HSR Health Services Research

DOI: 10.1111/j.1475-6773.2011.01316.x
RESEARCH BRIEF

Can Claims-Based Data Be Used to Recruit Black and Hispanic Subjects into Clinical Trials?

Ana M. Palacio, Leonardo J. Tamariz, Claudia Uribe, Hua Li, Ellen J. Salkeld, Leslie Hazel-Fernandez, and Olveen Carrasquillo

Objective. Evaluate the accuracy of an algorithm at identifying ethnic minorities from administrative claims for enrollment into a clinical trial.

Data Sources/Study Setting. Claims data from a health benefits company.

Study Design. We compared results of a three-step algorithm to self-reported race/ethnicity.

Data Collection/Extraction Methods. Using the algorithm, we identified subjects with high probability of being minority and ascertained self-reported race/ethnicity.

Principal Findings. We identified 164 subjects as likely minority based on our algorithm. Of these, 94 completed the survey and 87 identified themselves as black or Hispanic. The positive predictive value of the algorithm was 93 percent (CI: 85–97).

Conclusions. Claims data can be used to efficiently identify minorities for participation in clinical trials.

Key Words. Racial/ethnic disparities, clinical trial design, administrative claims data

Since 1993, the National Institute of Health (NIH) has required that racial and ethnic minority groups be adequately represented in all NIH (2009) sponsored research. Increasing the participation of racial and ethnic minorities in clinical trials is recognized as a key strategy for the reduction of health disparities (Freedman et al. 1995; Institute of Medicine 1999; Corbie-Smith, Moody-Ayers, and Thrasher 2004). Yet the representation of minorities in clinical trials continues to be suboptimal in many areas of biomedical research (Swanson and Ward 1995; Hall 1999; Heiat, Gross, and Krumholz 2002; Sateren et al. 2002; Swanson and Bailar 2002; Sullivan et al. 2007; Ford et al. 2008; Williams et al. 2010).

Reasons for underrepresentation of minorities in clinical trials include lower socioeconomic status, lack of insurance, non-English language, disease

burden, and fear of harm (Corbie-Smith et al. 2003; Hussain-Gambles, Atkin, and Leese 2004; Ford et al. 2008). Investigators also commonly cite limited access to minority communities as an additional barrier (Williams and Corbie-Smith 2006).

The use of administrative claims data may be one potential approach to rapidly and efficiently identify, and recruit minorities into clinical trials. A major barrier to using such an approach has been incomplete ascertainment of race and ethnicity in private and public datasets (Lauderdale and Goldberg 1996; Arday et al. 2000; National Research Council 2004). However, over the past decade investigators have been able to validate several probabilistic approaches to assign race and ethnicity to large cohorts of patients in such claims data (Fremont et al. 2005; Fiscella and Fremont 2006; Elliott et al. 2008).

These existing methodologies include geocoding techniques and the use of Census-based surname lists. Further, the combination of both methodologies, particularly when using a Bayesian approach, has shown to improve the sensitivity and predictive value of either one alone at identifying minority groups (Fiscella and Fremont 2006; Wei et al. 2006; Elliott et al. 2008; Elliott 2009). To date, this approach has been primarily used in observational studies. We hypothesize this approach may also be a powerful tool for recruiting minorities into clinical trials. In this paper, we describe a pilot study that evaluates a three-step algorithm that incorporates these existing methodologies and the Medicare Race Code (MRC) to identify racial/ethnic minority subjects from a large health benefits carrier database. The study was performed in preparation for a NIH-funded randomized clinical trial consisting of a phone-based intervention for which we plan to recruit 250 black and 250 Hispanic subjects having recently undergone a percutaneous coronary intervention (PCIS).

