

Exercise 0:

Introduction to CHIRON

This introduction is mostly reading-based. If you're working in a group, go through it together, either by taking turns reading aloud or reading silently at your own pace. Use the discussion questions as natural pause points to reflect and talk as a group.

What is Biorepository Research?

Biorepository research refers to the use of existing biological samples or data — such as DNA, blood, or health records — for research purposes that go beyond the original reason those samples or data were collected. This is often called secondary research.

Unlike primary research, which involves collecting new data directly from participants, secondary research uses existing materials. This makes it easier and faster to conduct — but also more removed from the people represented in the data.

Pause and discuss:

- What kinds of data do you or your institution work with?
- In your experience, how closely are researchers connected to the communities represented in their data?

What is Group Harm?

Even when researchers have good intentions, studies that rely on large datasets can unintentionally:

- Misrepresent certain communities
- Reinforce bias or stigma

- Lead to inequitable decisions in healthcare or policy

We call these types of effects **group harm**.

Group harm is different from the kind of harm we often think about in research — like a breach of privacy or a painful procedure that affects someone as an individual.

Instead, group harm refers to the negative consequences that arise because of your connection to a larger group, such as your race, ethnicity, gender identity, sexual orientation, disability status, where you live, the type of work you do, your socioeconomic background, or even patterns in your DNA. In some cases, the group itself may not exist until research instantiates it (like people with the BRCA gene mutation).

Examples

Let's look at a couple of examples of group harm.

1. **Summary of “Dissecting racial bias in an algorithm used to manage the health of populations,” by Obermeyer et. al.**

An algorithm is a set of rules that a computer follows, and it is “trained” on large amounts of information. There is lots of evidence that algorithms produce outcomes that are worse for women and people of color. However, it can be hard for researchers to investigate these algorithms because if they are made by a private company, they are kept secret.

In this study, the researchers were able to obtain both the algorithm and the data it was trained on. The algorithm that they obtained is used in large health systems in the US to predict which patients will need the most healthcare resources. When a patient is identified by this algorithm as “high risk,” the patient gets access to lots of helpful things, like teams of nurses and appointment slots.

The research team studied whether Black and white patients were being treated equally by this algorithm. What they found was that at the same “risk score” given by the algorithm, Black patients were more sick than white patients. In other words, Black patients had to have more illness in order to qualify for the same services as white patients.

The reason for this is that in order to determine who has the greatest health needs, the algorithm predicts who has the greatest health costs. But Black patients create less health system costs than white patients. Black patients may have more barriers to accessing health care, or they may not get as much health care due to discrimination from doctors.

Because health care costs are not a perfect substitute for health care needs, this algorithm helps white patients more than Black patients.

2. Summary of “The Pain Was Unbearable, so why did doctors turn her away?” from *Wired*.

The story begins with a woman named Kathryn who has endometriosis, which is a painful condition that causes uterine cells to grow in the wrong places.

Kathryn had been taking opioids for years to manage her pain. When she was hospitalized in extreme pain, she was accused of seeking opioids for the wrong reasons, and she later was told by her gynecologist that she couldn’t come to their office anymore due to her “risk score” in the “NarxCare database.”

NarxCare has algorithms that take into account the other medications a patient is using, medical conditions they have, and their criminal justice background to come up with a score that predicts how likely a patient is to abuse opioids. In Kathryn’s case, because she has older dogs that were prescribed opioids by their veterinarian, those medications were factored into Kathryn’s risk score, making it look like she had a lot of opioid prescriptions.

This risk score is only meant to be a suggestion for doctors, but because doctors are nervous about their patients abusing opioids, they seem to just turn away patients who have a high risk score. People who have received high risk scores feel like they are being discriminated against for things in their background, such as having mental health issues from sexual abuse or having chronic pain.

Some researchers have found that this algorithm is very bad at predicting who ends up abusing opioids. Because NarxCare is owned by a private company, researchers are not able to investigate the algorithm very well, and if patients like Kathryn complain, there isn't a lot they can do.

|| Pause and discuss:

- What went wrong in these examples?
- Where in the process could harm have been prevented?
- How would you explain group harm in your own words?

