A Church-Based Sampling Design for Research with *Latina* Immigrant Women

Melanie R Wasserman[1], Deborah E Bender[1], William D Kalsbeek[2]

[1]Carolina Population Center and Department of Health Policy and Administration, University of North Carolina at Chapel Hill, USA; [2]Carolina Population Center, Department of Biostatistics, and Center for Health Statistics Research, University of North Carolina at Chapel Hill, USA.

Version: 3.22.2003

Please address all correspondance to:

Melanie Wasserman, Carolina Population Center, University of North Carolina at Chapel Hill, CB# 8120, University Square, 123 West Franklin St., Chapel Hill, NC 27516-2524.

E-mail: wasserma@email.unc.edu, Telephone: 919-843-4448, Fax: 919-966-6638 .

A Church-Based Sampling Design for Research with *Latina* Immigrant Women

Abstract

The U.S. is experiencing its highest immigration rate since the 1930s. The largest proportion of immigrants come from Latin America. Women constitute a large and growing fraction of *Latino* immigrants to the United States. Although our knowledge of the relationship between migration and women's health is increasing, we still have only limited knowledge of migration and the health of women of reproductive age. Research in the area of *Latina* immigrant women's health is particularly needed in states such as North Carolina where *Latino* immigrants are a relatively new arrival. Yet research is impeded by the absence of an adequate sampling frame: *Latina* immigrants remain a largely hidden population. This study tests in four North Carolina counties a church-based sampling frame for *Latina* immigrant women in their reproductive years. In the study area, on an average week, 20% of the Spanish-speaking population attends church (2/3 Catholic). Compared against Census data for the study area, the study sample (N=706) provided a comparable representation of different national origins. New entrants to the U.S. and married women were over-represented in the church-based sample. The young (under 30), and women at the lowest and highest extremes of educational attainment were under-represented in churches. While a church-based sample is not entirely comparable to the Census, churches can provide rapid and cost-effective access to a large number of new immigrants. Church-based research should be complemented with research in other settings, adjusting sampling weights for overlap between sampling frames.

Keywords: Hard-to-reach populations, immigration, *Latino*, sampling, women

INTRODUCTION

Women constitute a large and growing fraction of *Latino* immigrants to the United States. Although our knowledge of the relationship between migration and women's health is increasing, we still have only limited knowledge of migration and the health of women of reproductive age(Iglesias et al. 2003).

Research in the area of *Latina* immigrant women's health is particularly needed in states where *Latino* immigration is a new phenomenon. Prior to the 1990s, immigrant utilization of healthcare was a policy issue relevant principally in the Southwestern United States. Since the early 1990s, Southeastern states have experienced an unprecedented growth in immigration, principally from Latin America. North Carolina hosts the nation's fastest-growing *Latino* population, rising by 396% between 1990 and 2000. Most *Latinos* arriving in North Carolina are new immigrants and their children (as opposed to second and third-generation US-born *Latinos*): 64.3% of all *Latinos* living in North Carolina are foreign-born, and an additional 25.7% are U.S.-born children under the age of 18 (US Census Bureau 2001). While the absolute numbers remain small – 5.4% of North Carolina's population was *Latino* in 2002 (US Census Bureau 2001), the speed of change has left healthcare providers struggling to find the best way to extend their services to this new population (NCIOM 2003).

The first step to fill the knowledge gap on *Latina* immigrants is to find an adequate probability sampling frame. Sampling minorities remains one of the frontiers of survey design (Kalsbeek 2003). Research is still needed to identify the study designs most likely

to secure adequate minority representation, particularly among women. In recent years, the UNC Center for Health Statistics Research has been testing the comparative merits of a variety of sampling frames for *Latina* immigrant women. As a contribution to these efforts, the Migration Transitions Study tested a church-based sampling approach in four North Carolina counties. This paper presents results on the comparability of a church-based sample with the Census of 2000, and describes the practical aspects of church-based sampling.

BACKGROUND

Public health research is often time-sensitive and subject to financial constraints. Thus, research on an entire population of interest (the target population) is rarely feasible. Studies generally focus on a subgroup of individuals (the study population). If we wish to generalize research results back to the population of interest, the study population must be representative of the target population. Ideally, one would start from a master-list of all members of the target population, and draw a probability sample from it. Very few such master-lists exist for hard-to-reach minority populations such as *Latino* immigrants.

A hard-to-reach or hidden population is defined as one for which no adequate sampling frame exists, and acknowledgement of membership in the population may be threatening (Heckathorn 1997). Among immigrant *Latinas*, issues of trust and fear tend to impede research participation (Marin&Marin 1991). These issues may be particularly relevant for undocumented immigrants. Geographic dispersion presents a further obstacle to sampling new *Latina* immigrants (Kalsbeek 2003). In the study area, 38% of all *Latinos*

live in census tracts where they constitute less than 10% of the total population (US Census Bureau 2001). The field of sampling research has produced a number of possible solutions to the problem of sampling hard-to-reach populations. Alternate options for drawing a probability sample of *Latina* immigrants are reviewed below, before justifying the choice of a church-based sampling approach. Note that chain referral methods such as snowball sampling are not discussed here, because they do not produce a probability sample.

[TABLE 1 ABOUT HERE]

A review of sampling frame alternatives

Standard sampling methods include the use of mail-in surveys and random-digit dialing. Low literacy can impede the use of the first method, while the second method only samples the portion of the population equipped with telephones. Both methods are inefficient in areas where there is a low density of *Latino* immigrants: the harvest there is more chaff than wheat.

Ideally, to sample hard-to-reach minority groups in an area of interest, one would draw a sample from a list of population members where race and ethnicity are known for all entries on the sampling frame (Kalsbeek 2003). For *Latina* immigrant women in the U.S., one of the few such lists may be derived from U.S. birth records, which include the parents' date and place of birth. This list may be used to good advantage when research focuses on new *Latina* mothers and their U.S.-born children. For example, the Centers for Disease Control and Prevention (CDC) fund the Pregnancy Risk Assessment

Monitoring System (PRAMS), an annual survey of new mothers (Buescher 2003). Participants are selected at random from birth certificates, and contacted by phone. This survey produces a representative sample at the state level. The cost of using this strategy for new surveys at the county level may be prohibitive, however, and mothers' county-to-county mobility may limit the strategy's local utility.

