January 2015

Strategies For Recruiting Representative Samples Of Asian Americans, Native Hawaiians, And Pacific Islanders For Population-Based Studies

Celeste Wong
Yale University, celestew@gmail.com

STRATEGIES FOR RECRUITING REPRESENTATIVE SAMPLES OF ASIAN AMERICANS, NATIVE HAWAIIANS, AND PACIFIC ISLANDERS FOR POPULATION-BASED STUDIES

By

Celeste K. Wong

A Thesis Presented to

The Faculty of the Yale School of Public Health

Yale University

In Candidacy for the Degree of

Master of Public Health

2015

ACKNOWLEDGEMENTS

This thesis would not have been possible without the expertise and support of the many talented and knowledgeable people I was privileged to work with during this process. I am particularly grateful to my thesis advisors, Dr. Melinda L. Irwin and Dr. Scarlett L. Gomez. Thank you, Dr. Irwin, for your constant mentorship and guidance, not only throughout this project, but throughout my entire two years at Yale School of Public Health as well. You have made my time here an incredibly fulfilling and wonderful learning experience. Thank you, Dr. Gomez, for your mentorship during my summer internship and throughout this past year, and for the opportunity to work on the Asian American Community Health Initiative (CHI) Study at the Cancer Prevention Institute of California, getting me started on this project, and guiding me through to the end. Your leadership and commitment to the CHI Study has made this project a success and a reality. I would also like to extend my gratitude to all the other dedicated and passionate individuals on the CHI Study team. Your expertise, advice, and support have been invaluable. Finally, I would like to thank my family and friends for the continuous encouragement, love and support during my graduate program. While this process was certainly challenging, it has truly been one of the most rewarding learning experiences I have ever had.

CONTENTS

LIST OF TABLES

ABSTRACT

There is limited data on effective methods for recruiting ethnically and culturally diverse populations into population-based studies. For case-control studies in particular, appropriate selection and successful recruitment of representative control subjects remain a challenge. In a population-based case-control study assessing novel risk factors for breast cancer among Asian Americans, Native Hawaiians, and Pacific Islanders (AANHPIs), we utilized a unique combination of population-based sampling, community-based recruitment methods, and internet- and media-based approaches for recruiting controls who were frequency matched to cases identified through a population-based cancer registry in the San Francisco Bay Area. We characterized the populations drawn from each recruitment source by comparing controls on a number of socio-demographic and medical characteristics across recruitment methods. We also compared characteristics of controls, in aggregate, to the overall source population to assess representativeness. Participants from each recruitment source differed with respect to many characteristics. For example, internet-based controls were more educated, had higher income, and were more likely to be born in the US, while controls recruited from community health centers were less educated, had lower income, and had limited English speaking skills. The combined control sample (N=483), however, appeared to be largely representative of the underlying source population with regards to most of the socio-demographic and medical factors under study, including nativity, education, marital status, and body mass index. Our simultaneous use of multiple alternative recruitment methods was found to be a feasible and cost-effective approach for recruiting a representative control series of diverse AANHPIs for population-based studies. Larger studies and further assessment of multiple strategies for recruitment of representative samples in various populations is needed.

INTRODUCTION

Appropriate selection and successful recruitment of controls (free of the disease under study) who are representative of the population from which the cases arose are two important steps in the design and implementation of population-based case-control studies. While this type of study design can produce important scientific findings for rare diseases with relatively little time and resources compared to other study designs (e.g., cohort studies), it can also be more susceptible to certain biases. Selection bias in case-control studies may exist when the control group is not representative of the underlying source population; whereas an ideal control group reflects the background prevalence of an exposure (or exposures) among individuals who are free of, but at risk for, the disease under study (1, 2). Thus, in order to avoid selection bias and to maximize study validity, the selection of a control group must adhere to a few guidelines: controls should come from the same source population as the cases, and should be selected independently of their exposure to the factors under study (1, 3).

Unfortunately, recruitment of representative controls that is both cost-efficient and feasible has become an increasingly challenging issue as US cultural norms and regulations have changed over the last 20 years (4-7). Traditionally, methods of population-based control recruitment have included random-digit dialing (RDD) or telephone directories, neighborhood block walking, and random sampling from population registries, beneficiary files, or voter or Department of Motor Vehicles (DMV) lists (4). However, with increasing confidentiality concerns, access to most of these population registry files has become restricted. Other methods, such as RDD, have become less efficient over time given lower response rates and an increased likelihood of selection biases that may invalidate study results. For instance, the exponential increase in cellular phone use (as well as a decline in the number of households maintaining landlines) and other technological advances, such as answering machines and caller identification, have resulted in fewer numbers of participants enumerated and recruited via RDD methods (4). In addition, as population-based controls are often a geographically-defined source population, an increasing problem in the pervasive use of cell phones is that area codes no longer represent residential areas and therefore such methods do not guarantee a random sample of the geographic population of interest (8, 9). Absences from the home during standard recruitment hours, incomplete directories that do not always cover the entire target population, and

regulations in certain geographic areas which preclude access to previously available databases (e.g., DMV), have exacerbated the issue, making these traditional methods inefficient, expensive, and increasingly ineffective for recruiting representative controls specific to race/ethnicity, age, and geography (5, 9). Indeed, studies have shown that these traditional methods tend to result in recruitment of controls who under-represent the older-aged and lower socioeconomic status groups, leading to selection bias (9-11).

Despite the aforementioned issues, population-based case-control studies continue to be a significant approach to identifying associations between risk factors and disease outcomes. As recruitment of representative controls becomes progressively more problematic and challenging, especially of underserved and diverse populations, it is crucial that researchers explore alternative methods for identifying and recruiting control subjects. Successful recruitment may require challenging traditional random sampling methods to recruit representative controls in ethnically diverse populations (12). A variety of alternative methods have been used in the literature, including recruitment from community organizations that serve the target population, advertisement in newspapers and other traditional media, and increasingly, recruitment via the internet and social media (e.g. Facebook and Twitter). In their recruitment of minority controls, Cabral et al. (6) and Bandera et al. (8) both found that controls recruited through community-based approaches (e.g. churches, health fairs, and senior centers predominantly serving African Americans or Latinos) were more representative of the source population than controls recruited through standard population-based approaches, such as RDD methods. However, these methods can still present a significant challenge in terms of recruitment of controls that represent the population from which the cases arose and can vary substantially depending on the community sources used. To our knowledge, there are no data comparing the representativeness of controls across multiple different recruitment sources. Clearly, any single recruitment source may introduce sampling bias. But, plausibly, the simultaneous use of multiple recruitment sources may temper and even balance out the biases associated with a single source. For example, samples coming from community organizations may over-represent persons from low incomes, whereas samples from social media outlets may over-represent those with higher incomes. In theory, understanding the types of bias introduced by various sources allows for the development of a sampling plan that maximizes the participation rate and monetary resources while minimizing bias.

To begin to address the issue of increasing rates of breast cancer among Asian American, Native Hawaiian, and Pacific Islander (AANHPI) women, the Asian American Community Health Initiative (CHI) Study was established as a pilot case-control study whose primary objectives are to explore novel breast cancer risk factors, with an emphasis on immigration-, discrimination-, early-life-, and contextually-related exposures. A secondary aim of the study is focused on identifying effective new approaches for recruiting AANHPIs for future population-based case-control and cohort studies, with the premise that through a careful and balanced selection of multiple recruitment methods, it is feasible to efficiently achieve a representative, population-based control series of AANHPIs.

Here, we focus on our experience using a unique combination of control selection strategies for recruiting controls frequency-matched to breast cancer cases identified through a population-based cancer registry. To characterize the control populations drawn from each source, we compared participants across recruitment methods on a variety of socio-demographic characteristics and breast cancer risk factors. To assess representativeness, we compared their characteristics, by ethnicity, to the overall source population.