Address correspondence to Ana M. Palacio, M.D., M.P.H., Division of General Medicine, Department of Medicine, The University of Miami Miller School of Medicine; The Miami VA Medical Center, 1120 NW 14th Street, Suite 967, Miami, FL 33136; e-mail: apalacio2@med.miami.edu. Leonardo J. Tamariz, M.D., M.P.H. is with the Division of General Medicine, Department of Medicine and the Miami VA Medical Center, Miami, FL. Olveen Carrasquillo, M.D., M.P.H., is with the Division of General Medicine, Department of Medicine. Hua Li, M.D., Ph.D., M.S., is with the Division of Biostatistics, Department of Epidemiology and Public Health. Ellen J. Salkeld, Ph.D., is with the Department of Family Medicine. All of the above are with the University of Miami Miller School of Medicine. Claudia Uribe, M.D., Ph.D., and Leslie Hazel-Fernandez, Ph.D., M.P.H., are with Competitive Health Analytics, Humana Inc., Miramar, FL.

METHODS

Study Design, Setting, and Participants

Subjects were identified from a large health benefits company (Humana) claims database of approximately 3.5 million members enrolled in a commercial health maintenance organization (HMO), preferred-provider organization (PPO), or Medicare plans. The database includes a member file containing demographic information for each member per encounter (age, gender, type of insurance, surname, and address); a medical file containing up to nine recorded International Classification of Diseases, ninth revision (ICD-9) codes per encounter; and a pharmacy file. From the database, we selected subjects older than 18 years of age who had an ICD-9 procedure claim for a drug eluting stent (36.07) or bare metal stent (36.06) between September 1 and October 31, 2009. We excluded subjects who died or disenrolled from the health plan, those living in states where neither of the two Humana study nurses held a nursing license, and self-insured employer groups for whom Humana only provides administrative services.

Assignment of Likely Race/Ethnicity

The subjects we identified as PCIS recipients were assigned a likely race/ethnicity according to a three-step sequential algorithm: (1) the 1990 Census Spanish Surname list (SSL), (2) the MRC, and (3) geocoding. Our objective was to identify minority subjects to then screen them for eligibility for enrollment into a clinical trial. As our goal was not to recruit a nationally representative sample, we decided to test this simpler methodology instead of using a Bayesian approach.

1990 Census Spanish Surname List. The surnames of all subjects identified as stent recipients were matched against the SSL (Word and Perkins 1996). This list classifies surnames into six categories according to the likelihood of being of Hispanic origin. We only assigned Hispanic ethnicity to those in the "Heavily Hispanic" category (i.e., >75 percent of householders with a surname in this classification are Hispanics).

Medicare Race Code. The MRC has a sensitivity of 97 percent for white persons and 95 percent for black persons, but <40 percent for Hispanic

Medicare beneficiaries (Arday et al. 2000). However, the positive predictive value (PPV) of the MRC is above 96 percent for both black and Hispanics persons (Arday et al. 2000). Thus, we used the MRC to identify blacks and any Hispanic subjects missed by the SSL in Step 1.

Geocoding. While racial integration is slowly increasing for blacks in major metropolitan areas of the United States (U.S. Census Bureau: The American Community Survey 2011), geographic racial segregation remains widespread and thus geocoding continues to be very accurate for the identification of black subjects (Glaeser and Vigdor 2001; Fremont et al. 2005). Though Hispanics tend to live in less segregated neighborhoods (Logan 2001), we also examined this approach to help identify any potential Hispanics who remained unclassified after the first two steps.

Using commercially available software (*Ingenix Geocoder 4.0* OptumInsight; Eden Prairie, MN, USA), we matched the geographical information of each subject (complete address and zip code) to census data to identify the demographic composition of the area in which each subject resided. We used the full address of a subject to match it to a census block group (Fremont et al. 2005). When we were unable to match at a census block level, we tried to match to the larger census tract level. In approximately 10 percent of cases, these matches were not possible and we matched at the more heterogeneous zip code level (Fiscella and Fremont 2006). We classified persons as black or Hispanic if they lived in the corresponding geographic unit where 75 percent or more of all persons belonged to that specific minority group. This threshold is more conservative than the 66 percent cutpoint used in other reports (Fremont et al. 2005; Fiscella and Fremont 2006).

Verification of Race/Ethnicity

We ascertained race/ethnicity information by self-report. Two bilingual (Spanish/English) Humana nurses made up to ten attempts to reach each subject during a 1-week period. For verification of race and ethnicity subjects we asked, "Are you of Hispanic or Latino origin?" and "Do you consider yourself black or African American, white or other race?"