In geographic areas with high concentrations of immigrants, an area-based sampling frame may be used to good advantage. For example, for the Health and Migration Survey, Donato and colleagues identified census tracts in Houston and San Diego with a high concentration of Mexican immigrants (Donato et al. 2003). Because Census data were out of date by the time of their 1996 and 1998 waves of data collection, researchers then spent several days walking through the neighborhoods, defining and redefining the boundaries (excluding, for example, neighborhoods containing principally commercial establishments). Then, they obtained a list of all household addresses in the neighborhoods defined, and randomly chose their sample of households from this list. Randomly selected households were each visited and screened for the presence of eligible persons. The cost of this approach may be quite high where there are lower concentrations of the population of interest (Kalsbeek 2003); and, despite the rapid growth in *Latino* population in North Carolina, concentrations in the overall population remain low. In the four counties of interest, of 108 census tracts, one, at 34%, passed the 30% mark, but 79 had fewer than 10% Hispanic/*Latino* population members; 21 had between 10 and 20%; and 5 had between 20 and 30 (US Census Bureau 2001).

Beyond cost, another concern with the area-based approach is that neighborhood segregation may affect health outcomes (Kawachi&Berkman 2003). Thus, if research is conducted only in neighborhoods with high concentrations of *Latino* immigrants, results obtained there may not apply to the target population at large. This is particularly problematic in locations such as our study area, where a large portion of the target population lives in areas where it is diffuse (US Census Bureau 2001). In national or state-wide studies, this difficulty can be overcome by including in the study sample some areas with lower concentrations of *Latinos*, and assigning a higher sampling weight to observations from these areas. For a local study, however, this approach may not be cost-effective.

The problem of sampling hard-to-reach minorities in low-density areas is similar to the problem of finding a small number of needles in a haystack. The thorough approach is to examine every stalk of hay until the needles are found, with the assistance, perhaps, of a large crew of farmhands. A faster approach might involve the use of a magnet. Similarly, several researchers have noted that diffuse, hard-to-reach populations may become more visible in institutional settings where they gather (Muhib et al. 2001; Watters&Biernacki 1989). For example, public health studies of *Latina* immigrant women have been conducted in healthcare facilities (Jones, Cason&Bond 2002) and schools (Sipan et al. 2003). Sampling members of these populations in the locations where they gather is certainly cost-effective. However, sampling in institutional settings brings into sharper focus the question of how the target population is connected to the sampling frame. For example, a school-based study will only reach women with children

over the age of 5. Similarly, clinic-based studies exclude women who do not access healthcare (Jones et al. 2002). This is particularly problematic for studies that focus on access to care.

Like other institutional settings, churches can potentially facilitate the rapid and cost-effective recruitment of a large number of *Latina* immigrant women. Church-based sampling may be particularly appropriate in North Carolina: in this state, churches play an important role in health promotion, act as brokers of community services, and serve as focal points for the social and political activities of the *Latino* community (Parra 1999). An important caveat for health researchers is that church attendance may positively affect health outcomes (Ellison&Levin 1998), so that a church-based sample may present a more optimistic picture of health outcomes compared to an area-based sample. That issue is beyond the scope of this paper, and will be addressed in forthcoming publications.

While some public health research has been conducted in churches (Carter-Edwards et al. 2002; Derose et al. 2000), to our knowledge, no prior studies have addressed the comparability of a church-based sample to Census data for a geographic area of interest. The comparison is useful not only for research, but also for interventions. Churches have become increasingly active participants in public health endeavors (Chatters, Levin&Ellison 1998; DHHS 2001), but no empirical studies have examined which portions of the population can be reached through churches, and which are excluded.

DATA

Data for the study were collected in fall 2002 and spring 2003 in eight churches from four North Carolina counties, as part of a larger endeavor, the Migration Transitions Study. The target population for the study was recently arrived immigrant women in their childbearing years, born in Spanish-speaking countries of Latin America and the Caribbean.

In the churches, women between the ages of 15 to 44 completed short screening forms to assess their eligibility for the Migration Transitions Study interview. For the purposes of this paper, only data from the screening form were used. Data from the full interview will be analyzed in forthcoming publications.

A total of 775 short questionnaires were collected in the eight churches combined. Final sample size is 706, after deleting repeat questionnaires and questionnaires outside the age-range of interest. Sample data were compared with Census 2000 data (Summary files 1, 3, and 4).

[Map 1 about here]

Sampling frame construction

To create the church-based sampling frame, a preliminary survey of churches with Spanish-language services was conducted in four contiguous counties of North-Central North Carolina. The Catholic Diocese provided the names and contact information of Catholic churches with Spanish-language services (Diocese of Raleigh 2003). Names of non-Catholic churches were gleaned from the North Carolina Office of Citizen Services, advocacy organizations, Spanish-language newspaper listings, and key informants in the community, until saturation was achieved (i.e. until all information gathered from new sources repeated existing information). Each priest or minister was interviewed by

telephone to confirm their church address, request the schedules of Spanish-language services, and ask for an estimate of the average number of parishioners. All Catholic priests, as well as 24 of the pastors and ministers of the remaining 27 churches, responded to our request for information. Churches that responded to our information request were included in the sampling frame.

[TABLE 2 ABOUT HERE]

Roman Catholic churches represent approximately 2/3 of the Spanish-speaking church-going population. A comparison of table 2 to Census data (US Census Bureau 2003b) reveals that the study area's estimated church-going population, at any given point in time, corresponds to approximately 20% of the foreign-born Spanish-speaking population.