MATERIALS AND METHODS

Recruitment of controls

The CHI Study was set in the San Francisco Bay Area counties of San Francisco, Contra Costa, Alameda, San Mateo, and Santa Clara. We employed several different strategies to recruit population controls: 1) address directory-based mailing, and recruitment through the 2) Army of Women, 3) Craigslist, 4 & 5) two community health centers (CHCs), and 6) community and media outreach by partnering with an AANHPI-serving community policy organization with wide-ranging connections to community-based organizations and ethnic media throughout the San Francisco Bay Area, disseminating study information via email listservs and attending ethnic-specific health fairs and community events. Each approach utilized easy-to-read study brochures, in English, Chinese, or Tagalog, which were developed in collaboration with ethnic- and language-specific focus groups in which recruitment messages, design of materials, and motivations for participation in health research were discussed. Recruitment of controls was conducted between March 2013 and October 2014.

Address directory-based mailing

Address-based sampling (ABS) is a sampling method that involves directly mailing a random sample of addresses from an electronic database (13). Selected commercial companies have access to the US Postal Service's Delivery Sequence File, which contains a listing of all addresses to which mail is delivered. Prior studies have suggested that ABS is more cost-effective compared to traditional sampling methods such as RDD, with greater coverage of residential homes and higher response rates (14, 15). For our study, we used a vendor that matched Asian surname from a list we provided to generate a random sample of 3,000 residential addresses in our study catchment area. The surname was based on head of household and the list included address, telephone number, age, and gender for head of household, plus person 2, if available. The vendor uses various commercial databases to provide the name and phone appends (e.g. Targus, InfoUSA, and Experian). The listings were stratified according to the following AANHPI ethnicities: Asian Indian or Pakistani, Chinese, Filipino, Guamanian/Chamorro, Japanese, Korean, native Hawaiian, Samoan, Vietnamese, Other Asian/Asian not otherwise specified (NOS), and Other Pacific Islander/Pacific Islander NOS. To

account for population mobility, the 3,000 listings were provided in 3 waves throughout the recruitment period. To each identified (likely) AANHPI address, we mailed a 4x6 red envelope that enclosed a visually attractive letter and flyer describing our study and a response form asking interested participants to verify or provide their contact information so that eligibility could be determined. Upon receipt of the completed response form (or response via phone, email, or our website), an interviewer contacted the interested woman, answered questions, determined eligibility, and scheduled a telephone interview as appropriate. Study staff did not follow up on mailing that resulted in no returned response forms.

Online approaches

Army of Women

The Army of Women (AOW) is a volunteer-based registry consisting primarily of women and some men, with and without breast cancer, who are interested in participating in breast cancer research. Registered members receive email updates from the AOW announcing new opportunities to participate in research studies approved by the AOW steering committee. Those who are interested can sign-up through an e-mail link where a set of screening questions is completed, eligibility is confirmed, and their contact information is subsequently relayed to the researcher. Since AANHPIs are the most active internet users among all US racial/ethnic groups (16) and our research involved participation in a study of breast cancer, the AOW conferred us a great opportunity to recruit interested women through this process. We sent out two "e-blasts" in April and August 2013.

Craigslist

We posted periodic advertisements on Craigslist between June 2013 and March 2014 for the geographically relevant areas under the “ETC” heading in the Jobs section. Interested women were instructed to contact the research coordinator via phone, email, or the study website.

Community health centers

We collaborated with two geographically-based, AANHPI-serving community health centers (CHCs) to assist with recruitment of control participants from their patient populations and the local communities they serve. In brief, each CHC conducted recruitment efforts which included contacting individuals to assess study eligibility based on ethnicity- and age-matching criteria, describing the study and disseminating study brochures, and transmitting to our staff the contact information of those who were interested in participating and willing to be contacted.

Asian Health Services (AHS) is a CHC located in Alameda County that offers primary health care services to the low-income and medically underserved, including the immigrant and refugee Asian community. Because many AANHPI communities are not reached through traditional social services due to cultural and linguistic barriers, AHS created Patient Leadership Councils, organized by language group and consisting of 10-30 community members each, to train patient volunteers to actively engage in their respective communities. During recruitment, our AHS staff and Patient Leadership Council members reached out to patients during medical visits and distributed brochures at numerous health fairs and other community events.

Asian Americans for Community Involvement (AACI) is a large, community-based organization located in Santa Clara County which provides an array of human services, such as primary medical care and mental health services, educational and advocacy programs, and shelter for individuals and families in the AANHPI community. Similar to AHS recruitment efforts, AACI staff conducted outreach and distributed brochures at AACI health and wellness centers, local businesses such as restaurants and nail salons, and at other community events such as health conferences.

General community recruitment

We partnered with the Asian and Pacific Islander American Health Forum (APIAHF), a national health policy organization headquartered in San Francisco, CA. APIAHF staff reached out to their networks of community-based organizations, media, and other contacts to generate interest and assist in recruitment for this study. The APIAHF placed advertisements in Chinese and Filipino newspapers and radio, distributed flyers at places of social congregation, such as

health fairs, senior centers, community events and fundraisers, as well as emailed and sent flyers and notifications to Asian-serving listservs and community groups. We also sent sporadic Tweets and made occasional Facebook postings promoting the study. Interested women were instructed to contact the research coordinator via phone or the study website.

Based upon the age and race/ethnicity distribution of the breast cancer cases enrolled in our study, in addition to our assumptions of distributions of socioeconomic indicators among controls that were expected from each recruitment source, we established pre-determined numbers of recruits from each of the sources that would be needed to achieve our targets. For example, we assumed that we would primarily recruit low socioeconomic status (SES) women from the CHCs, while women who were US-born and of higher SES would more likely be recruited from internet sources.

Data collection

Data were collected through an approximately 1-hour telephone interview and an approximately half-hour self-administered, mailed survey, which were available in English, Chinese (Mandarin and Cantonese), or Tagalog. Materials in Chinese and Tagalog were translated and independently back-translated. As a token of appreciation, participants who completed the telephone survey received a $35 check and those who further completed the self-administered survey received a $15 check.

To ensure that the controls we recruited, collectively, were representative of the underlying "at-risk" population, we periodically compared the nativity and SES (education and income) distribution of recruited controls to data from the California Health Interview Survey (CHIS), the largest state health survey in the United States that provides statistics for all age groups on demographics and a number of important health conditions and health-related behaviors (17). As data from neither the 2010 US Census nor the American Community Survey were available for the specific AANHPI ethnic populations of interest for the target geographic area, CHIS data were used to provide statistics for the source population. Beginning about eight months into the recruitment process, we conducted these interim analyses on an every-other-

month basis to ensure that our controls consistently matched the SES and nativity distribution of the CHIS population, redirecting our recruitment efforts as needed.

California Health Interview Survey

The CHIS is a population-based telephone survey of randomly selected households in California representing the non-institutionalized population residing in California. To provide estimates for California's overall population and major racial/ethnic groups, and to produce valid estimates at the county level, CHIS employs a complex weighting process and a multi-stage sample design that includes both landline and cellular telephone samples, in addition to surname list samples (17). CHIS data are released biennially via a free, online tool called *Ask*CHIS (http://www.ask.chis.ucla.edu) that allows users to produce customized health statistics on specific subpopulations. CHIS data from the 2011-2012 survey were used in this analysis, except for data on parity and hormone therapy use, which were from the 2009 CHIS. We extracted weighted prevalence estimates for Chinese, Filipina, other AANHPI, and total AANHPI female population, aged 20 years and older, residing in the CHI study catchment area.

As part of participation in our case-control study, participants were asked about a range of socio-demographic, medical, and lifestyle factors, some of which overlapped with the data collected by CHIS. All participants completed the telephone survey, but not all had completed the self-administered survey. Thus, the variables we selected for inclusion in this study were only those captured in the telephone survey. To assess whether our recruited control group, collectively, was representative of the underlying source population, we compared survey responses to similar data collected by CHIS.

