Engaging a Hard-to-Reach Population in Research: Sampling and Recruitment of Hired Farm Workers in the MICASA Study

M. T. Stoecklin-Marois, T. E. Hennessy-Burt, M. B. Schenker

ABSTRACT. Hired farm workers provide the majority of the workforce for California’s labor-intensive agricultural sector. Agriculture is one of the most hazardous occupations, but there has been little research into the etiology of poor health outcomes that occur disproportionately in hired farm worker populations. MICASA is a cohort investigation of occupational and environmental health risks in hired farm worker households in Mendota, California, that employed a two-stage sampling process, including random selection of census blocks and door-to-door enumeration. The aim of this analysis was to evaluate the success of the sampling process and compare demographics of the enumerated population to other regional samples of Latino populations. In the enumeration, 1257 addresses were mapped and 729 hired farm worker households were enumerated. Findings showed no significant differences between the enumerated population and the resulting MICASA study sample; however, the MICASA population was more likely to be male, from Central America, work in agriculture, and have fewer years residency in the U.S. than California Health Interview Survey (CHIS) respondents. Additionally, 9.4% of the enumerated dwellings were back houses or unofficial dwellings and may have been missed by the U.S. Census 2000. Demographic comparisons between the enumerated population, census data, and CHIS data highlight the differences in these sampling methods and suggest possible demographic changes in hired farm workers in California. While difficulties in accessing hired farm workers often account for the lack of population-based research, the MICASA cohort provides an opportunity to examine occupational health patterns relevant to other farm worker populations.

Keywords. Farm workers, Hard-to-reach populations, Occupational health, Sampling.

California’s agriculture industry is the largest and most diverse in the country, generating products worth over $36 billion in 2007 (CDFA, 2008-2009). Over 400 commodities are produced in California, accounting for almost 13% of the nation’s total cash receipts for agricultural goods (CDFA, 2008-2009). California agriculture differs from other regions of the country with dry climate farming practices, near universal crop irrigation, and labor-intensive production methods. Consequently, the size, diversity, and uniqueness of agriculture in California present challenges to
research studies of hired farm worker populations, but may serve as a model for re-
search in other regions of the country.

California agriculture employs up to one million hired farm workers each year, who
provide more than 85% of the state’s farm labor. The majority of California’s hired
farm workers are Latino, and most are Mexican immigrants. Nationally, the Bureau of
U.S. Citizenship and Immigration Services reported that from 1980 to 2005, among
countries of birth from which immigrants came to the U.S., Mexico ranked highest,
representing over three million legal immigrants and 28% of all immigrants during the
same time period (Rytina, 2008).

Hired farm workers contribute substantially to the productivity of California agri-
culture, yet they incur a disproportionate share of occupational injuries and disease
(May and Kullman, 2002; Villarejo, 2003; Villarejo and Baron, 1999; Villarejo et al.,
2005). Farm work is a traditional entry job for Mexican migrants to California and is
also a known cause of adverse health outcomes (Kirkhorn and Schenker, 2002;
Agricultural exposures include numerous and diverse environmental and occupational
health hazards, including chemical, physical, and biological toxicants (Donham and
Thelin, 2006; Schenker, 1996, 1998b). While working in agriculture is a risk factor for
acute and chronic diseases, no study has separated the effects of agricultural work
from the effects of migration to California and immigration-associated changes in
health, nor have there been cohort studies that tracked changes in health over time.

Sampling Challenges of Farm Worker Health Studies

One factor presenting challenges for studies addressing occupational and environ-
mental risks to farm worker health is the perception that farm workers are a relatively
inaccessible population. Indeed, farm workers are a hard-to-reach population that pre-
sents a variety of methodological challenges to epidemiologic research (Zahm and
Blair, 2001). They can be difficult to locate, face cultural and language barriers, have
low levels of education and literacy, and have complex occupational histories (Cooper
et al., 2004; Kamel et al., 2001; McCauley et al., 2006; Zahm and Blair, 2001). They
may also have fears regarding immigration status, distrust government and research-
ers, and be more reluctant to participate in studies and reveal personal information
(Shedlin et al., 2009).