Church selection

A stratified sampling approach was used to represent the diversity of church denominations and congregation sizes. Three strata were used: Roman Catholic churches; non-Catholic churches with congregations of 70 persons or more; and non-Catholic churches with congregations of fewer than 70 persons.

In the Catholic stratum, there were only four churches, one per county. This stratum was sampled with certainty, i.e. all Roman Catholic churches were included in the sample. Two churches were randomly selected from each of the other two strata. Baptist, Pentecostal, Episcopalian, and United Church of Christ denominations are represented in the final sample.

Recruitment within churches

In the eight churches sampled, the study goal was the complete enumeration of all women in the age-range of interest (15-44). The Spanish-speaking priest or minister made announcements about the study, in Spanish, on two consecutive weeks before the data collection visits. On the appointed dates, Spanish-speaking interviewers greeted parishioners as they entered the church and distributed short screening forms to all women in the age-range. The team of interviewers included bilingual university students and women from the *Latino* immigrant community. Following services, the priest and researchers explained the purpose of the study, and requested women's participation in completing the short screening form. Interviewers stood by after services to answer questions and assist women who had difficulty reading or writing.

Sampling time-periods

Priests and ministers indicated that church attendance is fluid. Only a portion of the congregation attends church on any given week. However, in the non-Catholic churches, a single visit was sufficient to gather screening forms from most parishioners in the age-range. Ministers were able to complete the roster of parishioners from their records, and put researchers in contact with the missing women. In Catholic churches, attendance is more fluid, congregations are larger, and registries do not reflect the population that actually attends services. Four visits were made to each Catholic church to ensure an adequate representation of their congregations.

While a larger portion of the *Latino* community attends church on holidays, priests and ministers unanimously rejected the idea of conducting research during holidays. Thus the sample obtained corresponds to the population that researchers might reasonably expect to reach in a cross-sectional church-based study.

A further question about the timing of data collection was whether church attendance might be higher during the growing season, with the influx of migrant laborers. Contrary to more Eastern areas in North Carolina (Wake County and points east), the economy of the study area is not heavily agricultural. The North Carolina Employment Security Commission estimates that the number of migrant Spanish-speaking farm workers in the area is a mere 405 persons, less than 2% of the total Spanish-speaking *Latino* population (NCESC 2002). The target population is women, the U.S. Department of Labor estimates that ¼ of all migrant farmworkers are women (US Dept. of Labor 1994), and it is unlikely that they all attend church; therefore waiting for the growing season would have only minimally added to sample size.

Data collection instrument

In the eight selected churches, all women aged 15-44 were asked to fill out a short screening form. This form replicated Census short and long-form questions regarding age, date of birth, marital status, country of birth, year of entry in the U.S. (for the foreign-born), years of education, and language spoken most frequently in the home. Women were also asked their year of entry into North Carolina, and ages of their U.S.-born and foreign-born children, if any. Finally, an item on the questionnaire asked for

their consent to be contacted later for an interview, and requested their phone number (women without telephones were asked for their addresses).

ANALYSIS METHODS

Descriptive statistics for the sample were obtained in Stata, using survey commands (`svyprop`, `svymean`, `svytotal, svygraph`) to adjust for weights and clustering of observations by church. Census data were then used as a gold standard against which to compare characteristics of the church-based sample. Sample proportions were compared to Census proportions using chi-square tests adjusted for sample weights and data clustering. The criterion for success in this experiment is that, if there is no statistically significant difference between sample data and the Census, then the church-based sample is comparable to the Census. Further details on methods used are provided below.

Issues with the use of the Census as a gold standard

Some concerns with the use of the Census as a gold standard are worth mentioning. First, there may be a differential undercount of *Latino* populations by the Census. The latest version of the Accuracy and Coverage Evaluation shows, however, that the differential undercount of the *Latino* population was not statistically significant in the 2000 Census (US Census Bureau 2003a). A second important concern with the use of the Census as the gold standard is the delay between Census data collection and diffusion. The delay means that data are out of date the instant they are published. This limitation cannot be remedied, except by assuming that in the space of three years, proportions within

population characteristics would have only minimally changed. Third, confidentiality issues affect the availability of the best data for comparison.

During research planning in 2001, the U.S. Census Bureau website stated that Census 2000 data would be available in fully searchable format by 2003. A remarkable amount of information is indeed currently available, but the mechanism for cross-tabulation of variables has some limitations that could not be anticipated at the time. For example, ideally data on place of birth (U.S. vs. foreign-born) would have been available in the counties of interest for Hispanic or *Latina* women between the ages of 15 and 44, as a group. The Census Bureau's Advanced Query function (available through State Census Data Centers) theoretically allows this cross-tabulation. However, because the age variable is shown in great detail, with increments of 5-year age groups, once it is cross-tabulated, cell counts become very small (<100), and data are suppressed to protect the confidentiality of Census participants. A less detailed age variable, with age 15-44 as a single category, is available for tabulations of Census short form data, but not for the long form data that are needed for this paper. For all its potential limitations, however, there is no standard more accurate than the Census.

Calculation of sampling weights

To generalize study results to the entire church-based population, we assigned sampling weights to individuals in the sample. For this purpose, the sample weight is calculated as the reciprocal of the probability of the selection of individuals in the sample, i.e.:

$$w = \frac{1}{p}$$

Where *w*= sampling weight and *p*= probability of selection.

Because the strategy of complete enumeration was used in each church, the probability of selection of each individual is the same as the probability of selection of the church. Thus, all observations from Catholic churches were given a sampling weight of 1, since they were sampled with certainty. Large, non-Catholic churches were given a sampling weight of 2/9 (since 2 of 9 churches were selected in this stratum), while smaller non-Catholic churches were given a sampling weight of 2/15 (since 2 of 15 churches were selected in this stratum).