Who is Responsible for Preventing Group Harm?

As of now, group harm is not formally recognized in most research regulations. This means that secondary research (studies that reuse existing data) typically does not undergo the same ethics review process required for primary research.

Because there are no formal guardrails to prevent group harm, individuals must make the choice to acknowledge the risk and take steps to mitigate it.

People in key positions to mitigate group harm:

- **Secondary Researchers** – Many choices can unintentionally cause group harm. Researchers must be mindful of the decisions they make throughout the planning, analysis, and dissemination of their research.

- **Ethics Committees** – Ethics committees are not usually required to review secondary research, or to consider group harm. But these committees may deal with secondary research on occasion, or some choose to engage with these issues anyway.¹ Even if regulations don't require it, ethics boards can shape better norms and advocate for more inclusive practices.
- **Data Access Committees** – DACs are often the only gatekeepers researchers must go through to access repository data. Even if ethics isn't their official role, they can ask meaningful questions and set expectations for community consideration.

Using CHIRON in Your Role

CHIRON was created to help people in these roles take meaningful steps to reduce group harm in research. Here's how we imagine it being used:

- **Secondary Researchers:**
Use the exercises while actively working on a project, or make up a project to discuss. Each exercise will help you think through your decisions with regard to a certain topic, such as choosing a research question, lumping or splitting variables, sharing results, etc.
- **Ethics Committees:**
Use CHIRON as a resource to guide your review process. You might adapt questions or prompts for inclusion in your application materials or discussion checklists. Choose a real research submission to discuss, or make up an imaginary one.
- **Data Access Committees:**
CHIRON can support your role as a safeguard against potential harms. You might integrate it into your review workflow or use specific exercises to inform your deliberations. You might also adapt questions or prompts for inclusion in your

application materials or discussion checklists. Choose a real research submission to discuss, or make up an imaginary one.

A Note on Research Roles

We recognize that roles aren't always clear-cut. A "secondary researcher" may also have been involved in collecting the data, or have other relationships with the community it comes from. You might be both a data steward and a researcher—or a committee member who also conducts studies.

Use CHIRON from the standpoint of your *current* role, while keeping in mind the full complexity of your work.

II Pause and discuss:

- Which role(s) do you most identify with?
- Where do you see opportunities to address group harm in your own work?
- Where do you see challenges to addressing group harm in your work?

Key Terms and Definitions

Knowing these terms will come in handy while using CHIRON.

- **Biorepository:** Any collection of data that has been amassed for research purposes that includes:
 - tissue samples and/or data derived from those samples
 - **plus** biological **and/or** health data of any kind.

These repositories may or may not be extremely large. These repositories may or may not be publicly funded.

- **Biorepository research:** Any study that uses data from a biorepository. This research often employs computational techniques like machine learning and artificial intelligence.

- **Algorithmically-defined community:** Groups formed through computational analysis (e.g., “BRCA carriers” or “high-risk patients”) that may not have existed before the research.
- **Group harm:** Harm experienced because of one’s connection to a larger group, rather than their individual identity. Examples include stigma, exclusion, or policy impacts.
- **Community:** A group of people with a characteristic in common. Can include traditional identity-based groups (race, gender, etc.), but also new or fluid categories that emerge from research itself.

|| Pause and discuss:

Are any of these terms new to you, or defined differently than you’re used to?

Trustworthiness Calculator

Before moving forward, **go to chiron.synapse.org and complete the Trustworthiness Calculator.** This tool helps you reflect on your current practices and assumptions toward community-centered research. You’ll return to it after completing the exercises to see how your perspective has evolved.

|| Pause and discuss:

- How did it feel to answer these questions?
- What, if anything, surprised you about your score?

Write it down:

Record your score somewhere safe so you can revisit it during the wrap-up exercise.

Final Thoughts

CHIRON isn't here to give you "right answers," because those will be different for each situation. Instead, it's designed to help you think critically.

Sources

1. Doerr M, Meeder S. Big Health Data Research and Group Harm: The Scope of IRB Review. *Ethics Hum Res.* 2022;44(4):34–38. doi:10.1002/eahr.500130