The Capstone Committee for Carlos H. Martinez Certifies that this is the approved version of the following Capstone

A comparison of participants and non-participants in a serum sample survey in Texas City, Texas.

Committee:

Edilma Guevara, Ph.D., Supervisor

Daniel H. Freeman, Ph.D.

Elizabeth Anderson, Ph.D.

Dean, Graduate School

A comparison of participants and non-participants in a serum sample survey in Texas City, Texas

by

Carlos H. Martinez, M.D.

Capstone

Presented to the Faculty of the Graduate School of

The University of Texas Medical Branch

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Public Health

The University of Texas Medical Branch

August, 2008

Dedication

To my son, Daniel, who brought me into this exciting adventure, and keeps giving me reasons to dream.

Acknowledgements

To the members of the Center for Population Health and Health Disparities of UTMB, especially Dr. Norma Perez, M.D., Dr.P.H., and Dr. Malcolm Cutchin, Ph. D.

A comparison of participants and non-participants in a serum sample survey in Texas City, Texas

Publication No._____

Carlos H. Martinez, M.D., MPH

The University of Texas Medical Branch, 2008

Supervisor: Edilma Guevara

Abstract: Lack of participation and response can affect the validity and generalizability of epidemiologic surveys. Participation has decreased in recent decades, and thus there is a need to determine reasons for and implications of this change. Usually respondents have better health, habits, and socioeconomic status and are interested in their health, being described as a "worried healthy" group. It is unclear if this explanation applies to minorities. The Environmental Risk, Coping, and Hispanic Health Project of the Center for Population Health and Health Disparities of UTMB, a population-based survey conducted in Texas City, Texas, included an interview of the randomly selected participants and an invitation to donate blood for biological measures. The study represented an opportunity to compare the characteristics of participants in the initial survey who, subsequently, provided blood sample and those who declined. In this Capstone the hypothesis that disadvantaged and less healthy groups are less likely to participate in serum sample collection as part of a sero-epidemiologic study was tested. The specific aims included: to determine the percentage of respondents declining to donate blood; to compare the characteristics of those who donated blood and those who declined, for the whole sample and across ethnic groups; and to examine the association with stress. We found that participation was lower than in similar surveys, with different response rates by ethnicity. We did not find great differences between respondents and nonrespondents, other than age, smoking habit, and perceived poor health. The only factors that were maintained in a multivariate analysis were age, family income, and smoking habit. We also found interaction between age and ethnicity, with advanced age being related to less likelihood of nonresponse in all ethnicities except in Non-Hispanic Blacks, where the direction was the opposite. A comparison of respondents and nonrespondents across ethnicities showed no additional factors, with the exception of results within US-born Hispanics. In addition to age, being a smoker and reporting poor health were associated with nonresponse, while low income was associated with less

likelihood of being nonrespondent. For all participants and across ethnicities perceived stress did not have an association with response.

Table of Contents

List of Tables

Research Plan

SPECIFIC AIMS

The decline in participation rates in epidemiologic surveys in the last decades increases the need to identify factors influencing participation in population-based research. The prevalent theory that asserts that the differences between participants and nonparticipants are predictable, with participants being healthier, part of a minority group, and in better socioeconomic status, has recently been challenged. Additionally, there is limited information about differences between respondents and nonrespondents across different ethnic and cultural groups. For surveys that include, besides the initial questionnaire, participation in a physical examination or provision of a blood sample, the information on participation rates and factors related to response is limited. The Environmental Risk, Coping, and Hispanic Health Study, a population-based survey in a multiethnic community, presents a unique opportunity to approach the problem of participation and response in epidemiologic studies that include blood sample collection. The identification of factors related to nonresponse will provide information that will help to improve the design of future research projects that include multiethnic populations and an invitation to donate blood for serological studies.

This Capstone will use data from the epidemiological and serological studies conducted as part of the Environmental Risk, Coping, and Hispanic Health Study to test whether disadvantaged and less-healthy groups are less likely to participate in the collection of a serum sample collection as part of an epidemiologic survey. This will be tested through the following Specific Aims and Hypothesis:

1. To determine the percentage of nonparticipation (nonresponse) in the blood collection stage of the survey, defined as response to the initial interview but failure to provide of a blood sample for analysis of stress biomarkers.