To compare the church-based sample to a random sample drawn from Census data, each variable of interest was divided into a small number of categories (4 or less), because test statistics function best with larger cell counts. Sample proportions, adjusted for weights, were then tabulated against Census proportions. Chi-square tests were used to assess the statistical significance of differences noted. The null hypothesis, in this case, is that there is no difference between Census and church-based sample data in the distribution of the variable of interest. In other words, if chi-square statistics are not statistically significant, then the church-based sample is comparable to the Census.

If sample data were not clustered, it would be appropriate to use the Pearson's chi-square statistic, calculated as follows (Black 1999).

$$\chi^2 = \sum_{i=1}^{m}\left[\frac{(O_i - E_i)^2}{E_i}\right] \sim \chi^2 \text{ with } (m\text{-}1) \text{ degrees of freedom}$$

Where

O_i= Obtained cell count

E_i = Expected cell count (based on Census proportions and sample size)

m= Number of categories

This formula assumes an underlying poisson distribution, where the estimated mean is equal to its variance. With clustered sample data, this assumption is violated, and the formula must be adjusted to reflect the true variance of the estimated proportion:

$$z^2 = \sum_{i=1}^{m}\left[\frac{\hat{p}_i - p_i}{s}\right]^2 \sim \chi^2 \text{ with } (m\text{-}1) \text{ degrees of freedom}$$

Where

i= category

$\hat{p}_i$= estimated (sample) proportion for each category

p_i= population (Census) proportion for each category

s= estimated (sample) standard error for $\hat{p}_i$

m= number of categories

Stata's `svytab` command returns an adjusted chi-square statistic to test the independence of rows and columns within the sample. However, there is no command in Stata to calculate an adjusted chi-square statistic comparing sample proportions with an outside

standard. The adjusted statistic was calculated in Microsoft Excel, using for the denominator standard errors returned by Stata's `svyprop` command.

RESULTS

Response rate

Seven-hundred six eligible women completed the questionnaire. Because women who do not wish to participate may simply walk away uncounted, assessing response rates for the study requires some assumptions. Across each of the four counties of interest, women aged 15 to 44 constitute 24% of the *Latino* population. Table 3, below, provides response rates for the short screening form, assuming that this proportion holds true for the church-based population. Our rough head-counts corroborate this assumption, except in one of the smaller churches selected, where parishioners appear to be older. Based on the assumption, 91% of the estimated eligible population participated in the survey.

[TABLE 3 ABOUT HERE]

Sample description

The following descriptive statistics are adjusted for sampling weights. Ninety-seven percent of women in the sample were foreign-born. Most women (74%) were from Mexico, and 69% of the sample moved to North Carolina in the same year that they moved to the United States. Note that in the various field sites, data were collected partway through 2002 or 2003, thus immigration data for these years is incomplete.

[GRAPH 1 ABOUT HERE]

The sample's age distribution is truncated because of the study eligibility criteria (age=15-44). The spike at age 29 suggests that some women may be under-reporting

their age. This is not a concern when comparing sample proportions to the Census, since women would be expected to answer Census questions in the same way. There are spikes in education levels corresponding to the completion of primary school (6 years); middle school (9 years); and high school (12 years) (See graphs 2 and 3).

[GRAPH 3 AND GRAPH 4 ABOUT HERE]

Most women in the sample (67%) were married and 5.5% were in an *union libre* (living with a partner). Eighty percent of the women had at least one child, and 60% of all women had at least one U.S.-born child. Comparing Roman Catholics with other confessions, we note that 19% percent of Catholics had an education of high school or higher, compared to 26% of non-Catholics (p=.05). Thirty-nine percent of women in Catholic churches had arrived in the U.S. since 1995, compared to 30% of women in other churches (p=0.09). On average, women in Catholic churches had 1.8 children, and women in other churches had 2.15 children (the difference is not statistically significant).

Comparability of the sample to Census data

Sample characteristics were compared with Census 2000 data, pooled for the four contiguous counties, focusing on age, marital status, education, country of origin, and, for the foreign-born, year of entry into the U.S. (US Census Bureau 2001). Table 4, below, summarizes study results.

[TABLE 4 ABOUT HERE]

Compared to the Census, the church-based sample over-represented the foreign-born, and appeared to provide a roughly similar distribution of countries of origin. Women in the sample tended to be older (>30) and were more likely to be married than in the Census.

Finally, compared to Census data, fewer women at the lowest and highest extremes of education are represented in the church-based sample.

Implementation issues

In contrast with previous church-based research (Carter-Edwards et al. 2002; Derose et al. 2000), this study did not require as extensive a period of rapport-building prior to the beginning of the study. Leaders of selected Spanish-language churches all accepted to participate after receiving a letter of introduction from a respected local *Latino* advocate (the state ombudsman for *Latino* affairs). Church participation was confirmed by telephone and by a personal visit to the church by the researchers. In general, priests, church staff, ministers, and pastors went out of their way to ensure the success of the research.

One lesson learned about the implementation of church-based sampling is the importance of minimizing demands made on parishioners' time. In keeping with the idea of *personalismo* and the importance of earning parishioners' trust, the study protocol initially emphasized rapport-building with parishioners through attendance at prayer meetings, baptism education classes, and other events. During the first two visits to Catholic churches, parishioners were initially invited to a meal following worship services, during which researchers explained the purpose of the study and requested women's participation in the study. Most parishioners did not have time for this, because Sundays are often the only free time available to them for family visits. Only 27% of the estimated pool of available women participated in the study under this earlier version of

the protocol. Once the protocol was revised – with questionnaires distributed before Mass, an announcement by priests and researchers, and less than 5 minutes of women's time required to fill out the screening form – an estimated 91% response rate was achieved. This protocol was also less costly than inviting all parishioners for a meal: the incentive given to women to complete the short screening form was a single-serving packet of cookies, at a cost of under 50 cents each.

Discussion

When *Latina* immigrants are geographically dispersed, churches with Spanish-language services can be a cost-effective venue to quickly reach a large number of recently arrived *Latina* immigrants. On the one hand, the church-based sample appears to provide a good representation of the regions of origin of *Latina* immigrant women. Compared to the Census, the church-based sample over-represents the foreign-born, particularly those who arrived in the second half of the 1990s. Where the study aims to recruit recent immigrants (as in the case of the Migration Transitions Study), this difference with Census data is helpful.