Traditional random sampling techniques can be difficult to employ since no exist-
ing list or sampling frame for hired farm workers is easily obtainable. Previous sur-
veys of farm workers have relied on a variety of methods, including site-based sam-
pling, often at farm worker residence camps (Anthony et al., 2010; Arcury and
Quandt, 1999; Arcury et al., 2009; Earle-Richardson et al., 2008; McCurdy et al.,
2003; Weigel et al., 2007), enrollment through community clinics or migrant health
centers (Goldman et al., 2004), employer or farm based recruitment (Gamsky et al.,
1992; Stack et al., 2006; U.S. DOL, 2005), and household-based sampling (Cameron
et al., 2006; Villarejo and McCurdy, 2008). The present study builds upon the house-
hold-based method of sampling with the goal of producing a randomly selected, repre-
sentative, community-based sample of farm worker families.

In the initial approach to data analysis, we evaluate the success of the enumeration
and sampling strategy by comparing the demographics of the resulting Mexican Im-
migration to California: Agricultural Safety and Acculturation (MICASA) study sam-
ple to the population in Mendota enumerated for the study. In addition, the California
Health Interview Survey (CHIS), a biennial statewide health survey conducted in all California counties, collects in-depth demographic and health data and provides a rare opportunity to compare Latino farm workers to the general Latino population in Fresno County. Therefore, the objectives of the present study are: (1) to compare enumerated family households and MICASA participant households on a number of demographic endpoints, and (2) to evaluate the demographic profile of the enumerated population in comparison to Fresno County Latinos interviewed for the 2005 CHIS and to data extracted from the U.S. Census 2000 for census tracts defining the area of Mendota, California (CHIS, 2005; U.S. Census Bureau, 2000).

Methods

Overall Design

The MICASA study is a population-based sample of 467 hired farm worker households in the town of Mendota, California. Mendota is located in California’s Central Valley and was chosen because of its large proportion of immigrants from Mexico and Central America and high proportion of hired farm workers. At the time of the household enumeration in 2005, the population of Mendota was estimated at 9791, with 97.1% reporting Hispanic/Latino ethnicity. Median household income was $25,422, and 32% of the population reported incomes below the poverty level (U.S. Census Bureau, 2005-2009). The primary crops grown in the area consisted of melon, tomatoes, cotton, nuts, and grapes.

The MICASA study is a cohort investigation of occupational and environmental exposures and other factors associated with health. A population-based household sampling strategy was used to avoid biases inherent in worksite or clinic sampling schemes, and to identify a population stable enough for follow-up. It was recognized that the sampling strategy would not capture the most mobile farm workers, but this tradeoff was considered appropriate to have a population that could be followed over time. Within this prospective cohort design, a series of cross-sectional and nested case-control studies were considered to examine a variety of specific health outcomes.

Household Enumeration

The goal of the sampling strategy was to select a random, community-based sample of Mexican and Central American hired farm worker families residing in Mendota, California. A stratified area probability sampling design was used with stratification into one of two census tracts and the census block as the primary sampling unit. Five extensively trained enumeration staff conducted both stages of the enumeration procedure. The effort was led by a consultant with extensive sampling experience in the National Agricultural Workers Survey (NAWS) and the California Agricultural Worker Health Study (CAWHS) as well as a study utilizing a similar methodology in Parlier, California (Sherman et al., 1997; Villarejo and McCurdy, 2008; Villarejo et al., 2010).

