ASPE ASSISTANT SECRETARY FOR PLANNING AND EVALUATION | OFFICE OF HEALTH POLICY

CONTRACTOR PROJECT REPORT
HP-2022-19

Imputation of Race and Ethnicity in Health Insurance Marketplace Enrollment Data, 2015 – 2022 Open Enrollment Periods

Prepared for
the Office of the Assistant Secretary for Planning and Evaluation (ASPE)
at the U.S. Department of Health & Human Services

by
RAND Health Care

June 2022

Office of the Assistant Secretary for Planning and Evaluation

The Assistant Secretary for Planning and Evaluation (ASPE) advises the Secretary of the U.S. Department of Health and Human Services (HHS) on policy development in health, disability, human services, data, and science; and provides advice and analysis on economic policy. ASPE leads special initiatives; coordinates the Department's evaluation, research, and demonstration activities; and manages cross-Department planning activities such as strategic planning, legislative planning, and review of regulations. Integral to this role, ASPE conducts research and evaluation studies; develops policy analyses; and estimates the cost and benefits of policy alternatives under consideration by the Department or Congress.

Office of Health Policy

The Office of Health Policy (HP) provides a cross-cutting policy perspective that bridges Departmental programs, public and private sector activities, and the research community, in order to develop, analyze, coordinate and provide leadership on health policy issues for the Secretary. HP carries out this mission by conducting policy, economic and budget analyses, assisting in the development and review of regulations, assisting in the development and formulation of budgets and legislation, and assisting in survey design efforts, as well as conducting and coordinating research, evaluation, and information dissemination on issues relating to health policy.

This research was funded by the U.S. Department of Health and Human Services Office of the Assistant Secretary for Planning and Evaluation under Contract Number HHSP233201500038I and carried out within the Payment, Cost, and Coverage Program in RAND Health Care. Please visit https://aspe.hhs.gov/topics/health-health-care/health-care-coverage-access for more information about ASPE research on health care coverage and access.

ASPE Executive Summary

Racial equity is a key priority of the Biden-Harris Administration. Executive Order 13985, "Advancing Racial Equity and Support for Underserved Communities Through the Federal Government," states the Biden-Harris Administration's commitment to "the consistent and systematic fair, just, and impartial treatment of all individuals, including individuals who belong to underserved communities that have been denied such treatment."[1] Measuring which communities are underserved and whether federal agencies are advancing racial equity requires accurate and accessible data on program participants.

Both state and federal Health Insurance Marketplaces collect many kinds of data that are both necessary for program operations and useful in assessing equity. For example, data on geographic location, needed to determine which plans are available to an applicant at what premiums, can also be used to identify urban and rural residents. Applicants for premium tax credits or cost-sharing reductions must submit information on family income. Self-reported information on race and ethnicity is collected, but responses to these questions are not required in order to obtain coverage or receive subsidies. As a result, data on race and ethnicity is often missing. Self-reported race and ethnicity is missing for 32.5 percent of HealthCare.gov Marketplace enrollees across the 2015 to 2022 Open Enrollment Periods, including 38.7 percent of enrollees during the 2022 Open Enrollment Period.

To address this data limitation, the Office of the Assistant Secretary for Planning and Evaluation (ASPE) contracted with RAND Health Care to develop methods for imputing race and ethnicity among people who selected Marketplace plans on HealthCare.gov but did not report their race or ethnicity, and to apply these methods to data from the 2015 to 2022 Open Enrollment Periods. The imputation approach involves two basic steps. For individuals that did not report race and ethnicity during an Open Enrollment Period in one year but did so in another year, race and ethnicity information was applied from the closest year for which it was reported. This step reduced the 2015-2022 percentage of individuals with no self-reported information from 32.5 percent to 23.5 percent.

The remaining individuals have no self-reported race and ethnicity in any Marketplace data from any year. RAND's modified Bayesian Improved First Name Surname and Geocoding (BIFSG) method uses data on first name, last name, and address to impute the probabilities that each individual belongs to six mutually-exclusive groups: American Indian or Alaska Native, Asian and Pacific Islander, Black, Hispanic, Multiracial, and White. Information on racial and ethnic patterns by last name and by address comes from Census data, while information on racial and ethnic patterns by first name comes from separate mortgage application data. The probabilities of belonging to each group sum to 100 percent for each individual.

[1] Executive Order 13985, "Advancing Racial Equity and Support for Underserved Communities Through the Federal Government" Available at: https://www.govinfo.gov/content/pkg/FR-2021-01-25/pdf/2021-01753.pdf

Individual-level probabilities computed using BIFSG were summed along with the self-reported information to estimate the proportions of Marketplace plan selections from each group, overall and by year, age, and Census Division. In other words, this approach does not assign a "definitive" race or ethnicity to individuals who did not report it, but instead calculates the probability that – based on their name and address – that they are in particular racial or ethnic groups. Comparison of enrollees who provided race and ethnicity information with those for whom these variables were imputed for all years combined, 2015-2022, suggest that self-reporters were more likely to be White (59.5 percent of self-reporters versus 49.3 percent of imputed enrollees) or Asian American, Native Hawaiian, and Pacific Islander (9.4 percent of self-reporters versus 6.7 percent of imputed enrollees). Self-reporters were less likely than imputed enrollees to be Black (10.9 percent of self-reporters versus 15.7 percent of imputed enrollees) or Hispanic (17.9 percent of self-reporters versus 26.1 percent of imputed enrollees).

Including both self-reporting and imputed enrollees, the estimated distribution of race and ethnicity of 2022 Open Enrollment Period HealthCare.gov plan selections is 0.7 percent American Indian or Alaska Native; 8.5 percent Asian American, Native Hawaiian, and Pacific Islander; 12.7 percent Black; 25.3 percent Hispanic; 1.9 percent Multiracial; and 51.0 percent White. Analysis of estimates using the modified BIFSG algorithm for individuals who did report race and ethnicity suggests that the imputation method performs best with regard to enrollees who are Asian American, Native Hawaiian, and Pacific Islander; Black; Hispanic; or White; but was less reliable for American Indian and Alaska Natives, as well as for respondents selecting multiple races. In addition, it is important to note that while imputation offers additional insights, it does not replace the need for improved data collection on race and ethnicity.

Overall, the results suggest that self-reported race and ethnicity data may lead to an underreporting of the share of Marketplace enrollees who are Black or Latino, and that imputation can provide useful additional information to support efforts to improve equitable health coverage in the U.S.