

Concurrent Session A2–5

**Variation in Surname Analysis and Geocoding in the Identification
of Hispanic Members in an HMO**

Debra P. Ritzwoller, PhD, Clinical Research Unit, Kaiser Permanente Colorado;
Ann Sukhanova, MA, Clinical Research Unit, Kaiser Permanente Colorado;
Bridget Gaglio, MPH, Clinical Research Unit, Kaiser Permanente Colorado;
Fabio Almeida, MSW, Clinical Research Unit, Kaiser Permanente Colorado;
Diego Osuna, MD, Clinical Research Unit, Kaiser Permanente Colorado;
Russell E. Glasgow, PhD, Clinical Research Unit, Kaiser Permanente Colorado

Background: In order to address racial and ethnic disparities in both health care and in the delivery and receipt of behavioral health interventions, researchers and policy makers need information regarding the racial and ethnic mix of their target populations. In the absence of self-report data, research suggests that combined surname analysis and geocoding may provide accurate and efficient means of inferring race/ethnicity across health plan membership.

Objective: The purpose of this study is to use the 'gold standard' of self-report to examine the variation in surname analysis, geocoding, and the combination of surname and geocoding in the identification of Hispanic or Latina members of an HMO, who were recruited for two behavioral health interventions.

Methods: We collected self-identified race/ethnicity data from telephone based recruitment data derived from two behavioral interventions; a tailored smoking reduction project, and a lifestyle behavior intervention, named Viva Bien, which targeted postmenopausal Latina women with type 2 diabetes at risk for CHD. We employed the Generally Useful Ethnic Search System (GUESS) along with geocoding to compare self-identification to surname and census data based proxies for Hispanic race/ethnicity. We computed sensitivity, specificity, positive predictive value, and negative predictive value for white non-Hispanics and white Hispanics.

Results: Only 593 of the 2,249 (26.3%) HMO members contacted regarding the smoking reduction study responded to the Hispanic self-identification question. Of those only 44 (8%) self-report as Hispanic. Of the 593 who responded to the Hispanic question, 46 (8%) map to GUESS surname software. Of the 44 subjects self identifying as Hispanics, 25 (56%) map to the GUESS surname software and only 15.5% (N=7) lived in a census tract with reporting a Hispanic population of > 50%. Of the 1,762 female HMO members who were contacted for the Viva Bien study, only 1,188 (67%), provided self identification of race/ethnicity. Only 178 (15%) self-identified as Latina/Hispanic, while 25% mapped to the GUESS system.

Conclusions: Contrary to recent publications, our findings suggest that the specificity of the GUESS system may vary in a more integrated region and that adding a census tract flag indicating that > 50% of the population is Hispanic does not always improve sensitivity or specificity of the GUESS system.