

Using Intersectionality in Data Analysis: An Example from Violence Prevention

TOOLS FOR EPIDEMIOLOGISTS AND ANALYSTS

What is intersectionality?

Intersectionality refers to the ways multiple forms of oppression interact with one another to affect individuals and communities. This framework was developed by Black feminists like those in the Combahee River Collective, and the term itself was coined by American scholar and civil rights lawyer Prof. Kimberlé Crenshaw, J.D., L.L.M.^{i,ii,iii} Each aspect of a person's identity, like their race, sex, gender, sexual orientation, class, disability, and age, has a different effect on how that person is treated in their daily life by the people, systems, and structures around them, which can have a powerful impact on health. When a person belongs to more than one group experiencing discrimination or inequity, this can impact their health in distinct ways by multiplying or compounding the amount of inequality they face. The effects of overlapping identities on health outcomes are often unique and can be greater than one would expect looking at each identity on its own.

Intersectionality is about overlap, but it is also about difference. Instead of shying away from difference, or controlling for it statistically, intersectionality challenges epidemiologists to recognize and learn from differences. An intersectional look at the data can reveal how people's identities, histories, and lived experiences influence health outcomes, public health programs, and how societal conditions affect them. More can be learned in seeking to recognize and understand these differences than from taking a universal but nonspecific approach. The framework of intersectionality has a long history of being applied to qualitative research in public health and is beginning to be recognized as a valuable lens for quantitative research in public health.

Why is it important to consider intersectionality in an analysis?

An epidemiologist's job includes documenting health inequities. However, critical nuances are missed in the data when an analysis looks at each individual variable in isolation from the other variables. Because communities are not a monolith, and no person is completely defined by just one category that they fit into, it's important to acknowledge differences in experience *within* categories, not just between them. For example, single axis disparities in prevalence data for intimate partner sexual violence (IPSV) in the 2019 Minnesota Student Survey (MSS) show:

- Students assigned female at birth (AFAB) are 3.5 times as likely to report IPSV as students assigned male at birth (AMAB) (11.6% vs. 3.3%).
- When combined, American Indian or Alaska Native, Asian or Asian American, Black, African or African American, Hispanic or Latino Middle Eastern or North African, and Native Hawaiian or Pacific Islander students report IPSV at similar rates (7.4% vs 7.7%) to white students.

- Lesbian, gay, bisexual, queer, pansexual, or questioning (LGBQ+) students are 2.9 times as likely to report IPSV as heterosexual students (17.9% vs. 6.2%). Also, transgender, genderqueer, or genderfluid students are 2.8 times as likely to report IPSV as cisgender students. (Cisgender is a term that means one's gender identity is the same as the sex that they were assigned at birth).

With these findings, one might mistakenly make the following incorrect interpretations:

- No students assigned male at birth are at elevated risk for IPSV.
- Differences in IPSV risk between students of color and white students is not a substantial inequity.
- All LGBQ+ and transgender, genderqueer, or genderfluid students are similarly affected by the large inequities in IPSV compared to heterosexual and cisgender students.

MDH's intersectional analysis challenges these interpretations by finding that:^{iv}

- Students who are assigned male at birth face elevated prevalence of IPSV victimization if they are transgender, unsure about their gender identity, or LGBQ+ (8.5-16.7%). This finding does not contradict conclusions about the role of sexism in IPSV distribution that can be drawn from the greater disparity between AMAB and AFAB students. Rather, it adds context about how sexism operates jointly with other forms of oppression like transphobia to shape different realities for different members of one identity.
- In certain contexts, race and ethnicity present very large disparities. For example, among students who are assigned male at birth that are LGBQ+ and transgender, genderqueer, genderfluid, or unsure of their gender identity, there is a greater than three-fold difference in IPSV risk between white students and students of color (9.8% vs. 32.0%). This finding illustrates two points: Firstly, that non-white students are not a monolith and that grouping various racial and ethnic identities together for ease of analysis might mask more nuanced disparities. Secondly, the fact that the eCHAID model did determine that grouping students from non-white racial and ethnic groups together in the specific context of AMAB students facing homophobia and transphobia uncovered the most statistically significant disparity, reveals how race and ethnicity, particularly whiteness, can interact in a unique way with sexual orientation and gender modality.

These findings would not be revealed without an analysis plan that can anticipate, recognize, and highlight instances of variation within categories. The complexity of the model is more appropriate to the complexity of real life. The ability of the model to highlight multiple disparities interacting in unique ways mimics the way that each person is affected by many identities and systems interacting in unique ways.

How can intersectionality be used in quantitative methods?