Data analysis

In order to determine the population(s) being reached by each recruitment approach, we examined differences and similarities with respect to socio-demographic characteristics and breast cancer risk factors of controls recruited by each recruitment source. We compared the distribution of each characteristic of interest between a given recruitment source and all other recruitment sources combined using Chi-square tests on distributions weighted to the age (20-39,

40-59, and ≥60 years) and ethnicity (Chinese, Filipina, other AANHPI) of the total control sample. Because the controls were sampled and matched to the cases based on these age and ethnic subgroups, we chose to use these same broad subgroups for our weighting process. The Rao-Scott correction was applied to the Chi-square tests to adjust for our multi-way sampling design and corresponding weighting procedures (18). To account for multiple comparisons and to control for family-wise errors, we also applied the Bonferroni correction to the two-sided alpha level of 0.05 by dividing by 112 (i.e., 16 variables times 7 recruitment sources) (19). Thus, *p*-values <0.00045 were considered statistically significant. All statistical analyses were conducted using the SAS statistical program, version 9.3 (SAS Institute, Cary, NC).

To assess our hypothesis that the combination of recruitment sources yielded a control group that is representative of the source population in our study catchment area, we compared controls recruited from all methods combined to data from the CHIS. *Ask*CHIS estimates are weighted to represent California's overall non-institutionalized population (17); thus, in order to ensure direct comparability between our control group and the CHIS sample, we weighted our control sample to the age distribution of the ethnic-specific population of females 20 years and older, residing in our study catchment counties. As Chinese and Filipina women represented the largest proportions of participants in both our control sample and the CHIS sample, and because there were too few numbers of recruits in each of the other ethnic subgroups, our comparisons are based on the Chinese, Filipina, other AANHPI, and total control sample. For household income, the categories used by CHIS were incomparable to those used in our survey; thus, to form the lowest level category (≤$30,000), we split the number of CHI control respondents in the "$25,000 through $35,000" category evenly into the prior and subsequent categories. Given the complex sampling design and weighting process applied to *Ask*CHIS data estimates, we were unable to statistically compare distributions of our control sample to that of the CHIS data. However, we calculated 95% confidence intervals for all estimates, which provide a general indication of statistical differences and similarities.

Because our surveys were only available in Chinese, Tagalog, and English, CHIS estimates for the other AANHPI subgroup were limited to those participants who had taken the CHIS survey in English in order to produce comparable estimates between CHI control data to CHIS data. As indicated in **Table 3**, CHIS estimates for the total control sample include Chinese and Filipina participants who took the survey in any language, and other AANHPI participants

who only took the survey in English. CHIS estimates for AANHPIs by English proficiency are not available via *Ask*CHIS, thus, to obtain weighted estimates for English-speaking AANHPIs (AANHPIs other than Chinese and Filipina) to match our CHI other AANHPI sample, we requested and received these estimates from the CHIS Data Access Center.

RESULTS

We recruited a total of 483 controls: 10.1% from the address directory-based mail approach (n=49), 13.0% from the Army of Women (n=63), 16.8% from Craigslist (n=81), and 20.1% from the two community health centers (n=39 and n=58 for AHS and AACI, respectively). An additional 23.8% of respondents identified one or more of our community organization efforts through APIAHF as the way that they heard about the study (n=115). For the remaining 16.2% of respondents (n=78), other specific or non-specific sources were identified; these women could not be classified into one of the six recruitment methods, although most were likely the result of our community-based recruitment or word of mouth (snowball sampling). It was not possible to know the true denominator to be able to calculate response rates for most of our methods. The address directory-based mail approach is the only method whose effectiveness can be easily quantified; we successfully recruited 49 participants from the 3,000 letters mailed for an effective participant response rate of 1.6%.

Controls recruited from each of the recruitment sources differed significantly with regards to most socio-demographic characteristics, but were similar in terms of reproductive and medical factors (**Table 1**). For example, controls recruited from internet sources (AOW and Craigslist) were more likely to be US-born (both, $p<0.0001$), proficient in spoken English (no respondents reported the lowest proficiency level), and more educated than all other recruitment groups ($p<0.0001$ for AOW and 0.0006 for Craigslist). AOW controls were also wealthier ($p<0.0001$), and Craigslist controls were more likely to be single and live alone ($p<0.0001$ and $p=0.0001$). On the other hand, community controls (particularly those from AHS) were more likely to be foreign-born (100% of AHS controls were foreign-born), less educated and less proficient in English, have lower income, and be covered by public health insurance plans such as Medi-Cal (all, $p<0.0001$ for AHS controls). Reproductive characteristics were generally consistent among recruitment groups, with some exceptions. Controls recruited from the address directory-based mail and the APIAHF approaches did not significantly differ from controls recruited from the other methods. With few exceptions (where Chi-square tests could not be computed), AACI controls were also similar to other groups. **Table 2** summarizes the differences in characteristics between controls recruited from each method.

The results of the representativeness analyses comparing controls to CHIS data are shown in **Table 3** for Chinese controls, Filipina controls, other AANHPI controls, and all controls, respectively. The combined control group appeared to mirror the California population of AANHPIs in our target area with respect to most socio-demographic and reproductive characteristics, including nativity, education, home ownership, marital status, and body mass index (BMI). Distributions for the highest household income level and being unemployed or not working were also similar. However, our control group differed from CHIS data with regards to a few of the other study variables. For example, study controls, in aggregate, tended to have slightly lower income, were less proficient in spoken English, and more likely to have public health insurance. Chinese and Filipina women in our study were representative of the ethnic-specific source population in terms of several factors (e.g. nativity, education, and marital status among Chinese controls, and usual source of care, household size and BMI among Filipinas), but they also differed with regards to other characteristics. For example, Chinese and Filipinas controls were less proficient in spoken English and less likely to own a home than the ethnic-specific target population. The other AANHPI control group was also similar to the other AANHPI CHIS sample with respect to most characteristics, such as household income, marital status, and home ownership, but tended to be less foreign-born, less likely to be working full time, and less likely to be uninsured.

DISCUSSION

In this population-based case-control study of breast cancer susceptibility among AANHPIs, we compared control subjects recruited from multiple different recruitment sources on a number of socio-demographic and medical characteristics. We also assessed representativeness by comparing our control sample, in aggregate, to the overall source population. Our study focused on recruitment of diverse ethnic groups within the AANHPI population, a traditionally hard-to-recruit population given language and cultural barriers (20-25). While we faced challenges meeting target goals for specific recruitment methods (i.e. address directory-based mailing), we were able to meet overall control recruitment goals in a timely and efficient manner. We found that controls recruited from each approach differed with regards to most of the variables under study, as we hypothesized, but the carefully selected combination of recruitment approaches yielded a control group that was largely representative of the underlying source population.

In general, recruitment of representative controls in population studies has become increasingly challenging in an era of changing cultural norms and regulations, technological advances, and privacy concerns. Although RDD was traditionally considered the "gold standard" of recruitment methods, changes in the ways people socialize and go about their daily routines have made traditional recruitment methods such as RDD inefficient and costly. RDD or random sampling from population registries poses an additional challenge to recruiting control subjects from across a wide range of demographic and socioeconomic distributions, which is especially important in studies involving ethnically diverse populations, as such methods tend to result in recruitment of participants who over-represent the younger-aged and higher socioeconomic status groups, resulting in selection bias (9-11).

Despite the vast literature documenting the problems associated with traditional methods used in population-based control recruitment, data assessing the effectiveness of alternative methods to recruit controls of ethnically diverse backgrounds, including AANHPIs, remain limited. Presumably, understanding the underlying distribution of socioeconomic factors and health behaviors of individuals from different recruitment sources can help with the development of a sampling plan that enhances participation while also minimizing potential selection bias.

We explored a variety of novel recruitment approaches, and to our knowledge, our study is the first to compare the representativeness of controls across multiple different recruitment sources.

Given that most of our methods were nonrandom, we sought to characterize the population reached by each recruitment method by comparing the distributions of a number of socio-demographic characteristics and breast cancer risk factors across recruitment methods and found differences in many of these factors, as was expected. For example, we found that online-based controls were more likely to be US-born and had higher SES status, while controls recruited through community-based organizations tended to be foreign-born and had lower SES status. Controls recruited via the APIAHF and address directory-based mail methods were similar to all other groups combined, as were AACI controls. We expected controls recruited from AACI to have a lower-income distribution like AHS controls, since AACI works predominantly with Asian immigrants and refugees. However, AACI is located in Santa Clara County, which actually has one of the highest median household incomes in the nation, approximately $90,000, and this may have contributed to the recruitment of controls with a more representative income distribution (26).