2.- To compare the demographic, health risks and behaviors, and disease profile of those who donated a blood sample and those who declined participation (nonrespondents) in blood sample collection.

Hypothesis 1: Nonresponse in the serological survey is associated with being less healthy, having a less healthy life-style, and having low socioeconomic status.

3. To examine the influence of ethnicity on response to the invitation to donate blood.

Hypothesis 2: The factors related to nonresponse in the serological surveywill vary by ethnicity.

4. To analyze the independent relationship between stress and participation in the blood sample collection.

Hypothesis 3: A high stress score is associated with a greater percentage of nonresponse in the serologic survey.

BACKGROUND AND SIGNIFICANCE

Introduction

Population-based research has been seen as one of the most powerful tools to use to obtain valid estimates of the prevalence of risk factors, certain behaviors, and health status that, ultimately, can be generalized to the whole population. The researcher usually expects a high response rate, as the first step in providing data to support community-based interventions and public health policy. Unfortunately, the observed response rate to surveys is highly variable, and according to some authors, is in decline. Recently, Galea et al. (1) presented data supporting the idea that participation rates are, in general, decreasing. This included some of the most frequently cited sources of epidemiologic information, as the Behavioral Risk Factor Surveillance Survey (71.4% participation in 1993 versus 51.1% in 2005), the Survey of Consumer Attitudes (with participation dropping from 62% in 1979 to 48% in 2003), and the National Health Interview Survey (whose participation rate went from 91.8% in 1997 to 86.9% in 2004).

There are multiple reasons public health professionals have to be concerned with the participation and response rates in epidemiologic studies. The first potential problem when participation is low is that the researcher and the users of the data will always have questions about the validity of the prevalence estimates (for diseases, health risks, and behaviors) provided from the study, especially in cross-sectional studies. A second reason to be interested in the participation rate is the theoretical potential to bias the estimates of the association between a potential risk factor and an outcome, when the response status is related to the exposure and/or disease status, due to a response bias or "self-selection" bias (2).

Available information about those who did not participate in a study is usually very limited, or totally unavailable, a factor that limits the development of corrective strategies to increase participation, and also limits the certainty about the population estimates of interest. If the nonrespondents are a random sample of the target population, the potential to endanger the validity and generalizability of the findings would be limited, but there is also a chance that respondents and nonrespondents differ in a systematic way. Therefore, the first step to use to improve participation should be identify such differences, something that has been set as a priority by the research community (3).

In order to address this question, we conducted a review of the literature on nonresponse in epidemiologic studies, with special focus on studies on ethnic differences in response. The search was performed on the Medline database, using the terms "nonresponse", "population", "epidemiologic studies", "cohort study", "cross-sectional study", "case-control study", "reasons", and "characteristics", in different combinations. The search was enriched with review of the references from the articles obtained, and with an additional search using the Sociological Abstracts database. All searches were restricted to the years 1970 to 2007. Articles focused on interventions to increase the response rate were excluded. From the literature review, it was apparent that at least two ways to investigate the characteristics of nonrespondents are in common use. The usual way is to perform a comparison between their demographic characteristics and health when at least some information is available, such as population databases, or to use the initial responses in multi-stage projects, as in some cohort studies. Another approach is to compare early respondents with late respondents, assuming that the more difficult to engage in the study could be more similar to the nonrespondents, a strategy applicable to cohort and cross-sectional studies. For this reason, we present together the information, independent of the design of the original study. The reasons for nonresponse, historically, have been considered to be homogeneous for the population, but due to the interest of the current project in a multiethnic community, the literature about ethnic differences will be presented separately. As the current project was focused on the response to the request to donate blood immediately after answering the survey's questionnaires, an additional search and section on response and participation in "serum surveys" is included. The following is a summary of the relevant findings.