Willingness to Participate in Clinical Trials

The survey also ascertained the willingness of each subject to participate in our proposed clinical trial. After a brief description of the study, subjects were asked, "If Humana launched this study to compare the effectiveness of two different types of services for their members who recently received a coronary stent, would you consider participating?

Statistical Analyses

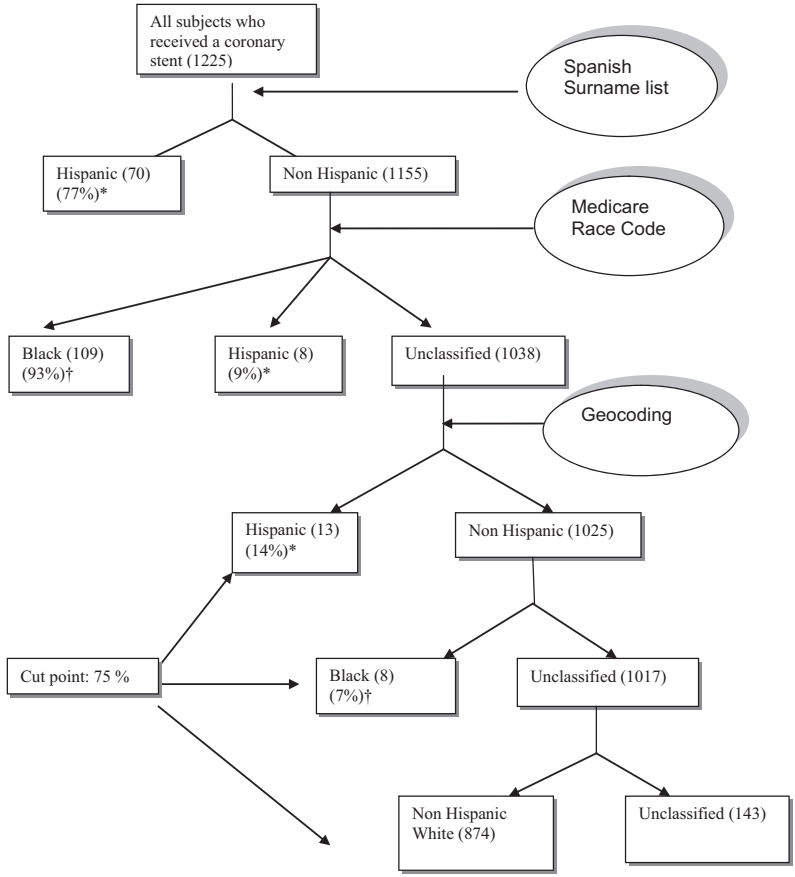
We report number and percentage of subjects who were reached and who agreed to participate in the survey, along with frequency distributions and the PPV of the algorithm to assign race/ethnicity. We measure the PPV of the algorithm at identifying a minority as well as the PPV by each race/ethnicity. In the calculation of overall accuracy of the algorithm, black Latinos were categorized as Hispanic to prevent double counting. However, for all race and ethnic specific analyses they were included in the black and Hispanic categories. Our main goal was to determine whether our approach could rapidly and efficiently enroll a large number of minority subjects. Had we wanted to determine other accuracy measures such as the sensitivity and specificity of our approach, we would have needed to sample a very large number of subjects to determine true and false negatives.

RESULTS

Baseline Characteristics

We identified 1,225 patients who had medical claim for PCIS during September and October of 2009 (Figure 1). The three-step algorithm identified 208 as likely minority subjects (black or Hispanic). These 208 subjects resided in 25 different states but 75 percent lived in one of the following five states: Florida (35 percent), Texas (17 percent), Illinois (6 percent), Louisiana (6 percent), and Virginia (6 percent). During a 1-week period, we were able to contact 99 of the 164 eligible subjects for a response rate of 60 percent. Of those, 94 agreed to participate in the survey for a participation rate of 95 percent. The baseline characteristics of this group are depicted in Table 1. Respondents were more likely to be black and most were Medicare beneficiaries. The average age was 71 ± 9 years and 46 percent of the reached subjects were female.