When we compared socio-demographic and medical characteristics of our control sample to data from CHIS, we found that the combination of recruitment sources produced a control group that was largely representative of the source population in our study catchment area. We conducted interim analyses throughout the recruitment process to ensure that our controls consistently matched the SES and nativity distribution of the CHIS population, and our results indicate that this approach was integral in helping us achieve representativeness. Our control sample had a slightly higher proportion of lower household income, and also differed from the source population with respect to a few other characteristics (e.g. usual source of care, health insurance, and English proficiency). One potential reason for the discrepancies is that the way the questions are asked and the response options may differ between our CHI and the CHIS surveys. For example, the CHI survey had five response options for English speaking proficiency, whereas the CHIS only had four. Because “OK” was not an option in CHIS, we combined CHI respondents who reported “OK” with those who reported “well” into a single response option. Additionally, the categories of household income used by CHIS were incomparable to those used in our survey, and although we attempted to rectify this dissimilarity by splitting responses evenly into the higher and lower income categories, residual differences

may have still resulted. Furthermore, most of our recruitment methods comprised nonrandom sampling, so it was expected that controls would differ across recruitment sources. Although our total control sample mirrored the source population with respect to most characteristics, it may well not be possible to achieve similarity for every characteristic included in our study.

Successes and challenges of recruitment strategies

Since AANHPIs represent the most active ethnic group of online users, we expected recruitment through our online approaches to be highly effective. We successfully achieved our target numbers in a timely matter, but we were unable to determine effectiveness rates for either the AOW or Craigslist, as it was not possible to know how many women saw our postings. The AOW was particularly effective due to the snowball effect (i.e. word of mouth), as several participants mentioned they had been referred by family, friends, and social media websites, such as Facebook and Twitter, to join the study via the AOW, which not only increased the number of potential participants for our study, but also increased the number of AANHPI members in the AOW. In our experience, we found that web-based methods were effective, inexpensive, and feasible strategies for identifying and recruiting AANHPIs, especially those who are US-born and with higher SES.

Although the overall response rate for the address directory-based mailing was only 1.6%, this rate is similar to another population-based study that used this technique to recruit cancer controls (13). The low response rate may have partially been a result of ineligible sampling units (e.g. addresses without an occupant), although this number is likely to have been small. In addition, we did not send duplicate mailings to individual addresses nor utilized a personalized approach, such as including personalized letters with the study brochure or flyer, which have previously been shown to increase response rates (14, 27). Despite the potential to reach large, representative target populations with relatively little staff effort, the low response rate and need to pay an external company to identify potential recruits make address directory-based mailing a less effective recruitment option, but perhaps might be useful as a supplement to sampling plans if resources permit.

Community-based recruitment through social gathering places and other social events, another example of snowball sampling, has traditionally been considered to result in selection

bias (28); however, it is considerably more efficient than traditional recruitment methods and can be a solution to the challenge of reaching lower SES subjects (6, 8). Our results confirmed this finding, as our collaboration with the two CHCs and their outreach to community leaders resulted in the increased participation of low SES women, many of whom were immigrants with limited English speaking skills. Again, effectiveness rates for community-based recruitment methods could not be quantified, as the true denominators were unknown. However, given that we were able to utilize existing resources and personnel through the community organizations we collaborated with, we found community-based approaches to be a cost-effective and feasible recruitment approach. The success of this approach depends greatly on establishing a strong relationship between researchers and the CHCs and the strong presence of the health center and its staff within the community.

The specific recruitment techniques employed by the APIAHF overlapped with some of our other major recruitment sources, such as distributing flyers at health fairs and other community events, and using the internet (email) to promote our study. The APIAHF also advertised our study in a number of media outlets, including bilingual newspapers and radio stations, which is a strategy previously shown to enhance societal support for minority women's participation in health research (29). We found that media outreach via ethnic language newspapers was considerably more effective for recruiting Chinese than Filipino or South Asian participants. In our experience, collaborating with an influential and multifaceted organization with far-reaching capacities such as the APIAHF was highly effective for recruiting a large, representative group of AANHPIs into our study.

The costs of recruiting controls from these alternative sources involve compensation to the community organizations and CHCs, fees to the Army of Women and for Craigslist postings, fees for purchasing ABS listings, and study staff time. Although we were not able to derive a per-control cost, we estimate that these costs are likely lower than those from traditional approaches such as RDD and neighborhood block walking.

Recruitment barriers

Despite notable attempts to recruit controls from a variety of sources, ultimately the success of effective recruitment approaches may depend largely on the willingness of potential

controls to participate in research studies, which is likely influenced by a number of factors, including prior experiences as study participants, beliefs about the disease under study, individual characteristics, and, particularly among minority populations, negative attitudes or mistrust toward researchers (21, 22, 25, 30-33). Personal factors that may affect recruitment include awareness (e.g. degree of understanding of the research process and project description), acceptance (e.g. having social/community support), and access to research participation (e.g. language barriers, transportation, and work/family responsibilities) (21, 31, 34). Due to the growing concerns relating to lower ethnic minority participation in population studies, researchers have suggested a number of strategies for successful recruitment of ethnically and culturally-diverse populations. Examples include partnering with local community organizations to build trust, employing and training bilingual/culturally-competent and personable recruitment staff, and snowball recruiting (12, 29, 35, 36). Other practical factors that may enhance participation include providing incentives for participation and the use of appealing, easy-to-read study brochures (4, 5, 36). We incorporated all of these strategies into our recruitment approach, which not only helped us overcome many of these traditional barriers to recruitment, but importantly, our unique approach allowed us to successfully and effectively achieve a representative group of controls.

Limitations

Our study has several limitations. For one, most of our recruitment methods were based on nonrandom sampling methods, which could potentially introduce selection biases. Since all women meeting eligibility criteria were invited to participate in our study, and criteria were unrelated to exposures of interest, we are confident that controls were enrolled independent of exposure status (one of the principles of appropriate control selection). We are also confident that our control sample was recruited from the same underlying population from which the cases arose (the other principle of appropriate control selection), given that our collective control group mirrored the source population with respect to many characteristics, which included both socio-demographics and breast cancer risk factors. However, it would nonetheless be valuable to determine what biases, if any, are present by comparing case and control characteristics and assessing whether the addition of any one of the control recruitment sources significantly affects

risk estimates. Secondly, given resource limitations, our surveys were only available in Chinese, Tagalog, and English, which likely limited the number of potential participants in our study. Although Chinese and Tagalog are the two most commonly spoken AANHPI languages in the San Francisco Bay Area, there are dozens of other languages spoken, including Vietnamese, Hindi and Korean (35). For many minority populations, especially immigrant populations, the primary spoken language is not English. Language barriers are often cited as a primary reason for low levels of participation in cancer research (20, 31, 36). Future studies targeting AANHPI populations should provide qualified bilingual staff and develop bilingual materials in multiple languages, if resources permit. In addition, we did not distinguish between specific modes of recruitment used in our community-based efforts (e.g. newspaper, health fair, or word of mouth), which would be helpful for determining whether any one specific recruitment technique was more effective than others. A challenge faced by many researchers conducting studies among small and heterogeneous populations is recruiting sufficient numbers from each ethnic subgroup. To address this issue, we combined non-Chinese and non-Filipina AANHPIs into one group; however, this restricted our ability to capture nuanced differences and similarities between AANHPI subpopulations. As mentioned previously, we were unable to statistically compare distributions of our control sample to that of the CHIS data, but the weighted estimates and 95% confidence intervals provide a general indication of statistical differences and similarities. Finally, our recruitment methods may not translate into other AANHPI communities (e.g. in rural areas), and they certainly may not translate into other racial/ethnic groups, which limits the generalizability of our results.