Definitions of nonresponse

There is not agreement in the literature on the definition for reporting the "response rate". Some authors suggest that the right term is "response proportion", and one of the definitions is "the number of completed or returned survey instruments divided by the total number of persons who would have been surveyed if all had participated" (3). This definition would fit better for cross-sectional designs. In longitudinal studies the same definition could apply at the beginning of the study, and the lack of response for subsequent follow-up at all the repeated measurement times could be called attrition. However, as the characteristics of nonrespondents in cross sectional and longitudinal studies have been found to be similar, the term could be used to describe the population in both designs. When a more complex design is used, as in surveys that include blood donation as part of a "serum survey", the same term, "nonrespondent", has been also used to define the subgroup that did not donate blood, or did not participate in the physical examination part of the study, after answering the initial questionnaire. This term has been used in this way in different Health Examination Surveys, as in NHANES and its Hispanic version, HHANES. With this background, it is appropriate to be aware that some authors can present data on different subgroups of their target population, even when they use the same term, "response rate" or "participation rate". It is recommended

that each definition be reviewed in detail. In the current study we used the term "nonrespondent" to describe those persons who, after answering the questionnaire, did not provide blood for the serologic part of the survey.

A more detailed description of the terms "response" and 'nonresponse", has been proposed, and includes different subsets of response, such as "contact rate" (percentage of potential participants identified with whom contact is successful), "cooperation rate" (percentage who answered of those eligible and contacted), and "over-all response rate" (percentage of respondents of those eligible and selected to participate). The use of these definitions has been inconsistent in the literature.

Differences between participants and non-participants in the general population

In 1973, Comstock and Helsing used data from a cross-sectional study in a randomly selected sample of adults in Washington County, Maryland (4), and compared the characteristics of the 571 respondents with the census information data available for the 78 nonrespondents. They found that young adults, less-educated people and males were more frequently represented among nonrespondents. Since then, subsequent reports, based on the experience of the major epidemiologic surveys, expanded the profile of the "nonrespondent". Data from the San Diego Medicare Preventive Health Project, reported in 1992, showed that in the random sample of 1,600 participants in a prevention and health education program, those who did not attend the first year follow-up were older males (5). The Cardiovascular Heart Study, using a random sample of 5,201 Medicare enrollees, had similar findings: nonrespondents were older, not married, and had a lower educational level (6). The analysis of the 6,021 participants in the Minneapolis subset of the Atherosclerosis Risk in Communities, a population-based cohort study of a random sample of adults in the community, showed that nonrespondents to the second wave of the survey were more frequently not married, less educated, perceived their health status as poorer, and had a higher frequency of smoking (7). As part of the Consumer Assessment of Health Plans Study Survey (8), a continuous survey of the client satisfaction with Medicare managed plans based on a national random sample of 136,062 participants in these plans, showed that nonrespondents usually are women, disabled, and less educated. All these surveys used data from previous interviews or local census data to compare those who participated with the nonrespondents. That is the reason the main differences found were in the demographic, socioeconomic, and health habits domain.

To inquire into potential differences in health profiles and outcomes according to response status, it is necessary to match the data from the survey with national or regional registries of healthcare, healthcare use, and vital statistics. In a population-based cohort study in Taiwan, with the participation of all of the community, and the use of data from a local census and visits to the nonrespondents, Chou et al. (9) found that the 1,538 nonparticipants, out of 4,451 habitants, were more frequently older and men, and the mortality during the first two years after declining participation was also higher than in

the respondents. Nonrespondents in a cross sectional studies had a higher burden of psychological symptoms, as showed by van der Akker et al. These investigators matched the demographic information of 6,680 adults invited to participate on a health survey in The Netherlands with the data from their personal physicians (10). In another cross-sectional study, the authors used the interviews with their personal physician from 115 nonrespondents and 999 respondents that previously were invited to be part of a community-based health survey in The Netherlands. They found that, in addition to age (higher in nonrespondents), marriage status (more frequently not married), and educational (lower in the nonrespondents) differences, the nonrespondents more frequently reported a history of stroke, diabetes, heart disease, and psychiatric illness (11). A similar approach to the problem was used by Korkelia et al. (12), who linked the demographic information of 52,739 randomly selected participants in a national health survey between adults in Finland, with data from the same country's population registry. This group found, again, that men, older individuals, not married, with less education, and smokers were more frequently nonrespondents, but there were no differences in the health profiles between respondents and nonrespondents. A community-based cohort study in Norway had similar results: in a randomly selected sample of 3,786 adults, those who did not respond to the questionnaires used on the second wave of the study, were more frequently unemployed, and older (13). In a case-control study evaluating risk factors for lung cancer, using 154 incident cases and 154 population-based controls in Turin, Italy, the nonrespondents were older and not married, but some characteristics of nonrespondents were different according to the case or control status (higher income in nonrespondent cases but lower in nonrespondent controls) (14). A more recent follow-up of a national cohort study in The Netherlands, based on a random sample of 12,786 adults, confirmed that the less educated, older, and those with more reported morbidity but lower frequency of use of medical services, are more frequently nonrespondents. In cohort studies of elders, the main reason for nonresponse or attrition was death, as has been shown by Markides et al. in a cohort study of elders in Texas (15).