The sampling and enumeration were conducted in July 2005. In the sampling stage, a complete list of all census blocks contained within the two census tracts defining Mendota was compiled. Then a random sample of 62 blocks with 1100 census-identified housing units was selected. Enumerators went door-to-door in Mendota and created a universal list of all dwellings (including houses, apartments, trailers, sheds, and garages) within these selected blocks. In the enumeration stage, study eligibility
and household composition were determined. Enumerators returned to each mapped dwelling and obtained general information about the people living in the household. The enumeration procedure ascertained the relationship of the dwelling’s occupants to the head of household (i.e., spouse, adult child, parent, sibling, or other), age, gender, ethnicity, involvement in farm work, and years of residence in Mendota. This process was repeated for additional heads of household in cases where several families resided together. Additionally, for each head of household, the total number of children under the age of 18 years was elicited. Households with no hired farm workers were mapped, but individuals were not enumerated. A total of 729 households with farm worker residents in the selected census blocks were enumerated.

At recruitment, family households were randomly selected from the enumeration results and contacted in this random order. Family households were defined as households comprised of an eligible head of household and spouse, with or without children residing in the home. Households with a single head of household (mostly female) and children residing in the home were also eligible and included as family households for recruitment. Unaccompanied male households were defined as those households with multiple unrelated adults residing together. A small sample of unaccompanied male households were contacted for completing an interview but were not included for cohort follow-up.

Participant recruitment began in November 2005 and continued through January 2007 with home visits to enumerated family households, with the goal of recruiting 400 family households into the cohort. For households with more than one eligible family, convenience sampling was used, and whoever answered the door was invited to participate. All family households were contacted for recruitment without achieving the desired number of participant households. Therefore, a supplemental second enumeration phase took place following the procedures used in the initial process. Twelve additional census blocks were randomly selected. However, in this second enumeration, households were mapped, enumerated, and if eligible, contacted for recruitment in one step.

**Household Eligibility**

At recruitment, trained interviewers contacted enumerated households, explained the purpose and procedures of the project and attempted to interview both the head of household and spouse in each household. Eligible participants included men and women between the ages of 18 and 55 years who self-identified as Mexican or Central American, resided in Mendota at the time of the baseline interview, and at least one member of the household must have been engaged in farm work for at least 45 days in the last year. During this phase, households either agreed to participate, declined participation, or were determined to be ineligible.

**Informed Consent**

After verification of eligibility, the study objectives and procedures were explained, and individuals were invited to participate. Spanish is the primary language of participants, and all explanations and consents were provided in Spanish. Written consent was obtained from each eligible respondent who agreed to participate in the study. The study procedures were approved by the University of California, Davis, Institutional Review Board.
Data Collection Instruments

The survey instrument collected information on demographic characteristics, occupational and environmental risk factors, home environment, diet and physical activity, food handling practices, food security, acculturation, smoking status, and health outcomes including respiratory health, injuries, mental health, and reproductive health. The questionnaire also obtained basic information on all children in the household, including age, schooling in the U.S., and work in agriculture.

Statistical Analysis

Comparisons on demographic characteristics were made between the enumerated population in Mendota and Fresno County Latinos interviewed in the CHIS in 2005 and with the U.S. Census 2000 data for census tracts defining Mendota, California (CHIS, 2005; U.S. Census Bureau, 2000). Data from the 2005 CHIS were used to correspond with the time frame for the enumeration conducted in Mendota. The CHIS is a random-digit dial telephone survey conducted every other year in California and designed to be representative of the non-institutionalized household population (CHIS, 2007). Sampling for the CHIS is designed to meet two purposes: (1) to provide estimates for large and medium-sized counties and for groups of the smallest counties, and (2) to provide statewide estimates for the overall population as well as racial and ethnic groups and some ethnic subgroups (CHIS, 2007). A weighting process is applied to the sample data to produce population estimates. AskCHIS, an online query system that contains weighted totals and percentages for nearly all variables in the CHIS datasets, was used to generate county-level cross-tabulated estimates of demographic variables for Fresno County Latinos (CHIS, 2005).

U.S. Census 2000 data for sex, age, years residing in the U.S., and occupational industry were compiled using American Fact Finder and restricted to the two census tracts that comprise the city of Mendota (U.S. Census Bureau, 2000). Census 2000 data were used since these populations estimates are the most stable for small communities such as Mendota.