CONCLUSION

Our results highlight the feasibility and effectiveness of simultaneously using various alternative approaches for recruiting AANHPI women into a case-control study assessing novel risk factors for breast cancer. The unique combination of community-based methods, internet or media-based strategies, and population sampling (i.e. address directory-based mailing), resulted in a control group of AANHPIs representing diverse socio-demographic backgrounds. Our methods not only facilitated the quick and efficient enrollment of participants who were more affluent and educated, it also allowed us to identify and recruit subjects who have traditionally been more difficult to recruit into health research (e.g. lower SES immigrants). Significantly, we were able to successfully yield a control sample that was representative of the source California population in our target area with respect to most socio-demographic characteristics.

In our experience, the active engagement of the AANHPI community contributed substantially to the success of our recruitment efforts. Community engagement and relationship building are essential for establishing and maintaining trust, particularly among the subsets of the population that have been historically disenfranchised or subject to discrimination, such as low SES immigrants or refugees with limited English speaking proficiency. We learned that face-to-face outreach was more successful than rote distribution of brochures or flyers. Community events where education and discussion between staff and potential participants is possible may enhance potential participants' understanding of the parallel between community goals and those of the researchers, ultimately resulting in increased participation (37, 38).

Based on results from our focus groups and what has been suggested in the literature, we employed a number of strategies to overcome recruitment barriers and enhance overall participation of AANHPI women, yet our experiences suggest that additional approaches may be helpful to optimize the recruitment of controls who are even more representative of the target population. Providing culturally-sensitive and bilingual recruitment staff and materials in multiple languages as well as distinguishing between different recruitment techniques may help improve recruitment efforts. In addition to assessing representativeness of controls to the source population and comparing case and control characteristics throughout the recruitment phase, it would also be beneficial for future studies to explore methods for quantifying the effectiveness of nonrandom sampling methods, and establish a tracking system that enables researchers to

systematically assess effectiveness of the various recruitment methods used. By capitalizing on established resources and networks with community organizations, we have shown that the simultaneous use of multiple different recruitment methods is a feasible and cost-effective approach for recruiting a representative sample of AANHPI controls from diverse demographic backgrounds into population-based studies.

Table 1. Age-adjusted[a] distribution of selected socio-demographic characteristics of controls (N=483) by recruitment method, females, Asian CHI Study, San Francisco Bay Area, 2013-2014[b]

	AACI (n=39)		AHS (n=58)		AOW (n=63)		Craigslist (n=81)		Mailing (n=49)		APIAHF (n=115)		Other (n=78)	
	%	*p*-value[c]	%	*p*-value[c]	%	*p*-value[c]	%	*p*-value[c]	%	*p*-value[c]	%	*p*-value[c]	%	*p*-value[c]
Age at interview (years)		0.3039		**0.0002**		0.2689		0.1245		0.2476		0.2432		0.0067
20-39	25.9		12.4		16.4		16.9		18.4		16.1		18.7	
40-59	52.2		62.3		61.4		60.7		56.6		62.0		71.7	
≥60	21.9		25.4		22.3		22.4		25.0		21.9		9.6	
Ethnic group		0.0057		N/A[d]		**0.0001**		0.0222		0.0015		0.0067		**<0.0001**
Chinese	82.4		71.5		48.9		49.0		58.5		50.6		43.7	
Filipino	6.0		28.5		20.8		16.3		9.4		20.4		23.7	
Other AANHPI	11.6		0.0		30.3		34.7		32.1		29.0		32.6	
Nativity		0.0132		N/A[d]		**<0.0001**		**<0.0001**		0.022		0.6558		**<0.0001**
US-born	13.2		0.0		65.6		62.3		25.6		31.7		22.4	
Foreign-born	86.8		100.0		34.4		37.7		74.4		68.3		77.6	
Annual household income		0.5561		**<0.0001**		**<0.0001**		0.0132		0.1332		0.6819		0.3268
≤$35,000	38.7		82.6		3.5		18.0		19.4		25.6		30.0	
$35,001 - $55,000	11.6		8.6		4.5		19.6		13.8		9.5		3.3	
$55,001 - $75,000	13.4		5.0		6.0		14.5		3.6		7.4		17.5	
$75,001 - $99,999	9.7		2.5		11.6		12.6		5.0		12.5		5.8	
≥$100,000	26.5		1.4		74.5		35.4		58.2		45.0		43.5	
Education level completed		0.0741		**<0.0001**		**<0.0001**		0.0006		0.2229		0.2589		0.0306
High school graduate or less	31.0		53.3		0.8		0.8		9.8		25.8		9.6	
Some college / vocational school / AA or AS degree	21.3		28.8		15.3		22.5		19.6		15.9		20.7	
College graduate or higher	47.7		17.9		84.0		76.7		70.7		58.3		69.7	
Employment status		N/A[d]		0.2991		0.5517		0.017		0.137		0.0077		0.0592
Full-time	27.5		40.8		48.6		45.3		47.4		49.2		32.1	
Part-time	28.7		18.5		21.5		31.3		10.6		19.1		38.5	
Unemployed / looking for work	0.0		8.6		10.9		9.7		10.7		1.2		9.5	
Retired	21.6		11.2		12.5		8.0		11.8		11.6		1.9	

Other[e]	22.3		21.0		6.5		5.8		19.6		18.9		18.0	
English proficiency[f]		0.2044		**<0.0001**		N/A[d]		N/A[d]		0.4653		0.4034		**0.0004**
Not well at all / poorly	35.3		46.9		0.0		0.0		29.0		21.4		10.3	
OK / well	53.3		50.0		9.3		26.8		49.6		55.2		67.3	
Very well	11.5		3.1		90.7		73.2		21.4		23.4		22.4	
Health insurance (multiple choices possible)														
Any public insurance	33.2	0.4542	73.0	**<0.0001**	10.5	0.0005	16.1	0.0208	32.1	0.6607	29.5	0.173	11.6	0.0388
Private insurance	67.0	0.5975	24.1	**<0.0001**	91.1	**<0.0001**	77.8	0.7173	81.3	0.1493	70.9	0.0914	85.1	0.311
No insurance	1.4	0.2154	7.1	0.3943	0.0	N/A[d]	5.6	0.1472	1.9	0.3875	1.2	0.1752	5.3	0.0209
Other[g]	13.1	0.1442	18.7	**<0.0001**	2.4	0.2808	4.1	0.8449	4.3	0.1633	3.1	0.3619	6.0	0.0734
Usual source of care		0.2826		**<0.0001**		N/A[d]		0.0006		0.2288		0.1208		0.2483
Doctor's office / Kaiser / HMO	86.0		32.5		87.2		69.2		84.2		76.6		78.7	
Clinic (hospital, community)	10.2		63.4		12.8		17.4		8.8		20.8		15.4	
Other[h], not one place, or none	3.8		4.2		0.0		13.5		7.0		2.7		5.9	
Marital status		0.7175		**<0.0001**		0.0769		**<0.0001**		0.2642		0.69		0.1335
Married / living with partner	72.4		70.2		76.8		43.2		72.1		65.6		73.3	
Divorced / separated / widowed	11.0		27.8		4.8		16.9		15.3		17.9		10.1	
Single / never married	16.6		20.6		18.4		39.9		12.6		16.6		16.7	
Household size		0.55		0.3985		0.334		**0.0001**		0.6006		0.3901		0.0016
1 person	8.8		5.7		15.7		33.3		8.9		15.2		8.5	
2-3 persons	53.6		51.1		58.6		47.8		58.4		44.6		47.1	
4 or more persons	37.6		43.3		25.7		18.9		32.7		40.3		44.4	
Home ownership		0.1623		**<0.0001**		0.0008		0.5604		0.0065		0.6687		0.8177
Yes	72.3		30.4		79.4		62.5		79.4		58.6		66.7	
No	27.7		69.6		20.6		37.6		20.6		41.4		33.3	
Body mass index (kg/m^2)		0.1627		0.8166		0.4262		0.2403		0.5757		0.073		0.6274
<25	82.2		65.7		57.0		60.7		71.4		62.8		78.1	
25-<30	15.9		25.0		36.4		29.4		24.9		18.0		16.1	
≥30	2.0		9.3		6.6		9.9		3.7		19.2		5.9	