Figure 1: Three-Step Sequential Algorithm to Ascertain Likely Race/ Ethnicity



† Percentage of Black subjects identified by this step
* Percentage of Hispanic subjects identified by this step

Accuracy of the Algorithm

Of the 94 subjects who participated in the survey, 87 identified themselves as black or Hispanic; 63 (67 percent) were black, 27 (29 percent) Hispanic, and three were black Hispanics (Figure 1). Thus, the PPV of the algorithm to

Table 1: Baseline Characteristics of Subjects Participating in the Pilot Survey

Characteristic	
Number	94
Age (mean ± SD)	70.6 ± 8.9
Female (%)	46
Likely black (%)	72
Likely Hispanic (%)	28
Drug eluting stents (%)	76
Medicare (%)	96

identify individual racial/ethnic minority subjects was 93 percent (CI: 85–97). All seven subjects whom our algorithm misclassified were non-Hispanic white. By race, of the 68 subjects who were identified as likely black, 62 were black for a PPV of 91 percent (CI: 80–96). Similarly, the algorithm identified 26 subjects as likely to be Hispanics, of whom 24 self-reported to be Hispanic for a PPV of 92 percent (CI: 75–99).

Of the 94 respondents, 58 (62 percent) noted that they would be interested in participating in our proposed clinical trial. Twenty-four respondents (26 percent) stated they would likely not participate in the study and another 12 percent were not certain they would participate. By race and ethnicity, 30 percent of the Hispanics and 20 percent of blacks reported they did not think they would participate in the proposed clinical trial.

DISCUSSION

In a 1-week period, two part-time bilingual research nurses (one FTE) were able to identify 60 black and 27 Hispanic subjects who had recently received a PCIS. Our algorithm had a PPV above 90 percent for both minority groups. Of these, 58 were very willing to participate in our proposed clinical trial.

Hispanics have been the group least represented in clinical trials (Hall 1999; Heiat, Gross, and Krumholz 2002) and hardest to identify from traditional claims data such as Medicare. In our study, most of the Hispanics (76 percent) were identified using the surname list. Others also have reported similar success in using surname lists to increase the accuracy of the MRC for Hispanics with reported sensitivities of above 90 percent (Howard et al. 1983; Hazuda et al. 1986; Perez-Stable et al. 1995; Word and Perkins 1996;

Morgan, II, and Virnig 2004; McBean 2006). However, our study identified 10 percent of the minority subjects through the geocoding method, suggesting this step may still be of value in this group. This has been confirmed in other reports (Elliott 2009). An additional advantage of geocoding is the ability to select additional sociodemographic characteristics such as median income.

Among elderly blacks, the MRC has been shown to have a high sensitivity and high specificity. In our study, the MRC indentified most of our black subjects. Thus, studies that seek to only identify black elders may use this as the primary strategy. However, the reported accuracy of geocoding techniques to identify blacks is 89 percent when compared to MRC. Thus, geocoding alone may be effective at identifying blacks, especially younger ones lacking Medicare data (Fremont et al. 2005). Thus, while SSL may be used for identifying most Latinos and MRC for black elders, for studies that seek to sample both groups, and in particularly when younger and more diverse populations are desired, our three-step approach that also incorporates geocoding seems most suitable.

In our study, over 60 percent of those surveyed were willing to participate in our clinical trial. This finding is not unexpected. Although minorities may have concerns about research participation, reports have shown that once minority populations are appropriately identified and informed using culturally sensitive approaches, they are at least as willing to participate in clinical trials as nonminority groups (Katz et al. 2008).

Our approach has some important limitations. First, researchers interested in this recruitment methodology need timely access to claims data having protected health information. This may be a challenge with some public claims files. Private claims databases may be more feasible but researchers will need to have HIPPA-compliant collaborative agreements with such private companies. Second, our planned clinical trial does not require face-to-face visits and thus subjects were recruited throughout the United States. Studies that require in-person encounters could still use such an approach but may need to identify subjects in narrowly defined geographical areas. The applicability of the algorithm in such cases would depend on the density of racial/ethnic minority populations in a particular area and the ability to collaborate with organizations collecting data in those groups.