Frequency data and cross-tabulations were conducted to describe the enumeration data. Comparisons between enumeration data with the 2005 CHIS and the U.S. Census 2000 data were assessed using chi-square tests, with significant differences reported at the 0.05 level. Data were analyzed using SAS version 9.2 for Windows (SAS Institute, Inc., Cary, N.C.).

Results

Enumeration Results

In the first enumeration stage, within the 62 randomly selected census blocks, 1257 addresses were mapped. Of these, 729 households with farm workers were enumerated (table 1). Within the remaining households, 442 had no farm worker residents, 40 were vacant, six were businesses or other non-resident dwellings, 14 did not want to provide any information, and 26 were mapped but not enumerated. Among addresses classified as a non-resident dwelling, most of these consisted of the manager’s office in an apartment complex. Dwellings were classified as vacant if they were uninhabited or if the occupants were not living there at the time of the enumera-
Table 1. Summary of enumeration of 62 census blocks in Mendota, California, July 2005.

<table>
<thead>
<tr>
<th>Disposition</th>
<th>Number</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total dwellings mapped</td>
<td>1257</td>
<td>--</td>
</tr>
<tr>
<td>Farm worker households enumerated</td>
<td>729</td>
<td>58.0</td>
</tr>
<tr>
<td>Households with no farm workers</td>
<td>442</td>
<td>35.2</td>
</tr>
<tr>
<td>Vacant</td>
<td>40</td>
<td>3.2</td>
</tr>
<tr>
<td>Refused (did not want to provide information)</td>
<td>14</td>
<td>1.1</td>
</tr>
<tr>
<td>Business or other non-resident dwelling</td>
<td>6</td>
<td>0.5</td>
</tr>
<tr>
<td>Mapped but not enumerated</td>
<td>26</td>
<td>2.1</td>
</tr>
</tbody>
</table>

tion due to work in another state or outside of the U.S. For addresses that were mapped but individuals within the dwelling were not enumerated, most consisted of occupied dwellings where there was no answer to enumerators after repeated visits. Of the 1257 dwellings mapped, 118 (9.4%) consisted of trailers, garages, or sheds with residents and were located on property with a principal residence. Sixty-six percent $(n = 78)$ of these “unofficial” dwellings housed farm worker residents, and 52 (44%) were households comprised of unaccompanied males.

The household enumeration identified 2441 individuals living in 729 households in Mendota. Mendota is comprised of two census tracts, with 52.8% of enumerated individuals residing in one tract and 47.2% within the second. Twenty-four percent were households with unaccompanied males, and 9% had two or more married couples. Among enumerated individuals, 85.8% were currently engaged in farm work. Sixty-six percent were men, and 34% were women. Over half were Mexican born (55%), 35% were from Central America (primarily El Salvador and Honduras), and 10% were U.S. born.

Response Rates for Interview Participation

From the enumeration, 467 households were recruited and participated in the MICASA study, and interviews were completed with 843 adults. Of these, both the head of household and spouse were interviewed in 376 households, only one household member was interviewed in 21 households, and 29 households consisted of single heads of household (table 2). In addition, 41 households comprised of unaccompanied males participated, with one household member completing an interview. Seventy percent of contacted households participated in the study and completed interviews, with 30% declining participation. The primary reasons given by those who declined to participate in the study included desconfiado (distrusted), no time, not interested, or did not want to share personal information. A total of 351 households were determined to be ineligible due to age (over the age of 55 years), no farm worker living in the household, or the family had moved away prior to study recruitment.

Comparisons between enumerated family households and MICASA participant households were made to examine the success of the enumeration and sampling strategy. No significant differences were found between the enumerated family population and the MICASA sample in terms of age, country of birth, and number of years living in the U.S. (data not shown).