In summary, data provided by different designs, in different populations, as part of different research questions, converge to give a more or less homogeneous profile of the nonrespondent in epidemiologic studies: usually an older individual, male, not married, with lower education, and with a higher burden of morbidity, including psychiatric symptoms and depression, who usually smokes and drinks with higher frequency (16), and does not use the healthcare services available. In contrast, the respondent is usually described as healthier, with a better socioeconomic status, married, with better health, healthier habits, and more frequent use of health services, and with better health outcomes during the follow-up (when follow-up data have been available).

The explanations given for the differences between the respondents and the nonrespondents suggest that they are members of two different segments of the population (17). The respondents can be described as a subgroup of "worried well" or "worried healthy", a term introduced by Criqui et al. in the 1980s (2), whose participation is determined by their interest in their health, probably related to a higher exposure and retention of information about health and health hazards. In contrast, the nonrespondents

have been described as socially disengaged, and with lower interest in discovering or sharing information about what they consider could be socially undesirable or unacceptable (16, 18). The theoretical model is also convergent with the evidence obtained from other types of research, such as controlled clinical trials, that show a higher participation rate of candidates with similar characteristics to the respondents in epidemiologic surveys and epidemiologic research (19).

Nonresponse in multiethnic populations

The previously described explanations for participation and response predict that, due to their higher mobility, lower socioeconomic status, and high burden of disease, minorities will have a lower participation in surveys, and that nonrespondents will fulfill the "social disengaged" stereotype. More recent surveys, designed to include a more diverse population, have confirmed, with some exceptions, a low participation of minorities, but they have failed to identify the predicted differences between respondents and nonrespondents (8, 20), and have challenged the validity of the "worried healthy" and the "social disengagement" explanations.

The body of available information shows that Non-Hispanic Blacks have lower participation rates in all types of medical research, from surveys to clinical trials, and that efforts to improve the participation of Non-Hispanic Blacks have not been successful. Distrust of the medical community and the belief that the poor and the minorities will be exposed to a higher burden of risks during the research process has a strong relation with the failure to recruit African Americans (21, 22). As the participation rate seems to be more dependent on these factors, the demographic differences between African American respondents and nonrespondents in epidemiologic studies are less clear. In a multi-stage population-based mental health survey in Alameda County, California, with initial enrollment of 1,763 participants, 47% Non-Hispanic Blacks, and 26% Mexican-Americans, the response rate in the second wave was lower in both groups, compared with the Non-Hispanic Whites, but the only difference between African American nonrespondents and respondents was age, without differences in educational level and marital status (23). During a two-year follow-up of a community-based cohort study designed to improve the use of preventive measures in the community, whose 5,918 initial participants were selected by a random digit dialing method, the rate of nonresponse in the second wave was 51.3% in Non-Hispanic Blacks, and 47% in Hispanics, in comparison with 37.5% in Whites. The main difference between nonrespondents and respondents was age, for all groups. Unexpectedly, in African Americans a higher educational level, better health, and higher frequency of physical activity were related to nonresponse (24). In a comparison of the frequency of nonresponse and the characteristics of the nonrespondents across ethnic groups in the four centers of the community-based cohort study of cardiovascular health, ARIC (Atherosclerosis Risk in the Communities), with 19,772 participants in the initial interview and 15,800 in the subsequent physical exam, the relation between nonresponse,

age, health habits, and socioeconomic status was again proved for Non-Hispanic Whites, but no clear differences appeared between Non-Hispanic Black respondents and nonrespondents (25).