Third, while the price of geocoding software is decreasing, the cost of some products remains beyond the reach of individual investigators. However, many organizations that work with large claims databases such as insurers and state public health departments already use these analytic tools.

Lastly, using the above algorithm we may select a sample that is not fully representative of the all minorities living in the United States. For example, we would not capture subjects who do not have heavily Hispanic surnames or younger black who live in nonsegregated communities. However, the methodological need of having a fully representative sample is less pressing when recruiting for a clinical trial. Other approaches currently used by investigators to recruit minorities into clinical trials, such as targeted recruitment in health systems that disproportionately serve minority groups, would also generate nonrepresentative samples. Declining and differential response rates are also challenging. Indeed, this is a problem that plagues most approaches to collect nationally representative samples of minority groups (Blustein 1994; Bickart 1999; Zaslavsky 2002; Kressin 2003; Fremont et al. 2005; Fiscella and Fremont 2006).

In conclusion, the lack of race and ethnicity information on claims data has limited its applicability for minority health research. Using a simple three-step algorithm, we could accurately identify, from a claims database, a large number of minority subjects who were willing to participate in a clinical trial. The main strengths of this methodology were the relatively low costs and resources needed, making it an attractive strategy for trials that require significant minority representation. The use of claims data also allows investigators to identify minority populations who fulfill other relevant inclusion criteria, such as having had a major medical procedure or to measure outcomes such as health care utilization, clinical outcomes, and medication adherence. Ultimately, the 2010 Patient Protection and Affordability Act (Andrulis et al. 2011) and the requirements for enhanced reimbursement under Medicare for meaningful use of electronic health records (Blumenthal and Tavenner 2010) will require all administrative data systems to collect data on race and ethnicity. Until these federal mandates are fully implemented, we propose this as a simple and feasible approach to identify and recruit minorities into clinical trials using medical claim databases.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: The sole sponsor of the study is the National Institute on Minority Health and Health Disparities/NIH (grant RC1MD004327). Humana is a subcontractor on this award. There were no financial conflicts of interest for any of the co-authors of the manuscript. The Humana employees collaborating in the study are co-authors and provided

their personal feedback to the manuscript but had not contractual right to review or modify the content as Humana representatives. The University of Miami has complete independence to publish any result. The authors thank Sylvia Desiree Garay for her assistance in preparing this manuscript. The University of Miami has received grants from Humana to collaborate in other privately sponsored research projects unrelated to this manuscript and study. The policy of The University of Miami is to always have publication independence.

Disclosures: None.

Disclaimers: None.

REFERENCES

Andrulis, D. P., N. J. Siddiqui, J. Purtle, and L. Duchon. 2011. *Patient Protection and Affordability Act: Advancing Health Equity for Racially and Ethnically Diverse Populations.* Washington, DC: Joint Center for Political and Economic Studies.

Arday, S. L., D. R. Arday, S. Monroe, and J. Zhang. 2000. "HCFA's Racial and Ethnic Data: Current Accuracy and Recent Improvements." *Health Care Financing Review* 21 (4): 107–16.

Bickart, B. 1999. "The Distribution of Survey Contact and Participation in the United States: Constructing a Survey-Based Estimate." *Journal of Marketing Research* 36 (2): 286.

Blumenthal, D., and M. Tavenner. 2010. "The "Meaningful Use" Regulation for Electronic Health Records." *New England Journal of Medicine* 363 (6): 501–4.

Blustein, J. 1994. "The Reliability of Racial Classifications in Hospital Discharge Abstract Data." *American Journal of Public Health (1971)* 84 (6): 1018–21.

Corbie-Smith, G., S. Moody-Ayers, and A. D. Thrasher. 2004. "Closing the Circle Between Minority Inclusion in Research and Health Disparities." *Archives of Internal Medicine* 164 (13): 1362–4.

Corbie-Smith, G., C. M. Viscoli, W. N. Kernan, L. M. Brass, P. Sarrel, and R. I. Horwitz. 2003. "Influence of Race, Clinical, and Other Socio-Demographic Features on Trial Participation." *Journal of Clinical Epidemiology* 56 (4): 304–9.