On the other hand, it is important to acknowledge the differences between the church-based population and the population of the geographic area of interest. It may come as no surprise that compared to the Census, more women in the church-based sample are married. Seventy-nine percent of women in the church sample are currently married, compared to 59% of the area population. Public health studies need to take this into

consideration, because marriage has generally been found to have positive effects on health (Waite&Lehrer 2003).

An even more important difference is that the church-based sample under-represents younger women (under the age of 30), and women with the lowest levels of education. To adequately represent these groups, researchers may need to complement their church-based research with work in other community settings. Where multiple sampling frames are used, however, they may overlap. Each woman's chance of inclusion would depend on the number of sampling frames in which she appears, so that sampling weights would need to be adjusted accordingly (Kalton&Anderson 1986).

Acknowledgements:

Funding for this study was provided by the UNC Center for Health Statistics Research, under CDC/ National Center for Health Statistics Contract No. UR6/CCU417428-01; the Mellon Foundation; the Carolina Population Center; the Agency for Healthcare Quality and Research (AHRQ); and the Department of Health Policy and Administration of UNC-Chapel Hill.

Ted Mouw (UNC-Chapel Hill), Chirayath Suchindran (UNC-Chapel Hill), Laura Showalter (U.S. Census Bureau); Bob Coats (North Carolina Census Data Center), Katharine Donato (Rice University) and Emilio Parrado (Duke University) contributed to our thinking about the methodology.

Our contacts with churches and their parishioners were facilitated by Rogelio Valencia, North Carolina ombudsman for *Latino* affairs, and the priests, ministers, staff members,

and community partners of churches included in our study.

References

Black, T.R. (1999). *Doing Quantitative Research in the Social Sciences: an Integrated Approach to Research Design, Measurement and Statistics*. Thousand Oaks, CA: Sage Publications.

Buescher, P.A. (2003). A review of available data on the health of the Latino population in North Carolina. *N C Med J* 64(3):97-105.

Carter-Edwards, L., J.T. Fisher, B.J. Vaughn, & L.P. Svetkey. (2002). Church rosters: is this a viable mechanism for effectively recruiting African Americans for a community-based survey? *Ethn Health* 7(1):41-55.

Chatters, L.M., J.S. Levin, & C.G. Ellison. (1998). Public health and health education in faith communities. *Health Educ Behav* 25(6):689-699.

Derose, K., J. Hawes-Dawson, S. Fox, N. Maldonado, A. Tatum, & R. Kington. (2000). Dealing with diversity: recruiting churches and women for a randomized trial of mammography promotion. *Health Educ Behav* 27(5):632-648.

DHHS. (2001). HHS and the President's Faith-Based and Community Initiative.

Diocese of Raleigh. (2003). The Catholic Diocese of Raleigh. Accessed 9/15/2003. url: http://www.dioceseofraleigh.org/

Donato, K.M., M. Stainback, R. C.C., & K. S.M. 2003. The Significance of Gender in Explaining the Effects of Migration on Mexican Child Health. Presented at Population Association of America, Annual Meeting.

Ellison, C.G.& J.S. Levin. (1998). The religion-health connection: evidence, theory, and future directions. *Health Educ Behav* 25(6):700-720.

Heckathorn, D.D. (1997). Respondent-driven sampling: a new approach to the study of hidden populations. *Social Problems* 44(2):174-199.

Iglesias, E., E. Robertson, S.E. Johansson, P. Engfeldt, & J. Sundquist. (2003). Women, international migration and self-reported health. A population-based study of women of reproductive age. *Soc Sci Med* 56(1):111-124.

Jones, M.E., C.L. Cason, & M.L. Bond. (2002). Access to preventive health care: is method of payment a barrier for immigrant Hispanic women? *Womens Health Issues* 12(3):129-137.

Kalsbeek, W.D. (2003). Sampling minority groups in health surveys. *Stat Med* 22(9):1527-1549.

Kalton, G.& D. Anderson. (1986). Sampling Rare Populations. *Journal of the Royal Statistical Society* 149(1):65-82.

Kawachi, I.& L.F. Berkman. (2003). *Neighborhoods and health*. Oxford ; New York: Oxford University Press.

Marin, G.& B. Marin. (1991). *Research with Hispanic populations*. Newbury Park: Sage Publications.

Muhib, F.B., L.S. Lin, A. Stueve, R.L. Miller, W.L. Ford, W.D. Johnson, & P.J. Smith. (2001). A venue-based method for sampling hard-to-reach populations. *Public Health Rep* 116 Suppl 1:216-222.

NCESC. (2002). Personal communication. 1999 Estimate of Migrant and Seasonal Farmworkers During Peak Harvest by County. Fax from Rural Manpower Services, North Carolina Employment Security Commission.

NCIOM. (2003). NC Latino Health, 2003: Final Report of the Latino Health Task Force. Durham, NC: North Carolina Institutes of Medicine.

Parra, I. (1999). Personal communication. Discussion of the role of churches in the North Carolina Latino community.

Sipan, C., E. Blumberg, M. Hovell, N. Kelley, K. Moser, M. Ocana, L. Friedman, J. Acosta, A. Vera, & M. Adams. (2003). Screening Latino adolescents for latent tuberculosis infection (LTBI). *Public Health Rep* 118(5):425-433.

US Census Bureau. (2001). Census 2000, Summary File 1. Accessed through American Factfinder, url: http://factfinder.census.gov/home/saff/main.html?_lang=en.

—. (2003a). A.C.E. Revision II.

—. (2003b). Census 2000 Summary File 4. Accessed through American Factfinder, url: http://factfinder.census.gov/home/saff/main.html?_lang=en.