The response, and in general the participation rate for Hispanics, is more variable. National surveys with random sampling strategies have showed that the response is lower than in Non-Hispanic Whites and similar to that reported for Non-Hispanic Blacks (8, 20, 23, 26, 27). The inquiries on the differences between Hispanic respondents and nonrespondents have found that the expected differences in socioeconomic status, demographics, and health status, based on the "worried healthy" theory, are not present in this minority (28).

In summary, the previously described theory of respondents as a "worried well" population, and nonrespondents as "socially disengaged", is difficult to sustain for ethnic groups different from the Non-Hispanic Whites. The reasons for nonresponse could be culture- and ethnicity-specific, and new data and theoretical frames are required in order to improve the response rate in minorities. And finally, an additional piece of information points against the "worried well" hypothesis: another survey in The Netherlands, based on a national random sample of 31,556 subjects, linked the baseline health status and risks with the mortality data five years later, and found that the nonrespondents had a lower frequency of cardiovascular disease at baseline, but higher mortality at five years (29). The authors have proposed the existence of a "worried ill" population. The findings, however, have not had independent confirmation in other studies. Altogether, the evidence summarized here shows that there is an urgent need for research in the nonresponse field, and in the factors related to nonresponse across different ethnic and cultural groups, since the previously accepted theoretical models do not fit the multicultural and complex composition of American society. The current project looks to answer some of these questions.

Nonresponse in surveys that include blood samples

Not many epidemiologic projects have included a physical examination or blood donation, in addition to the questionnaire part of the survey. Probably the best known is the National Health and Nutrition Examination Survey (NHANES), a U. S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for health Statistics survey, that includes an in-person interview in the household, followed by a physical examination and medical tests in a mobile examination center. A similar design was used for the Hispanic Health and Nutrition Examination Survey (HHANES), a nationwide probability sample of approximately 16,000 persons, 6 months-74 years of age, conducted between 1982 and 1984 by the National Center for Health Statistics (30). Another population-based survey with a design similar to NHANES, recently completed, is the New York City Health and Nutrition Examination Survey, 2004 (31). Other local or regional projects that involved interview and physical

examinations or biological samples include the previously discussed ARIC (25), the different waves of the Framingham Cardiovascular Study, and the 1989-1990 and 2000-2001 phases of the Nurses' Health Study. As the design of the Environmental Risk, Coping, and Hispanic Health Study resembles the NHANES and its Hispanic and New York City replications, is of interest to discuss the frequency and factors related to nonresponse in those surveys.

For the NHANES III, performed between 1989 and 1994, the response rate for the interview was 86%, and 78% for the examination. In the NHANES 1999-2000, 82% accepted the invitation to be interviewed, and 76% agreed to the medical examination. In 2003-2004 the response to the interview was 79%, and 76% for the examination. Data for the period 2005-2006 are 80.4% for the interview, and 77.3% for the examination (32). Non-Hispanic Blacks had response rates of 81.7% (interview), and 78.8% (examination), and the figures for Whites were 78.4%, and 75.2%, while for those described as Mexican-Americans the rates were 82.7%, and 79.6%, respectively. In the HHANES, the response rate for the physical examination varied between 60.7% for those described as Cuban-Americans, and 76.1% for Mexican-Americans (33). The nonresponse rate in the HHANES prompted different evaluations of the factors related and the potential bias of the estimates. In general, the response rate was higher in younger people and varied according to the reported cultural background (higher in Mexican-Americans and lower in Cuban-Americans). After multiple comparisons between the examined and the screened population, and between subsamples of the respondents, no clear explanation was found, and the results for the Cuban sample are considered difficult to adjust. These findings point to even greater culture-specific differences in participation and response that still need to be evaluated. The NYC-HANES reported a response rate of 76% for the interview part, and 66% of the surveys were completed. This, according to their definition, means full questionnaire and at least one examination component. The Nurses' Health Study invited 32,826 participants (27% of the cohort) to provide blood samples in 1989-1990, and had a response rate of 97% (34).