Elliott, M. N. 2009. "Using the Census Bureau's Surname List to Improve Estimates of Race/Ethnicity and Associated Disparities." *Health Services and Outcomes Research Methodology* 9 (2): 69–83.

Elliott, M. N., A. Fremont, P. A. Morrison, P. Pantoja, and N. Lurie. 2008. "A New Method for Estimating Race/Ethnicity and Associated Disparities Where Administrative Records Lack Self-Reported Race/Ethnicity." *Health Services Research* 43 (5p 1): 1722–36.

Fiscella, K., and A. M. Fremont. 2006. "Use of Geocoding and Surname Analysis to Estimate Race and Ethnicity." *Health Services Research* 41 (4 Pt 1): 1482–500.

Ford, J. G., M. W. Howerton, G. Y. Lai, T. L. Gary, S. Bolen, M. C. Gibbons, J. Tilburt, C. Baffi, T. P. Tanpitukpongse, R. F. Wilson, N. R. Powe, and E. B. Bass. 2008. "Barriers to Recruiting Underrepresented Populations to Cancer Clinical Trials: A Systematic Review." *Cancer* 112 (2): 228–42.

Freedman, L. S., R. Simon, M. A. Foulkes, L. Friedman, N. L. Geller, D. J. Gordon, and R. Mowery. 1995. "Inclusion of Women and Minorities in Clinical Trials and the NIH Revitalization Act of 1993—The Perspective of NIH Clinical Trialists." *Controlled Clinical Trials* 16 (5): 277–85; discussion 286–9, 293–309.

Fremont, A. M., A. Bierman, S. L. Wickstrom, C. E. Bird, M. Shah, J. J. Escarce, T. Horstman, and T. Rector. 2005. "Use of Geocoding in Managed Care Settings to Identify Quality Disparities." *Health Affairs (Project Hope)* 24 (2): 516–26.

Glaeser, E. L., and J. L. Vigdor. 2001. *Racial Segregation in the 2000 Census: Promising News.* Washington, DC: Brookings Institute Center on Urban & Metropolitan Policy.

Hall, W. D. 1999. "Representation of Blacks, Women, and the Very Elderly (Aged > or = 80) in 28 Major Randomized Clinical Trials." *Ethnicity and Disease* 9 (3): 333–40.

Hazuda, H. P., P. J. Comeaux, M. P. Stern, S. M. Haffner, C. W. Eifler, and M. Rosenthal. 1986. "A Comparison of Three Indicators for Identifying Mexican Americans in Epidemiologic Research. Methodological Findings from the San Antonio Heart Study." *American Journal of Epidemiology* 123 (1): 96–112.

Heiat, A., C. P. Gross, and H. M. Krumholz. 2002. "Representation of the Elderly, Women, and Minorities in Heart Failure Clinical Trials." *Archives of Internal Medicine* 162 (15): 1682–8.

Howard, C. A., J. M. Samet, R. W. Buechley, S. D. Schrag, and C. R. Key. 1983. "Survey Research in New Mexico Hispanics: Some Methodological Issues." *American Journal of Epidemiology* 117 (1): 27–34.

Hussain-Gambles, M., K. Atkin, and B. Leese. 2004. "Why Ethnic Minority Groups are Under-Represented in Clinical Trials: A Review of the Literature." *Health & Social Care in the Community* 12 (5): 382–8.

Institute of Medicine. 1999. *The Unequal Burden of Cancer: An Assessment of NIH Research and Programs for Ethnic Minorities and the Medically Underserved.* Washington, DC: National Academy Press.

Katz, R. V., S. S. Kegeles, N. R. Kressin, B. L. Green, S. A. James, M. Q. Wang, S. L. Russell, and C. Claudio. 2008. "Awareness of the Tuskegee Syphilis Study and the US Presidential Apology and Their Influence on Minority Participation in Biomedical Research." *American Journal of Public Health* 98 (6): 1137–42.

Kressin, N. R. 2003. "Agreement between Administrative Data and Patients' Self-Reports of Race/Ethnicity." *American Journal of Public Health (1971)* 93 (10): 1734–9.