Age at first birth (years)		N/A[d]		**<0.0001**	N/A[d]	**0.0001**	0.0631	0.3955	0.0078
Never given birth	27.1		8.2	32.1	53.6	24.5	24.9	20.0	
≤19	0.0		6.1	0.0	2.4	3.9	9.6	5.8	
20-29	51.5		58.9	38.9	25.9	25.4	35.4	33.9	
≥30	21.4		26.8	29.0	18.1	46.2	30.1	40.4	
Ever have menopausal hormonal therapy?		0.7308		0.7554	0.1943	0.2625	0.4334	0.6628	0.3277
Yes	9.1		6.4	8.0	7.9	11.0	11.1	3.5	
No	90.9		93.7	92.0	92.1	89.0	88.9	96.5	
Ever have a screening mammogram?		0.4217		0.2101	0.2897	0.6509	0.3868	0.6577	0.9291
Yes	73.6		80.1	86.3	80.3	79.8	82.4	82.3	
No	26.4		20.0	13.8	19.7	20.3	17.6	17.8	

Abbreviations: CHI, Community Health Initiative; AACI, Asian Americans for Community Involvement; AHS, Asian Health Services; AOW, Army of Women; APIAHF, Asian and Pacific Islander American Health Forum; AA, Associate of Arts; AS, Associate of Science

[a]Estimates were adjusted to the age distribution of the ethnic-specific population of all study controls combined (N=483)

[b]Table values are column percentages based on non-missing values only; percentages may not sum to 100% due to rounding

[c]Based on Chi-square test comparing controls from each recruitment method to controls from all other methods, weighted to the age and ethnicity distribution of all controls

[d]No test could be computed because one or more table cells had missing frequencies

[e]Responses may include on disability, homemaker, student, or volunteer

[f]Limited to participants who spoke another language at home other than English (N=267)

[g]Responses may include single-service plan (e.g. dental, vision, prescriptions), or insurance through another family member or organization

[h]Responses may include acupuncturist, websites, or self

Table 2. Summary of differences[a] in socio-demographic characteristics of controls (N=483) by recruitment method, females, Asian CHI Study, San Francisco Bay Area, 2013-2014

	AACI (n=39)	AHS (n=58)	AOW (n=63)	Craigslist (n=81)	Mailing (n=49)	APIAHF (n=115)	Other (n=78)
Age at interview		Less likely to be in the younger age category					
Ethnic group		All participants were either Chinese or Filipina[b]	Less likely to be Chinese				Less likely to be Chinese
Nativity		All participants were foreign-born[b]	More likely to be US-born	More likely to be US-born			More likely to be foreign-born
Annual household income		More likely to have low income	More likely to have high income				
Education level completed		More likely to have only completed high school or less	More likely to have completed college or higher				
Employment status							
English proficiency		More likely to have poor/limited English speaking skills	No participants reported speaking English poorly[b]	No participants reported speaking English poorly[b]			More likely to speak English "OK / well"
Health insurance		More likely to have public or other form of insurance; least likely to have private insurance	More likely to have private insurance; no participants reported being uninsured[b]				
Usual source of care		More likely to visit a clinic for care					
Marital status		More likely to have been previously married		More likely to be single			
Household size				More likely to live alone			
Home ownership		Less likely to own a home					
Body mass index							
Age at first birth		Less likely to have never given birth		More likely to have never given birth			
Ever have menopausal hormonal therapy?							
Ever have a screening mammogram?							

Abbreviations: CHI, Community Health Initiative; AACI, Asian Americans for Community Involvement; AHS, Asian Health Services; AOW, Army of Women; APIAHF, Asian and Pacific Islander American Health Forum; AA, Associate of Arts; AS, Associate of Science

[a]Based on *p*-values <0.00045 for Chi-square tests comparing controls from one method to all other methods combined, weighted to the age and ethnicity distribution of all controls combined

[b]No test could be computed because one or more table cells had missing frequencies

Table 3. Age-adjusted[a] distribution of selected socio-demographic characteristics of controls (N=483) by AANHPI ethnicity, compared to CHIS data, females, Asian CHI Study, San Francisco Bay Area, 2013-2014[b]

	% (95% CI)							
	Chinese		Filipinas		Other AANHPIs[c]		All participants[c]	
	Controls (N=243)	CHIS	Controls (N=94)	CHIS	Controls (N=146)	CHIS	Controls (N=483)	CHIS
Age at interview (years)								
20-39	9.2 (5.7, 12.6)		18.2 (10.6, 25.8)		36.3 (27.4, 45.2)		17.0 (13.5, 20.5)	
40-59	71.6 (66.4, 76.9)	N/A	65.6 (56.5, 74.7)	N/A	54.3 (45.5, 63.1)	N/A	67.3 (63.1, 71.6)	N/A
≥60	19.2 (14.8, 23.6)		16.2 (9.8, 22.6)		6.4 (3.4, 9.4)		15.7 (12.7, 18.6)	
Nativity								
US-born	28.3 (22.6, 34.1)	31.7 (23.4, 40)	21.6 (13.4, 29.8)	35.0 (19.5, 50.5)	58.1 (49.5, 66.7)	28.9 (21.2 (38.6)	34.8 (30.3, 39.3)	32.6 (27.0, 38.3)
Foreign-born	71.7 (65.9, 77.4)	68.3 (60, 76.6)	78.4 (70.2, 86.6)	65.0 (49.5, 80.5)	41.9 (33.3, 50.5)	70.1 (61.4, 78.9)	65.2 (60.7, 69.7)	67.4 (61.7, 73.0)
Annual household income								
<$30,000	30.0 (23.8, 36.2)	23.2 (16.8, 29.7)	23.6 (12.2, 35.0)	13.7 (2.4, 25.0)*	12.9 (6.7, 19.1)	13.2 (6.7, 19.8)	24.7 (20.2, 29.2)	18.0 (13.7, 22.3)
$30,000 - $999,999	31.4 (25.5, 37.4)	37.5 (29.6, 45.4)	39.5 (27.0, 52.1)	48.9 (32.7, 65.1)	36.1 (27.4, 44.8)	40.7 (32.2, 49.2)	33.4 (28.8, 38.0)	40.9 (35.1, 46.7)
≥$100,000	38.6 (32.0, 45.2)	39.3 (30.9, 47.7)	36.9 (24.6, 49.2)	37.4 (21.7, 53.2)	51.0 (42.2, 59.9)	46.1 (37.5, 54.6)	41.9 (36.9, 47.0)	41.2 (35.3, 47.0)
Education completed								
High school or less	22.6 (17.4, 27.8)	23.0 (16.8, 29.1)	14.0 (7.1, 20.9)	20.5 (7.9, 33.2)*	4.2 (0.8, 7.5)	8.9 (3.8, 14.0)	16.9 (13.4, 20.5)	17.6 (13.4, 21.8)
Some college / vocational school / AA or AS degree	17.6 (12.9, 22.4)	11.5 (6.3, 16.6)	23.5 (15.1, 31.9)	27.5 (11.7, 43.4)	21.5 (14.4, 28.5)	13.4 (7.8, 19.1)	19.3 (15.7, 22.9)	15.5 (10.8, 20.1)
College graduate or higher	59.7 (53.6, 65.9)	65.6 (58.2, 73)	62.6 (52.9, 72.2)	51.9 (35.7, 68.2)	74.4 (66.8, 81.9)	77.7 (70.5, 84.8)	63.7 (59.2, 68.2)	66.9 (61.4, 72.5)
Employment status								
Full-time	42.8 (36.5, 49.1)	57.5 (49.5, 65.5)	59.6 (49.9, 69.3)	49.0 (32.8, 65.2)	48.4 (39.6, 57.2)	66.1 (58.3, 73.9)	46.4 (41.7, 51.0)	58.4 (52.7, 64.1)
Part-time	20.9 (15.7, 26.1)	11.8 (6.9, 16.7)	21.3 (13.2, 29.3)	16.6 (4.8, 28.4)*	24.6 (17.2, 31.9)	7.5 (3.5, 11.6)	21.8 (18.0, 25.7)	11.5 (7.8, 15.2)
Not working[d]	36.3 (30.3, 42.3)	30.7 (23.7, 37.8)	19.1 (11.4, 26.9)	34.4 (19.6, 49.1)	27.0 (19.3, 34.8)	26.4 (19.2, 33.6)	31.8 (27.5, 36.1)	30.1 (25.0, 35.3)
English proficiency[e]								
Not well at all / poorly	37.2 (29.8, 44.6)	31.6 (24.0, 39.3)	10.1 (2.2, 18.0)	0.0	0.0	7.0 (1.5, 12.4)	28.2 (22.5, 33.9)	18.3 (14.0, 22.6)
OK / well	48.6 (40.9, 56.3)	40.5 (30.9, 50.2)	58.6 (45.5, 71.7)	27.2 (8.8, 45.7)*	36.7 (21.9, 51.4)	32.4 (22.7, 42.0)	48.0 (41.7, 54.3)	35.5 (28.7, 42.3)
Very well	14.2 (8.8, 19.6)	27.8 (19.3, 36.4)	31.3 (18.9, 43.7)	72.8 (54.3, 91.2)*	63.4 (48.6, 78.1)	60.7 (50.4, 71.0)	23.8 (18.5, 29.0)	46.2 (39.2, 53.3)