US Dept of Labor. (1994). Migrant Farmworkers: Pursuing Security in an Unstable Labor Market. *Research Report* (5).

Waite, L.& E. Lehrer. (2003). The benefits from marriage and religion in the United States: a comparative analysis. *Population and Development Review* 29(2):255-275.

Watters, J.K.& P. Biernacki. (1989). Targeted Sampling: Options for the Study of Hidden Populations. *Social Problems* 36(4):416-430.

Table 1. Probability Sampling Frames for Health research with *Latina* immigrants

Sampling Frame	Target Population	Research questions	Potential biases	Practical considerations	Example
Telephone-based (random-digit dial)	Adult population in area codes and exchanges of interest	Any research question	*Non-response bias and "yea-saying" bias*. Cold calls may reduce trust and lead to non-response or "socially desirable" answers. *Selection bias:* Participation depends on % of population with phone service	Difficult to oversample minorities, due to geographic dispersion (Kalsbeek 2003).	Behavioral Risk Factor Surveillance System (BRFSS) survey. (Centers for Disease Control (CDC) 2002)
Record-based	Women who recently gave birth or experienced another registered health event	Distribution of disease and practices among women who recently experienced a registered health event	*Selection bias:* mobile women (whose address has changed) will be excluded from the sample.	Survey can be conducted by mail and/or telephone. Very convenient for interviewers.	Pregnancy Risk Assessment Monitoring System (PRAMS) survey (NC State Center for Health Statistics 2001)
Area-based	Immigrants living in a defined geographic area	Prevalence and distribution of disease and practices in a geographic area	*Selection bias:* If only neighborhoods with a high density of immigrants are included, residential segregation may affect outcomes (Kawachi and Berkman 2003). For results applicable to a broader population, some low-density areas must be included and weighted.	Screening residents for eligibility may be costly, particularly in areas where there is a low density of the eligible population. Interviewer safety may be a concern.	Houston Migration and Health Survey (Donato 2000)
Institution based: Health care facility	Persons who access health care facilities	Prevalence and distribution of disease and practices among persons who access health-care facilities	*Selection bias:* persons who do not access health services are excluded	Long waiting times are an opportunity for researchers. New privacy regulations (HIPAA) limit recruitment based on records.	Intimate partner violence and substance abuse among minority women receiving care from an inner-city emergency department.(El-Bassel et al. 2003)
Institution based: Schools	Children and teens in schools, and their parents	Prevalence and distribution of disease and practices among school-aged children and/or their parents	*Selection bias:* preschool children and their parents are excluded.	Children in schools are a "captive audience", but parental consent may be difficult to obtain, particularly among minorities with limited English proficiency and/or low literacy	Screening *Latino* adolescents for latent tuberculosis infection (Sipan et al. 2003)
Institution based: Churches	Persons who attend churches	Any research question not offensive to the church (excludes abortion and family planning)	*Selection bias:* persons who do not attend church are excluded	(See methods section of this paper). Good venue for establishing trust with participants. Parishioners may only have limited time after worship services.	A church-Based Sampling Design for Research with *Latina* Immigrant Women (Wasserman et al., forthcoming)

Table 2. Church-based sampling frame: number of churches and estimated attendance (based on priests' and pastors' self-report)

Stratum	Number of churches in the study area	Estimated *Latino* attendance
Roman Catholic	4	2,950
Other – large (>=70 persons)	9	960
Other – small (<70 persons)	15	689
TOTAL	24	4,599

MAP 1. North Carolina counties and churches included in the study.