Lauderdale, D. S., and J. Goldberg. 1996. "The Expanded Racial and Ethnic Codes in the Medicare Data Files: Their Completeness of Coverage and Accuracy." *American Journal of Public Health* 86 (5): 712–6.

Logan, J. 2001. *Ethnic Diversity Grows, Neighborhood Integration Lags Behind.* Albany, NY: Lewis Mumford Center, University at Albany.

McBean, A. M. 2006. "Improving Medicare's Data on Race and Ethnicity." *Medicare Brief* 15: 1–7.

Morgan, R. O., W. II, and B. A. Virnig. 2004. "Improving Identification of Hispanic Males in Medicare: Use of Surname Matching." *Medical Care* 42 (8): 810–6.

National Institute of Health. 2009. *Biennial Report to the Director, 2009: Monitoring Adherence to the NIH Policy on the Inclusion of Women and Minorities as Subjects in Clinical Research. Comprehensive Report: Tracking of Human Subjects Research As Reported in Fiscal Year 2007 and Fiscal Year 2008* [accessed on August 8, 2011]. Available at: http://report.nih.gov/biennialreport/ViewSection.aspx?sid=67&cid=5.

National Research Council. 2004. "Eliminating Health Disparities: Measurement and Data Needs." In *Committee on National Statistics, Division of Behavioral and Social Sciences and Education*, edited by the Panel on DHHS Collection of Race and Ethnic Data Michele Ver Ploeg, and Edward Perrin, pp. 41–61. Washington, DC: National Academies Press.

Perez-Stable, E. J., R. A. Hiatt, F. Sabogal, and R. Otero-Sabogal. 1995. "Use of Spanish Surnames to Identify Latinos: Comparison to Self-Identification." *Journal of the National Cancer Institute. Monographs* 18: 11–5.

Sateren, W. B., E. L. Trimble, J. Abrams, O. Brawley, N. Breen, L. Ford, M. McCabe, R. Kaplan, M. Smith, R. Ungerleider, and M. C. Christian. 2002. "How Sociodemographics, Presence of Oncology Specialists, and Hospital Cancer Programs Affect Accrual to Cancer Treatment Trials." *Journal of Clinical Oncology* 20 (8): 2109–17.

Sullivan, P. S., A. D. McNaghten, E. Begley, A. Hutchinson, and V. A. Cargill. 2007. "Enrollment of Racial/Ethnic Minorities and Women with HIV in Clinical Research Studies of HIV Medicines." *Journal of the National Medical Association* 99 (3): 242–50.

Swanson, G. M., and J. C. BailarIII. 2002. "Selection and Description of Cancer Clinical Trials Participants—Science or Happenstance?" *Cancer* 95 (5): 950–9.

Swanson, G. M., and A. J. Ward. 1995. "Recruiting Minorities into Clinical Trials: Toward a Participant-Friendly System." *Journal of the National Cancer Institute* 87 (23): 1747–59.

U.S. Census Bureau: The American Community Survey. 2011. "U.S. Census Bureau: The American Community Survey" [accessed on April 20, 2011]. Available at http://www.census.gov/acs/www/

Wei, I. I., B. A. Virnig, D. A. John, and R. O. Morgan. 2006. "Using a Spanish Surname Match to Improve Identification of Hispanic Women in Medicare Administrative Data." *Health Services Research* 41 (4 Pt 1): 1469–81.

Williams, I. C., and G. Corbie-Smith. 2006. "Investigator Beliefs and Reported Success in Recruiting Minority Participants." *Contemporary Clinical Trials* 27 (6): 580–6.

Williams, M., M. Powers, Y. G. Yun, and E. Foa. "Minority Participation in Randomized Controlled Trials for Obsessive-Compulsive Disorder." *Journal of Anxiety Disorders* 24 (2): 171–7.

Word, D. L., and R. C. Perkins. 1996. *Building a Spanish Surname List for the 1990s—A New Approach to an Old Problem.* Population Division, Technical Working Paper 13. Washington, DC: U.S. Bureau of the Census.

Zaslavsky, A. M.. 2002. "Factors Affecting Response Rates to the Consumer Assessment of Health Plans Study Survey." *Medical Care* 40 (6): 485.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix SA1: Author Matrix.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.