Latino Population Comparisons

Although Mendota lies within the geographical boundaries of Fresno County, it is much less diverse economically, demographically, and occupationally than the county as a whole. To evaluate the demographic characteristics of the MICASA enumerated
population in the context of other California survey data, comparisons were made to both Fresno County Latinos interviewed for the 2005 CHIS and to data extracted from the U.S. Census 2000 for census tracts defining the area of Mendota, California (CHIS, 2005; U.S. Census Bureau, 2000) (table 3). There were no significant differences between the census data and enumeration data on gender, age, and years living in the U.S. However, the enumeration data captured a greater proportion of agricultural workers than reported by the census data, as expected given the parameters of enumeration and sampling in the study.
Alternatively, the 2005 CHIS data were significantly different from the enumeration data on all of these parameters, except age group. While an equal proportion of men and women were interviewed for the 2005 CHIS, two-thirds of the enumerated Mendota sample consisted of men ($p = 0.03$). The enumerated sample included more individuals born in Central America than were interviewed for the 2005 CHIS (35.0% vs. 0.4%). Additionally, CHIS 2005 interviews included a much larger proportion of U.S. born Latinos ($p < 0.0001$) and those residing in the U.S. for 15 or more years ($p = 0.0003$), reflecting a Latino population that is less likely to work in agriculture.

**Discussion**

In general, hired farm workers are a hard-to-reach, underserved population. The goal of the MICASA study was to obtain a random, representative sample of hired farm worker families living in Mendota, and it appears that the use of the two-stage enumeration and sampling procedure succeeded in this endeavor. Age, country of birth, and number of years living in the U.S. were comparable between the enumerated family population in Mendota and the resulting MICASA sample. Studies in California and Florida have utilized a similar two-stage, household-based sampling approach with success (Cameron et al., 2006; Villarejo and McCurdy, 2008). The California Agricultural Workers Health Study (CAWHS) conducted enumeration and sampling of households in seven communities representing agricultural regions in California, reported high response rates, and assessed comparability through the proportion of agricultural workers surveyed from each community with agricultural employment reported by the California Department of Employment Development (Villarejo and McCurdy, 2008). In Florida, high response rates were achieved and demographic profiles were consistent with the NAWS data (Cameron et al., 2006). The MICASA enumerated population is demographically similar to both CAWHS in California and the NAWS in terms of age, country of origin, and years residing in the U.S., lending further evidence toward the success of the sampling procedure (U.S. DOL, 2005; Villarejo and McCurdy, 2008).

The household enumeration and sampling process identified 1257 dwellings, including 118 consisting of “unofficial” dwellings, such as trailers, garages, or sheds. The U.S. Census 2000 counted 1100 dwellings in these same census blocks, suggesting that these “unofficial” dwellings were likely uncounted (Sherman et al., 1997). These “unofficial” dwellings accounted for over 9% of the mapped dwellings, consistent with 11% informal dwellings identified in the CAWHS (Villarejo, 2011). These types of dwellings are more likely to house unaccompanied males, have a greater likelihood of overcrowding, and lack plumbing or food preparation amenities, such as refrigerators, stoves, and washing facilities (Sherman et al., 1997; Villarejo, 2011). Furthermore, these substandard housing conditions are beginning to be linked to negative health outcomes, including infectious disease outbreaks and gastrointestinal illnesses (Villarejo, 2011). Utilizing a household enumeration procedure allows identification of these informal dwellings and inclusion of individuals likely engaged in farm work and among the more vulnerable population.

The enumerated population reflects the changing nature of hired farm workers in California. The proportion of males and those engaged in farm work reported in the enumeration was not surprising, given the design and *a priori* knowledge of farm worker communities. The higher proportion of males is likely attributable to the unac-
companied males residing in the area for the purposes of agricultural work. Additionally, as unaccompanied males are more likely to reside in back houses or informal structures, the enumeration procedure identified and included them where they would be unlikely to be counted with other methods. Similarities between the demographic profiles of the enumerated population with the NAWS and CAWHS support this finding (U.S. DOL, 2005; Villarejo and McCurdy, 2008).