Even when data on the frequency of nonresponse are known, the evaluation of nonresponse and the potential for nonresponse bias have been limited in these studies. In particular, limited information is available about the specific components of the physical examination and blood donation. In an early evaluation of nonresponse in the NHANES II, the comparison between those who participated in both the interview and the examination and those who did not completed the examination showed that nonrespondents for the physical examination part were older, their working status was uncertain, and they had less concern about medical problems that they wanted to discuss with a physician, which was a surrogate for the presence of diseases (35). The same profile of the nonrespondent had appeared in similar evaluations from the Framingham Study and the Pittsburgh Cardiovascular Study, whose analyses were also limited to participation in the physical examination portion. Whether the same characteristics of the nonrespondents apply to the serum-survey part of a population-based survey is unclear, but some information has recently emerged. An analysis of 485 persons who initially refused the medical examination part of the NHANES, but latter on converted to

respondents showed that these "converted refusals" were more likely to be current smokers. They, otherwise, did not have greater differences in risk factors when compared with respondents, and had a lower educational and income level. Unexpectedly, a parallel comparison between converted refusal at the time of the interview and the initial respondents did not show any difference in socioeconomic factors or risk factors (36, 37). In summary, for health examination surveys, the main differences between full participants and participants that did not participate in the serum collection or physical examination part, are still unclear, and are limited to age, work status, and ethnicity.

The identification of factors influencing the decision to donate blood as part of a survey will be even more critical in the future, as serum markers and seroprevalence studies will supplement other available information. If the response rate and the factors related to response are unknown, the potential to bias the estimates and the design of evidence-based preventive interventions will be endangered (38). The interpretation of the everyday more-common data from genetic tools in epidemiology could also be limited if the response rate is different for some specific subgroups. A recent analysis of the participants in the NHANES 1999-2000 who were asked to provided consent for genetic research using the blood samples they provided, showed that, consistently in both years, being Non-Hispanic Black was associated with refusal to consent for genetic tests, and that in 2000 being Mexican-American was also associated with refusal (39). The potential applications of the findings from research relating to the possible reasons for nonresponse are promising, as they could facilitate the design of future surveys, and clarify in what stage of the survey the researcher should make additional efforts to prevent nonresponse.

Effect of response rate on measures of association

A major concern about nonresponse in epidemiologic surveys derives from the potential to bias the estimation of risks in any kind of study. In other words, there would be concern about whether the participants represent the source population from which they were originally selected (7).

In a case-control study, if the response rate is lower than expected and the response status depends either on the case status (case or control) or the exposure status (exposed or not exposed), the estimation of risk will not change.

Suppose that for the following 2x2 table, the true odds ratio (OR) is (a*d)/(b*c),

	Case	Control
Exposed	a	b
Not Exposed	c	d

then, if the proportion of response varies according to the disease status, with only a proportion (*p1*) of cases and controls, the new table will be,

	Case	Control
Exposed	*p1* a	*p1* b
Not Exposed	*p1* c	*p1* d

and the resulting OR will be the same (a*d)/(b*c), since the proportion *p1* of response will cancel in the equation.

If the response rate in cases is different from controls, with a proportion *p1* for cases and *p2* for controls,

	Case	Control
Exposed	*p1* a	*p2* b
Not Exposed	*p1* c	*p2* d

then the resulting OR will be the same (a*d)/(b*c), since the proportions *p1* and *p2* of response will cancel one another in the equation.

If the response rate in the exposed subjects is different from the not exposed subjects, with a proportion *p1* for exposed and *p2* for not exposed,

	Case	Control
Exposed	*p1* a	*p1* b
Not Exposed	*p2* c	*p2* d

then, the resulting OR will still be the same (a*d)/(b*c), since the proportions *p1* and *p2* of response will cancel one another in the equation. Whether there is a similar or a differential response or participation, the magnitude of response does not affect the estimate of risk. In order to bias the OR, the proportion of response should differ on the exposure and the outcome, giving origin to at least three different probabilities of participation.

The cohort design has specific problems. If the relative risk (RR) in a population sample is

	Disease	Not Disease	Total
Exposed	a	B	N1
Not Exposed	c	D	N0

then, as shown in the above table the RR will be (a/N1)/(c/N0)

In a cohort study, the differential response between the exposed subjects and the not-exposed can have a different effect: If the response rate in exposed is different from not-exposed, with a proportion *p1* for exposed and without changes for those not exposed,

	Disease	Not Disease	Total
Exposed	*p1* a	*p1* b	*n1*
Not Exposed	c	D	N0

then the resulting RR can change (*p1*a/ *n1*)/(c/ N0), since the proportion of the response will change only the numerator in the equation.