Health insurance (multiple choices possible)								
Any public insurance	30.1 (24.4, 35.7)	15.8 (10.8, 20.8)	24.3 (15.9, 32.7)	22.5 (9.6, 35.5)	12.9 (7.4, 18.4)	6.0, (3.0, 9.0)	25.1 (21.1, 29.0)	14.1 (10.4, 17.9)
Private insurance	69.6 (63.9, 75.9)	75.9 (69.5, 82.3)	65.7 (56.3, 75.1)	74.7 (61.3, 88.1)	86.1 (80.3, 92.0)	83.3 (77.1, 89.4)	73.2 (69.1, 77.2)	78.1 (73.6, 82.6)
No insurance	3.0 (0.8, 5.2)	8.3 (3.8, 12.7)	11.8 (5.4, 18.2)	2.8 (0, 7.4)*	2.4 (0.3, 5.2)	10.7 (5.4, 16.1)	4.0 (2.2, 5.8)	7.8 (5.0, 10.6)
Other[e]	7.6 (4.2, 10.9)	N/A	3.9 (0.1, 7.6)	N/A	5.6 (1.6, 9.5)	N/A	6.6 (4.2, 9.0)	N/A
Usual source of care								
Doctor's office / Kaiser / HMO	75.4 (69.9, 80.8)	62.3 (54.0, 70.5)	59.6 (49.8, 69.4)	55.5 (39.2, 71.7)	84.8 (78.5, 91.1)	63.7 (55.0, 72.4)	75.6 (71.7, 79.6)	61.3 (55.6, 67.0)
Clinic (hospital, community, neighborhood)	21.2 (16.1, 26.3)	23.2 (15.9, 30.4)	27.8 (18.9, 36.7)	31.3 (17.1, 45.5)	7.6 (2.7, 12.5)	23.2 (16.6, 29.9)	18.7 (15.1, 22.3)	24.9 (19.9, 29.9)
Other[f] / not one place / none	3.5 (1.1, 5.9)	14.6 (8.5, 20.7)	12.6 (6.1, 19.1)	13.2 (0, 27.1)*	7.6 (3.2, 12.0)	13.1 (6.1, 20.1)	5.7 (3.6, 7.7)	13.8 (9.4, 18.2)
Marital status								
Married / living with partner	69.2 (63.4, 75.0)	63.5 (55.4, 71.7)	64.0 (54.6, 73.5)	69.1 (54.9, 83.4)	68.2 (60.1, 76.2)	63.9 (55.4, 72.3)	68.3 (63.9, 72.6)	64.5 (58.9, 70.0)
Divorced / separated / widowed	13.0 (8.9, 17.0)	12.5 (8.3, 16.8)	17.7 (10.4, 25.0)	7.6 (0.5, 14.6)*	10.0 (5.0, 15.1)	9.6 (5.3, 13.9)	12.9 (9.9, 15.8)	10.8 (8.0, 13.6)
Single / never married	17.9 (12.9, 22.8)	23.9 (16.1, 31.8)	18.3 (10.6, 25.9)	23.3 (10.1, 36.5)	21.8 (14.6, 29.0)	26.6 (17.8, 35.3)	18.9 (15.2, 22.6)	24.8 (19.5, 30.0)
Household size								
1 person	15.8 (11.3, 20.2)	9.1 (5.3, 13.0)	6.7 (1.8, 11.6)	8.9 (1.7, 16.1)*	15.6 (9.4, 21.8)	10.2 (4.6, 15.8)	14.5 (11.3, 17.8)	9.61 (6.8, 12.4)
2-3 persons	56.1 (49.8, 62.4)	49.0 (40.7, 57.3)	39.3 (29.5, 49.0)	40.9 (25.2, 56.5)	44.7 (36.0, 53.3)	41.7 (33.3, 50.2)	51.1 (46.4, 55.8)	44.5 (38.7, 50.4)
4 or more persons	28.2 (22.4, 34.0)	41.9 (33.6, 50.1)	54.0 (44.1, 63.9)	50.2 (34.1, 66.4)	39.8 (31.0, 48.5)	48.0 (38.7, 57.3)	34.4 (29.9, 38.8)	45.9 (40.0, 51.7)
Home ownership								
Yes	65.5 (59.5, 71.4)	72.4 (65.2, 79.6)	49.1 (39.1, 59.0)	71.1 (57.9, 84.3)	67.8 (59.7, 76.0)	61.7 (52.7, 70.7)	63.9 (59.5, 68.4)	66.5 (63.3, 73.8)
No	34.5 (28.6, 40.5)	27.6 (20.4, 34.8)	50.9 (41.0, 60.9)	28.9 (15.7, 42.1)	32.2 (24.0, 40.3)	36.7 (27.8, 45.5)	36.1 (31.6, 40.5)	31.5 (26.2, 36.7)
Body mass index (kg/m^2)								
<25	76.7 (71.0, 82.3)	82.2 (76.4, 88.1)	51.9 (40.3, 63.5)	58.8 (42.6, 75.0)	58.5 (48.9, 68.0)	71.6 (63.3, 79.8)	69.6 (65.0, 74.2)	73.6 (68.2, 78.9)
25 - <30	21.0 (15.6, 26.4)	12.5 (8.2, 16.8)	35.5 (24.3, 46.7)	30.0 (14.1, 45.9)	23.1 (14.9, 31.2)	21.3 (14.1, 28.4)	23.1 (18.9, 27.3)	19.3 (14.5, 24.1)
≥30	2.4 (0.3, 4.4)	5.2 (0.8, 9.7)*	12.6 (5.1, 20.1)	11.2 (2.1, 20.3)*	18.5 (10.8, 26.1)	7.2 (2.9, 11.5)	7.3 (4.8, 9.8)	7.1 (4.0, 10.2)
Ever have a screening mammogram?[h]								
Yes	88.6 (84.6, 92.5)	66.9 (57.3, 76.5)	72.3 (63.4, 81.1)	81.3 (66.6, 95.9)*	61.6 (52.8, 70.5)	66.6 (57.9, 75.4)	79.8 (76.1, 83.5)	70.0 (63.8, 76.2)
No	11.4 (7.5, 15.4)	33.1 (23.5, 42.7)	27.7 (18.9, 36.6)	18.7 (4.1, 33.4)*	38.4 (29.5, 47.2)	33.4 (24.6, 42.2)	20.2 (16.5, 23.9)	30.0 (23.8 36.2)

Abbreviations: AANHPI, Asian American, Native Hawaiian and Pacific Islander; CHIS, California Health Interview Survey; CHI, Community Health Initiative AA, Associate of Arts; AS, Associate of Science