Experts have proposed many ways to model intersectionality using quantitative data, including multilevel models, decision trees, and descriptive cross-classification. While there is not yet one agreed-upon best way to do an intersectionality analysis, most experts agree that an ideal process goes beyond including an interaction term in the regression model. The authors of the

recent MDH study discussed above used exhaustive Chi-square automatic interaction detection (eCHAID), a supervised machine learning method.^v This method takes a large, diverse sample and repeatedly divides observations into smaller groups until groups with more uniform risk for the health outcome of interest are identified.

The resources section below links to many sources that discuss methodological options for intersectionality analysis in depth, including the strengths and limitations of each. Though some specific methods like eCHAID require large sample sizes and advanced statistical software, the lens of intersectionality can be applied to an analysis of any size and nature with enough deliberation and thoughtfulness. Whatever method is chosen, epidemiologists are encouraged to apply the following best practices:^{vi}

- Include a clear definition of intersectionality and correctly cite one or more foundational authors from the field. To do this, it will be necessary to read, deeply engage with, and ground oneself in the literature on this topic, particularly Black feminist scholarship.
- Choose social position/identity variables that are meaningful to the outcome of interest and reflect real-world societal power imbalances. For example, adding disability status to a model is a decision that draws on the history, scholarship, and experiences of disabled people and communities overcoming structural inequities.
- Be clear about the limitations of the categories used, especially if variables are multi-dimensional or proxies for something else. For example, if your dataset only has sex assigned at birth, it's important to acknowledge that this variable is not interchangeable with gender identity and might not accurately reflect the experiences of intersex people.
- Review the literature or team up with an evaluator or community engagement specialist working on a given topic area to identify intersections of interest with qualitative data and community expertise. When possible, pair quantitative analyses with qualitative data collection or cite existing qualitative resources when reporting quantitative results.
- Include opportunities for community members to review and give feedback early and often in the data sharing process. Doing so is important not only because it allows community members to access data to inform their work and exercise decision-making power about their own data. It also allows community members affected by health inequities to describe in their own terms how these disparities are created and experienced on the ground and note findings that surprised them or do not line up with their experiences. Differences between quantitative results and community wisdom can be opportunities to dig deeper into the data and find out whether any nuances are being missed, or whether the community you're working with is not fully representative of the community in your data, or vice versa.
- Continuously evaluate tools and procedures for data collection and analysis to ensure that no community or identity is being systematically excluded or made invisible. This includes not only providing options for individuals to describe themselves as a particular identity but ensuring that the language used is inclusive and reflective of how communities describe themselves and considering usefulness and interpretability carefully before using analytical techniques that group more than one identity together.

Resources

- [Incorporating intersectionality theory into population health research methodology: Challenges and the potential to advance health equity - ScienceDirect](#)
- [The intersectionality-based policy analysis framework: demonstrating utility through application to the pre-vaccine U.S. COVID-19 policy response - PubMed \(nih.gov\)](#)
- Recommended reading from the Intersectionality Training Institute (ITI): [Reading Recommendations - Intersectionality Training](#)
- [Intersectionality primer - Women of Color Policy Network.pdf \(intergroupresources.com\)](#)

Minnesota Department of Health
Violence Prevention Programs Unit
625 Robert St. N.
PO Box 64975
St. Paul, MN 55164-0975
651-201-5000
health.violenceprev@state.mn.us
www.health.state.mn.us/communities/svp/index.html

05/14/2024

To obtain this information in a different format, call: 651-201-5400.

ⁱ Combahee River Collective (1977). 'A Black Feminist Statement' (pp. 210-218).

ⁱⁱ Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory, and antiracist politics. In K. T. Bartlett & R. Kennedy (Eds.), *Feminist legal theory: Readings in law and gender*. (pp. 57–80). Routledge.

ⁱⁱⁱ National Association of Independent Schools. (2018, June 22). *Kimberlé Crenshaw: What is intersectionality?* [Video]. <https://www.youtube.com/watch?v=ViDtnfQ9FHc>

^{iv} Cole, C., Raguette, M., Rider, G. N., & McMorris, B.J. (2024). Predictors of adolescent intimate partner sexual violence victimization: patterns of intersectional social positions in a statewide, school-based sample. *Journal of Interpersonal Violence*. 2024 Jan 16:8862605231221504. doi: 10.1177/08862605231221504. Epub ahead of print. PMID: 38229477.

^v Supervised machine learning refers to a computerized process that uses algorithms and statistical models to analyze patterns in a dataset without step-by-step instructions.

^{vi} Bauer, G. R., Churchill, S. M., Mahendran, M., Walwyn, C., Lizotte, D., & Villa-Rueda, A. A. (2021). Intersectionality in quantitative research: A systematic review of its emergence and applications of theory and methods. *SSM—Population Health*, 14, 100798. <https://doi.org/10.1016/j.ssmph.2021.100798>