In both types of designs, cohort and case-control, a complete certainty of the outcome and exposure status is required in order to assess and detect any bias related to participation and response, or to follow-up. This type of information is rarely available, so this is the reason to develop preventive strategies during the design of the study to help to guarantee a high participation and retention of participants, with the expectation that this will decrease the chances of response bias.

Although the theoretical frame for the presence of a response bias is clear, the limited empirical information available to evaluate the association between response rate and bias has failed to prove an association between nonresponse rate and nonresponse bias (2, 7, 12, 40-4). This is a very controversial finding, and probably reflects the difficulties in gathering the relevant data needed to understand all the consequences of nonresponse. For this reason, the comparison between respondents and nonrespondents is usually based on the information available, usually demographics. When the distribution of a variable or characteristic in the population of interest is unknown, and it happens to be a variable related to the exposure or outcome of interest, it is almost impossible to measure the bias effect. When additional characteristics of the population have been collected, and mathematical simulations have been run to evaluate the effect of the nonresponse proportion over nonresponse bias, some simulations have found a bias effect or, at least, a potential to bias the estimates (16, 45, 46). These contradicting findings are an example of the limited information we have in this area and another reason to call for additional sounded research in the nonresponse field (3).

Description of the Environmental Risk, Coping, and Hispanic Health Study

The Environmental Risk, Coping, and Hispanic Health Study, conducted by the Center for Population Health and Health Disparities of the University of Texas Medical Branch, is a population-based survey designed to evaluate the relation between environmental stress and stress-related diseases and biological markers of stress in Texas City, Texas. The study included a census of about 11, 000 households. Ethnicity and age

was ascertained within each household to select three ethnic strata: Mexican Americans 25 to 64 years of age, non-Hispanics, and Hispanics 65 years of age and older.

One in eight households and one adult in non-Hispanic households having at least one adult age 25 and over were selected. For Hispanic households, all adults age 65 and older were selected. For Hispanic households with no adults age 65 and older, one adult was selected. In this way a total of 3,428 adults between 25-64 years of age were eligible for interview of which 2,706 were interviewed at home and invited to participate in a health exam and donate blood to measure biologic markers of stress.

The interviewing followed standard U.S. Census Bureau Current Population Survey methods (U.S. Census, 2002) with appropriate local modifications. When a Hispanic household or 8th non-Hispanic housing unit was identified, the project was summarized and an oral informed consent for enumeration was sought. For households that consented, adults were enumerated, and one adult age 25 or older was selected at random using the Kish selection method. If the selected individual was present, he or she was invited to participate in the survey and an additional oral informed consent is obtained. Once the participant had consented to participate, a blood pressure and pulse reading was obtained along with height and weight. Each interview was completed within one hour to one hour and 15 minutes, and respondents received a $15 compensation grocery store card after successfully completing the interview. Two versions of the questionnaire were available (English and Spanish) and participants were allowed to choose between these. Information from the questionnaire provided demographic data that included age gender, race, ethnic group, marital status, income, education, health insurance and health habits and behaviors. Also included was data on physical activity, smoking, alcohol use, perceived health and stress, and environmental stress.

Blood was collected from the surveyed population 2-3 weeks following, the initial in-home interview. With prior written consent the blood was collected before 12 pm at the subject's residence or at the local clinic by a trained nurse or phlebotomist. The following additional information was obtained before the blood draw: fasting status, infection status, certification that the participant was disease free, and a list of current medications, with dose and frequency. Blood pressure, pulse, height, and weight were also measured. A total of 35 ml of blood were drawn to measure inflammatory and stress markers such as TNF - a, CRP, IL-6, cortisol; and antibody titers that include HSV -1, Latent EBV-capsid antigen (VCA), early antigen (EA), and EBV nuclear antigen (EBNA). In addition to monetary compensation in the form of a gift card worth $30, each participant was provided with additional readings for glucose, total hemoglobin, hemoglobin A1C, %A1C, high density lipids, low density lpids, total cholesterol and triglycerides. The selected strategies of offering the two options (home or clinic), a monetary compensation, and the time lag between the interview and the blood collection could affect the response proportion (47), but as they were not randomly applied, they are not included as factors in the current analysis.