[a]Distributions (except age at interview) for controls were adjusted to the age distribution of the general ethnic-specific California population residing in study catchment area

[b]Table values are column percentages based on non-missing values only; percentages may not sum to 100 due to rounding

[c]Other AANHPI estimates were calculated including only respondents who took the CHIS in English

[d]Responses may include unemployed, retired, on disability, homemaker, student, or volunteer

[e]Limited to participants who spoke another language at home other than English (controls, N=267)

[f]Responses may include single-service plan (e.g. dental, vision, prescriptions), or insurance through another family member or organization

[g]Responses may include acupuncturist, websites, or self

[h]CHIS respondents limited to females ≥30 years of age

*Statistically unstable; has not met criteria for minimum number of respondents neede

REFERENCES

1. Grimes DA, Schulz KF. Compared to what? Finding controls for case-control studies. *Lancet* 2005;365(9468):1429-33.
2. Rothman KJ, Greenland S, Lash TL. Modern epidemiology: Lippincott Williams & Wilkins; 2008.
3. Schulz KF, Grimes DA. Case-control studies: research in reverse. *Lancet* 2002;359(9304):431-434.
4. Bernstein L. Control recruitment in population-based case-control studies. *Epidemiology* 2006;17(3):255-7.
5. Hartge P. Participation in population studies. *Epidemiology* 2006;17(3):252-4.
6. Cabral DN, Napoles-Springer AM, Miike R, McMillan A, Sison JD, Wrensch MR, et al. Population- and community-based recruitment of African Americans and Latinos: the San Francisco Bay Area Lung Cancer Study. *Am J Epidemiol* 2003;158(3):272-9.
7. Olson SH. Reported participation in case-control studies: changes over time. *Am J Epidemiol* 2001;154(6):574-81.
8. Bandera EV, Chandran U, Zirpoli G, McCann SE, Ciupak G, Ambrosone CB. Rethinking sources of representative controls for the conduct of case-control studies in minority populations. *BMC Med Res Methodol* 2013;13:71.
9. Brogan DJ, Denniston MM, Liff JM, Flagg EW, Coates RJ, Brinton LA. Comparison of telephone sampling and area sampling: response rates and within-household coverage. *Am J Epidemiol* 2001;153(11):1119-27.
10. Glaser SL, Clarke CA, Keegan TH, Gomez SL, Nugent RA, Topol B, et al. Attenuation of social class and reproductive risk factor associations for Hodgkin lymphoma due to selection bias in controls. *Cancer Causes Control* 2004;15(7):731-9.
11. Wang PP, Dicks E, Gong X, Buehler S, Zhao J, Squires J, et al. Validity of random-digit-dialing in recruiting controls in a case-control study. *Am J Health Behav* 2009;33(5):513-20.
12. Kagawa-Singer M. Improving the validity and generalizability of studies with underserved U.S. populations expanding the research paradigm. *Ann Epidemiol* 2000;10(8 Suppl):S92-103.
13. Clagett B, Nathanson KL, Ciosek SL, McDermoth M, Vaughn DJ, Mitra N, et al. Comparison of address-based sampling and random-digit dialing methods for recruiting young men as controls in a case-control study of testicular cancer susceptibility. *Am J Epidemiol* 2013;178(11):1638-47.
14. Link MW, Battaglia MP, Frankel MR, Osborn L, Mokdad AH. Address-based versus random-digit-dial surveys: comparison of key health and risk indicators. *Am J Epidemiol* 2006;164(10):1019-25.
15. Iannacchione VG, Staab JM, Redden DT. Evaluating the use of residential mailing addresses in a metropolitan household survey. *Public Opinion Quarterly* 2003;67(2):202-210.
16. Spooner TL, R.; Meredith, P. Asian-Americans and the Internet: The young and the connected. Washington, DC: Pew Internet and American Life Project; 2001.
17. California Health Interview Survey. CHIS 2011-2012 Methodology Series: Report 5-Weighting and Variance Estimation. Los Angeles, CA: UCLA Center for Public Health Policy Research; 2014.

18. Rao J, Scott A. On simple adjustments to chi-square tests with sample survey data. *The Annals of Statistics* 1987:385-397.
19. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. *BMJ* 1995;310(6973):170.
20. Lim JW, Paek MS. Recruiting Chinese- and Korean-Americans in Cancer Survivorship Research: Challenges and Lessons Learned. *J Cancer Educ* 2015.
21. Nguyen TT, Somkin CP, Ma Y, Fung LC, Nguyen T. Participation of Asian-American women in cancer treatment research: a pilot study. *J Natl Cancer Inst Monogr* 2005(35):102-5.
22. Gollin LX, Harrigan RC, Calderon JL, Perez J, Easa D. Improving Hawaiian and Filipino involvement in clinical research opportunities: qualitative findings from Hawai'i. *Ethn Dis* 2005;15(4 Suppl 5):S5-111-9.
23. Chao SZ, Lai NB, Tse MM, Ho RJ, Kong JP, Matthews BR, et al. Recruitment of Chinese American elders into dementia research: the UCSF ADRC experience. *Gerontologist* 2011;51 Suppl 1:S125-33.
24. Brugge D, Kole A, Lu W, Must A. Susceptibility of elderly Asian immigrants to persuasion with respect to participation in research. *J Immigr Health* 2005;7(2):93-101.
25. George S, Duran N, Norris K. A systematic review of barriers and facilitators to minority research participation among African Americans, Latinos, Asian Americans, and Pacific Islanders. *Am J Public Health* 2014;104(2):e16-31.
26. U.S. Census Bureau: State and County QuickFacts: Santa Clara County, CA. In: U.S. Census Bureau. Washington, DC; 2009-2013.
27. Kiernan M, Phillips K, Fair JM, King AC. Using direct mail to recruit Hispanic adults into a dietary intervention: an experimental study. *Ann Behav Med* 2000;22(1):89-93.
28. Heckathorn DD. Snowball Versus Respondent-Driven Sampling. *Sociol Methodol* 2011;41(1):355-366.
29. Brown BA, Long HL, Gould H, Weitz T, Milliken N. A conceptual model for the recruitment of diverse women into research studies. *J Womens Health Gend Based Med* 2000;9(6):625-32.
30. Trauth JM, Musa D, Siminoff L, Jewell IK, Ricci E. Public attitudes regarding willingness to participate in medical research studies. *J Health Soc Policy* 2000;12(2):23-43.
31. Ashing-Giwa KT, Padilla GV, Tejero JS, Kim J. Breast cancer survivorship in a multiethnic sample: challenges in recruitment and measurement. *Cancer* 2004;101(3):450-65.
32. Ford JG, Howerton MW, Lai GY, Gary TL, Bolen S, Gibbons MC, et al. Barriers to recruiting underrepresented populations to cancer clinical trials: a systematic review. *Cancer* 2008;112(2):228-42.
33. Hussain-Gambles M, Atkin K, Leese B. Why ethnic minority groups are under-represented in clinical trials: a review of the literature. *Health Soc Care Community* 2004;12(5):382-8.
34. Giuliano AR, Mokuau N, Hughes C, Tortolero-Luna G, Risendal B, Ho RCS, et al. Participation of minorities in cancer research: the influence of structural, cultural, and linguistic factors. *Ann Epidemiol* 2000;10(8 Suppl):S22-34.
35. U.S. Census Bureau, 2006-2010 American Community Survey, 5-Year Estimates. 2010.

36. Wang JH, Sheppard VB, Liang W, Ma GX, Maxwell AE. Recruiting Chinese Americans into cancer screening intervention trials: strategies and outcomes. *Clin Trials* 2014;11(2):167-77.
37. Yancey AK, Ortega AN, Kumanyika SK. Effective recruitment and retention of minority research participants. *Annu Rev Public Health* 2006;27:1-28.
38. Ma GX, Seals B, Tan Y, Wang SY, Lee R, Fang CY. Increasing Asian American participation in clinical trials by addressing community concerns. *Clin Trials* 2014;11(3):328-335.