

- Talk openly about their methods and why they made certain choices
- Use tools to avoid false positives, like correcting for multiple comparisons

- Make it easy for others to try to replicate the work

## Working with Communities

We also recommend partnering with members of the communities in the data when possible. One good example of this is a recent study<sup>18</sup> from the African Ancestry Neuroscience Research Initiative<sup>19</sup> about how genetic ancestry affects gene expression in Black Americans. The researchers running this study collaborated with the organization Black in Neuro.<sup>20</sup> They included community leaders in decision-making from recruitment to publication. This helped make sure their work was accurate and didn't repeat harmful ideas.<sup>18</sup>

## Final Thoughts

Research is not done in a vacuum. It is always shaped by its context.

To avoid harm and make sure their science is accurate, repository researchers must keep that context in mind at every step.

This means:

- Knowing where the data came from, and
- Knowing what limitations the study data could bring to a new study
- Designing the study in a way that handles those limitations and uses good science
- Sharing your findings carefully so people don't misunderstand or misuse them
- Accounting for any personal biases that might shape your research

## Sources and Further Reading

1. The Protection of Non-Subjects from Research Harm | HHS.gov. March 11, 2022. Accessed November 10, 2025.  
<https://www.hhs.gov/ohrp/sachrp-committee/recommendations/tabc-the-protection-of-non-subjects-from-research-harm.html>
2. Horgan J. Code rage: The “warrior gene” makes me mad! (Whether I have it or not). Scientific American. Accessed November 10, 2025.  
<https://www.scientificamerican.com/blog/cross-check/code-rage-the-warrior-gene-makes-me-mad-whether-i-have-it-or-not/>
3. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 2019;366(6464):447–453. doi:10.1126/science.aax2342
4. Swilley-Martinez ME, Coles SA, Miller VE, et al. “We adjusted for race”: now what? A systematic review of utilization and reporting of race in American Journal of Epidemiology and Epidemiology, 2020–2021. *Epidemiol Rev*. 2023;45(1):15–31. doi:10.1093/epirev/mxad010
5. Dhejne C, Lichtenstein P, Boman M, Johansson ALV, Långström N, Landén M. Long-Term Follow-Up of Transsexual Persons Undergoing Sex Reassignment Surgery: Cohort Study in Sweden. *PLOS ONE*. 2011;6(2):e16885. doi:10.1371/journal.pone.0016885
6. Reed E. Debunked: “The Swedish Study” Doesn’t Say What Anti-Trans Activists Claim It Says. July 25, 2024. Accessed November 10, 2025.  
<https://www.erininthemorning.com/p/debunked-the-swedish-study-doesnt>
7. Williams C. Fact check: study shows transition makes trans people suicidal. TransAdvocate. November 2, 2015. Accessed November 10, 2025.  
[https://www.transadvocate.com/fact-check-study-shows-transition-makes-trans-people-suicidal\\_n\\_15483.htm](https://www.transadvocate.com/fact-check-study-shows-transition-makes-trans-people-suicidal_n_15483.htm)
8. Genetic variants underlying human bisexual behavior are reproductively advantageous | Science Advances. Accessed November 10, 2025.  
<https://www.science.org/doi/10.1126/sciadv.adj6958>
9. Leffer L. Is Bisexuality Genetic? It’s More Complex Than Some Studies Imply. Scientific American. Accessed November 10, 2025.  
<https://www.scientificamerican.com/article/is-bisexuality-genetic-its-more-complex-than-some-studies-imply/>

10. Ventresca C, Martschenko DO, Wedow R, et al. The methodological and ethical concerns of genetic studies of same-sex sexual behavior. *Am J Hum Genet.* 2024;111(10):2107–2116. doi:10.1016/j.ajhg.2024.08.007
11. Winnett AV. Genetic Studies Can Perpetuate Stereotypes of Bisexual People. Undark Magazine. June 13, 2024. Accessed November 10, 2025. <https://undark.org/2024/06/13/opinion-genetic-studies-bisexual-stereotypes/>
12. Tatchell P. Don't fall for the myth that it's 50 years since we decriminalised homosexuality. *The Guardian.* <https://www.theguardian.com/commentisfree/2017/may/23/fifty-years-gay-liberation-uk-barely-four-1967-act>. May 23, 2017. Accessed November 10, 2025.
13. Bisexual Monogamy: Twice the Temptation but Half the Fun? – Vernalis – 1999 – Journal of Social Philosophy – Wiley Online Library. Accessed November 10, 2025. <https://onlinelibrary.wiley.com/doi/10.1111/0047-2786.00022>
14. Al-Sibai N. Controversial New Research Find That Bisexuals Are a Bunch of Rascals. Futurism. January 7, 2024. Accessed November 10, 2025. <https://futurism.com/neoscope/controversial-research-bisexual-genetics>
15. Elbein S. New study helps explain 'Darwinian paradox' of same-sex attraction. The Hill. January 3, 2024. Accessed November 10, 2025. <https://thehill.com/policy/equilibrium-sustainability/4387638-same-sex-attraction-genetic-factors-darwinian-paradox-new-study/>
16. Men with bisexual, risk-taking gene variants tend to have more kids: study. Yahoo News. January 6, 2024. Accessed November 10, 2025. <https://www.yahoo.com/news/men-bisexual-risk-taking-gene-010713090.html>
17. Bisexual behavior genetically tied to risk-taking, controversial DNA analysis finds. Accessed November 10, 2025. <https://www.science.org/content/article/bisexual-behavior-genetically-tied-risk-taking-controversial-dna-analysis-finds>
18. Benjamin KJM, Chen Q, Eagles NJ, et al. Analysis of gene expression in the postmortem brain of neurotypical Black Americans reveals contributions of genetic ancestry. *Nat Neurosci.* 2024;27(6):1064–1074. doi:10.1038/s41593-024-01636-0
19. Bridging the Gap In Neuroscience & African Ancestry. AANRI. November 22, 2023. Accessed November 10, 2025. <https://aanri.org/>
20. Black In Neuro. Black In Neuro. October 2, 2025. Accessed November 10, 2025. <https://blackinneuro.com>