

## **Reconciling race and ethnicity data across multiple time points and datasets**

**AUTHORS:** Lisa P. Oakley\*, S. Marie Harvey, Jangho Yoon, Annie Larson, Jeff Luck

**TOPIC/TARGET AUDIENCE:** Public health professionals and researchers interested in collecting and analyzing race/ethnicity data in state-level datasets for summary statistics or disparity research.

**ABSTRACT:** Although race and ethnicity data are central to health disparities analyses, they are sometimes reported differently in different datasets and/or for the same individual over time. This presentation describes an innovative approach to classify complex race/ethnicity data from linked Oregon Birth Certificate and Medicaid Eligibility datasets into a single race/ethnicity category.

Our analysis included 271,896 birth certificates from 211,080 Oregon women between 2008 and 2012. The majority (93%) of birth certificates reported a single maternal race/ethnicity. Including Hispanic ethnicity, race/ethnicity proportions were calculated for women reporting multiple races on a single birth certificate and for women who reported different races across two or more birth certificates. A single race/ethnicity was assigned using six different methods based on the calculated proportions. Sensitivity analyses were used to compare prenatal care adequacy and birth outcomes across the different assignment methods. Because the classifications did not significantly differ on outcomes, the assignment of “least prominent racial group” was chosen to address underrepresentation of small racial groups, specifically American Indian/Alaska Native. We employed the same procedure described above to calculate race/ethnicity proportions within the Medicaid Eligibility dataset. Differences between race/ethnicity classifications across datasets, additional findings, and implications for disparities analyses will be discussed.

### **OBJECTIVE(S):**

- Explain how race and ethnicity were reported in Oregon Birth Certificate and Medicaid Eligibility datasets, over time
- Discuss how multiple races at a single time point, across time points, and across datasets were combined to create a single-category race variable
- Describe the six classification methods used to compare race assignment: more than 50% of a single race, more than 75% of a single race, most prominent racial group, least prominent racial group, most prominent non-White racial group, and an all-inclusive multiple race category
- Explain the importance and consequences in classifying race across time and datasets

### **PRIMARY CONTACT INFORMATION:**

Lisa P. Oakley, MPH, PhD  
Postdoc/Research Associate  
Oregon State University  
5417371387 | Lisa.Oakley@oregonstate.edu

### **CO-PRESENTER(S):**

S. Marie Harvey | Marie.Harvey@oregonstate.edu

Jangho Yoon | [Jangho.Yoon@oregonstate.edu](mailto:Jangho.Yoon@oregonstate.edu)  
Jeff Luck | [Jeff.Luck@oregonstate.edu](mailto:Jeff.Luck@oregonstate.edu)