Over one-third of the enumerated population consisted of individuals from Central America, specifically from El Salvador and Honduras. This distribution by country of birth was not reflected in the CHIS data collected during the same year. The Census 2000 data for Fresno County estimated that 10.5% of Latinos were of Central American descent, substantially higher than the representation in the 2005 CHIS (U.S. Census Bureau, 2000). Additionally, the 2010 U.S. Census Bureau estimates for Mendota show a doubling of those of Central American descent to almost 25%, with this increase largely accounted for by individuals from El Salvador (U.S. Census Bureau, 2010). Enumerated individuals from Central America reported significantly shorter residency in the U.S. compared to those from Mexico (7.9 years vs. 11.8 years, p < 0.0001). This finding suggests that the newest immigrants are more likely to be from Central America, with individuals from Mexico comprising longer-term residents. The random digit dial sampling used by the 2005 CHIS appears less likely to reach the less settled and newer immigrant population, and there seems to be a demographic shift with an increase in migration from Central America, particularly with immigrants from El Salvador.

The enumerated population reported a higher prevalence of agricultural work than the 2005 CHIS and U.S. Census 2000 data. This was not surprising since the sampling strategy was designed to selectively count hired farm workers, and individuals in other occupations were not enumerated in the second stage. However, the small prevalence of agricultural workers in the 2005 CHIS suggest that data sources such as these, while useful for many purposes, would not provide an adequate inclusion of this underserved population.

The enumeration segment of the sampling plan enabled us to develop a reasonably complete sampling frame from which we were able to draw a random sample of hired farm worker families. Typical methods used to create a sampling frame or “list” can be problematic to employ with this population. Census data do not accurately count all farm workers, not all households or families maintain telephones, and inclusion of persons in the U.S. without documentation precludes the use of immigration records (Link et al., 2006). In accessing a hard-to-reach population, the importance of utilizing community-based assets cannot be overemphasized (Spring et al., 2003). Through community-based sampling, the study maintains a presence in the community and builds rapport with participants, furthering efforts of longitudinal follow-up. Employing community members for the local field team to conduct the enumeration and interviews was critical to gain the trust and confidence of potential study participants and overcome obstacles often faced in community-based research (Eskenazi et al., 2005).

A major limitation of the study is the focus on families of hired farm worker households. While the enumeration included unaccompanied male households and a small number were recruited for an interview, these households were not included for cohort recruitment. This concession was made to improve our ability to follow participants over time, since the population of unaccompanied males is the most migratory and difficult to track. While restricting the cohort to more settled family households
limits assessment of an at-risk segment of hired farm workers in California, this limitation was considered a necessary tradeoff to examine occupational health concerns of the MICASA cohort over time with acceptable loss to follow-up rates.

The hired farm workers included in the MICASA study have demographic characteristics similar to other farm worker populations, including the CAWHS and the NAWS. However, we recognize that farm worker populations in other parts of the country work in different crops, engage in numerous types of agricultural tasks, and follow distinct employment practices. These data are not intended to be conclusively related to all farm workers across the U.S., but the MICASA cohort likely share characteristics of other farm workers in California and possibly suggest occupational health patterns relevant to agricultural workers in other part of the country.

Conclusion

Hired farm workers supply the vast majority of the agricultural labor in California and increasingly in other states. Yet despite this dependence on hired, largely Latino, farm workers, few data exist on their health status as well as any significant guidelines for improving their health. The household enumeration strategy described here reflects a worthwhile investment of cost and time to produce a representative, randomly selected cohort because it captures the breadth of dwelling types in which farm worker families reside. The focus on the household is a resourceful approach for creating a cohort that not only characterizes the population of interest but also has an increased potential of successful follow-up over time to examine the health of hired farm worker populations.

Acknowledgements

We are grateful to all of the local field staff, particularly Gloria Andrade, Ana Cervantes, Alex Cervantes, and Giselle Garcia, and all the participants who contributed to this study. Funding was provided by the National Institute for Occupational Safety and Health (2U500H007550 and RO1OH009293) and a grant from the California Endowment.

References