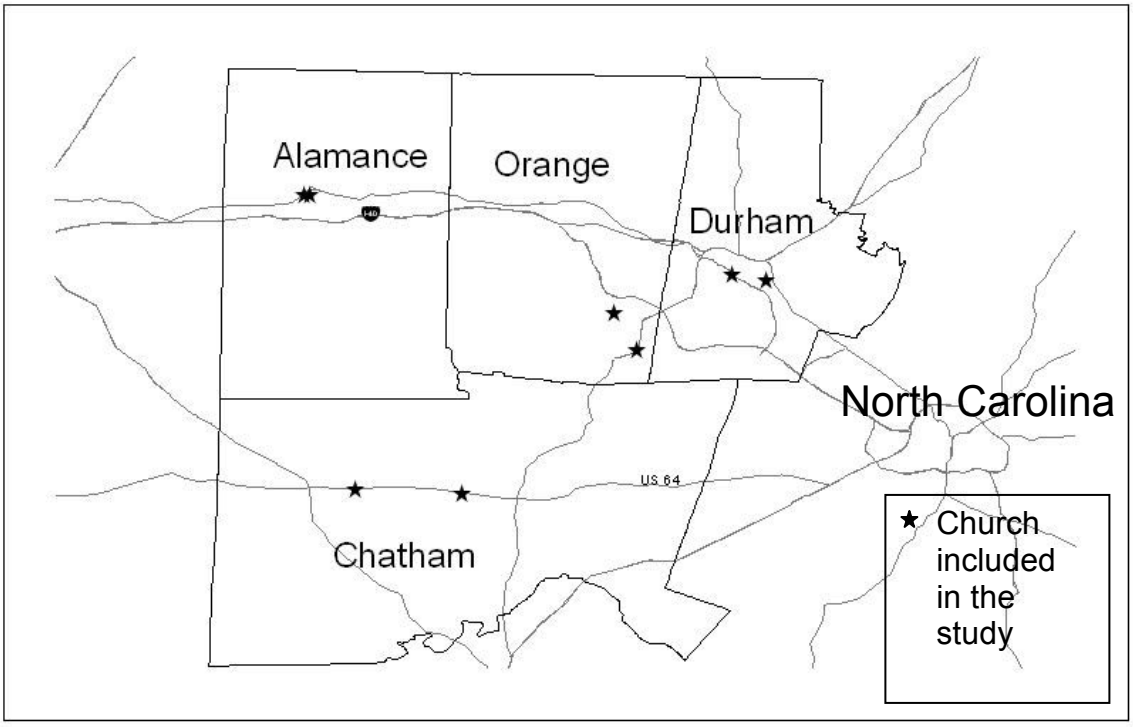
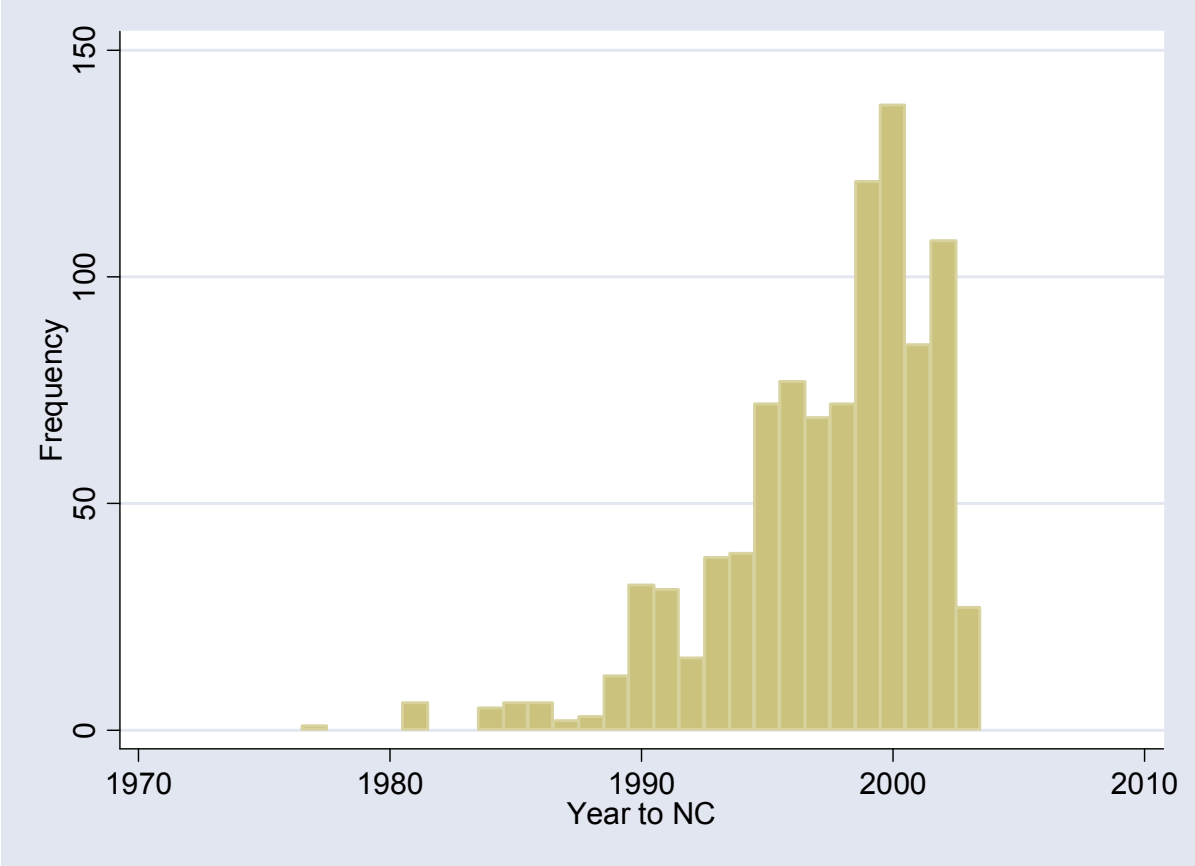


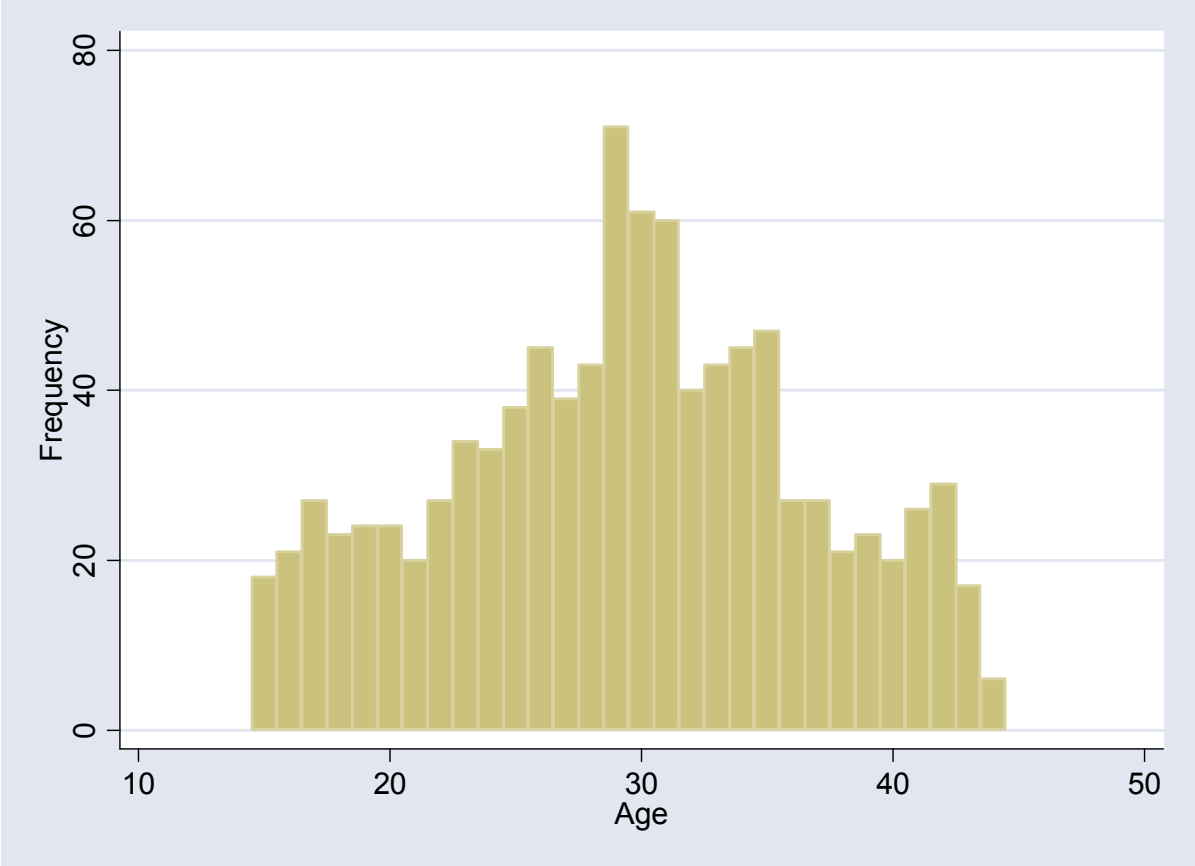
Table 3. Estimated response rates, by stratum

Stratum	Number of churches in stratum	Number of churches selected	Estimated attendance in selected churches	Estimated number of eligible parishioners	Partici-pation	Partici-pation rate
Roman Catholic	4	4	2,950	708	650	91%
Other – large (>=70 persons)	9	2	190	45	42	93%
Other – small (<70 persons)	15	2	90	21	14	58%
TOTAL	28	8	3,230	774	706	**91%**

Graph 1. Year of migration to North Carolina: women in the church-based sample



Graph 2. Age distribution of sample



Graph 3. Educational attainment of sample

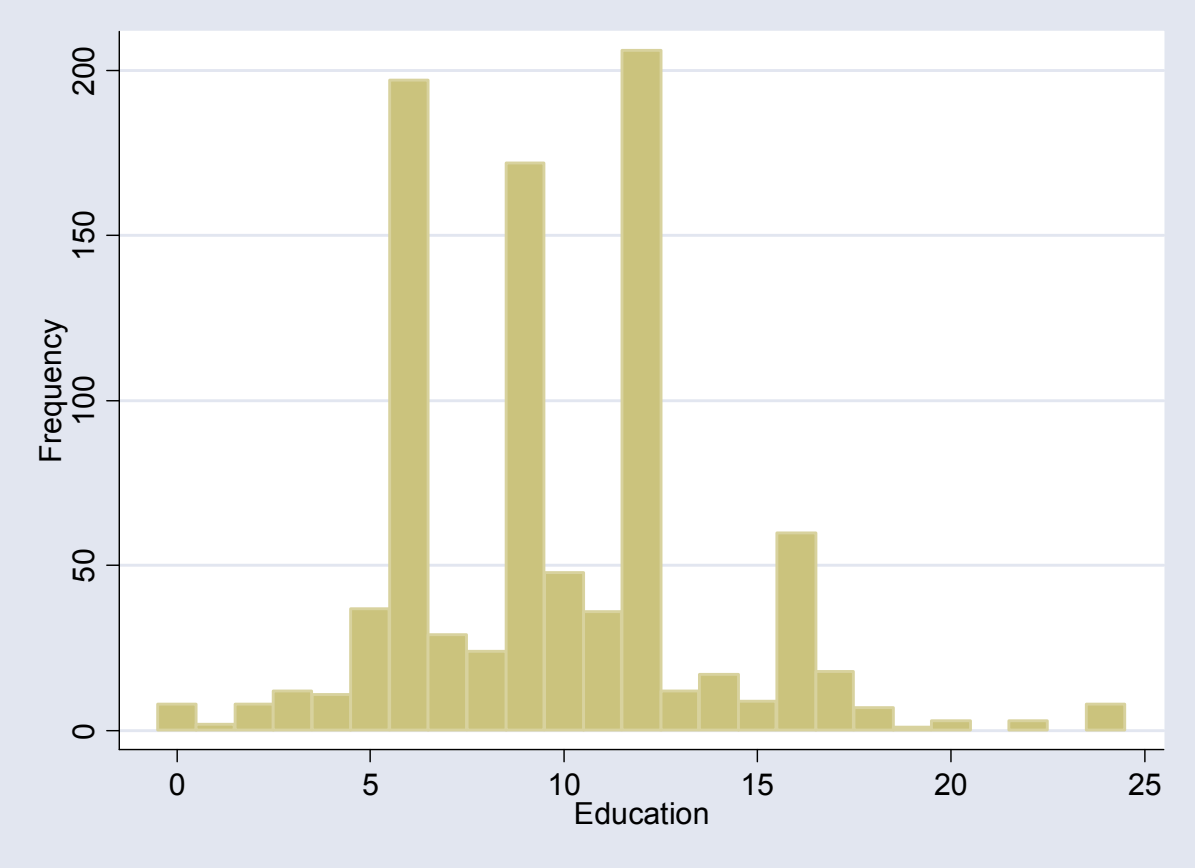


Table 4. Comparison of church-based sample characteristics against census data for 4 North Carolina counties

Characteristic	Census %	Sample %	Test Statistic (χ^2)	*Df*
Place of birth[a]				
United States	23	3		
Foreign-born	77	97	1077 **	1
Origin of the foreign-born[b]				
Mexico	75	78		
Central Am/Caribbean	17	15		
South America	8	7	0.81	2
Year of Entry[b]				
Pre-1990	33	14		
1990-1994	20	22		
1994-2000	47	63	44.57**	2
Age[c]				
15-19	15	12		
20-29	49	38		
30-39	28	41		
40-44	8	10	31.96**	3
Marital Status[c]				
Now married	59	73		
Never married (includes "union libre")	37	23		
Other	4	4	41.93**	2
Education[c]				
Less than 9th grade	38	34		
9th to 12th grade	37	46		
More than 12th grade	25	20	14.94**	2

N=706

* $p<=.05$; ** $p<= .01$

[a]Women 18 and older

[b]Census cross-tabulation by age unavailable. Proportions are not age-adjusted.

[c]Women age 15-44.