This survey design presented an unique opportunity to address questions relevant to factors related to nonresponse in a tri-ethnic community. First, the survey was rich in Hispanic and Non-Hispanic Black and White subjects, residing in the same neighborhood, and it provided data to potentially compare the factors related to nonresponse across these cultural/ethnic groups. Second, as information about scores in different environmental and emotional stress scales was obtained during the interview, the relation between the exposures of interest and the response status could be evaluated in this population, which was an initial step in the evaluation of the response bias.

The objectives of the current project were: a) to determine the percentage of nonparticipation (nonresponse) in the serological part of the survey, defined as response to the initial interview but failure to provide a blood sample for analysis of stress biomarkers, b) to compare the demographic, health risks and behavior, and disease profile between respondents and nonrespondents in the overall population and across different ethnic groups, and c) to examine the association between emotional and environmental stress and response status by ethnicity in the Environmental Risk, Coping, and Hispanic Health Study.

The main areas of professional development and public health competences related to MPH training that were addressed in this project included:

a) Apply epidemiological principles and methods to describe the distribution and occurrence of disease in human populations.

b) Demonstrate statistical thinking in formulating and answering questions, and employ statistical methods to draw quantitative inferences and construct formal models in biomedicine and human populations.

c) Assess social and behavioral determinants of health.

d) Consistently employ high ethical and professional standards in public health practice and research activities.

One of the end-products of this capstone was a report to the original project Principal Investigators and the research community about the characteristics of the studied population that need special attention in order to improve participation in future surveys.

RESEARCH DESIGN AND METHODS

Study Design

A cohort design was used to analyze data from the Environmental Risk, Coping, and Hispanic Health Study. The current Capstone sought to identify demographic, socioeconomic, and health factors related to nonresponse to the serum collection part of

the survey. The outcome of interest was the response status. The data collected during the interview were the independent variables that were explored.

Dependent Variable

The main outcome was "nonresponse" to the serum survey, defined at the individual level as participation in the interview but failure to subsequently donate blood for biologic measures for the second stage of the survey. This definition of nonresponse in Health Examination Surveys has been used in projects with similar design, as the Hispanic Health and Nutrition Examination Survey (30 and the ARIC (25). In both, the initial interview was followed by an invitation to donate blood samples for biological measures.

Independent Variables

The independent variables were selected out of the information collected during the interview, based on the theoretical frame of nonresponse (that included the relationship with demographics, socioeconomic status, and health), and on the relationship with the outcomes of interest of the survey (stress).

Demographic Variables

Age was self-reported by each individual, and was recorded in years. Ethnicity was also self-reported by each participant, and for those who described themselves as Hispanic, an additional question inquired about the place of birth, and later the subject was classified as U.S.-born Hispanic, or Foreign-born Hispanic. Marital status (married, not married) and health insurance status were self-reported. The educational level was described by each individual as number of years of education completed, and was recoded as "less than high school", and "completed high school or higher". The data on household income were categorized using the census categories, and recoded as "lower than $20,000 per year", or "$20,000 or more per year".

Health Habits and Behaviors

Moderate exercise was defined as a minimum of 30 minutes at least five days a week. Smoker status was defined by the response to the question "Do you now smoke cigarettes everyday, some days, or not at all". Drinking alcohol was based on the answer to the question "Have you ever drunk any type of alcohol?"

Perceived Health and Disease Profile

The disease profile was based on the response to questions about the presence or absence of diabetes, hypertension, heart attack, stroke, respiratory disease (asthma, bronchitis or emphysema), and cancer. The subject also described his or her health in response to the question "In general would you say your health is excellent, very good, good, fair, or poor". This is part of the SF-36 questionnaire, and reclassified in "good health" for those that answered "excellent, very good, or good", or reclassified as "poor" for those answering "fair or poor".

Stress Measures

The measures of stress that were collected included the score on the Cohen's Perceived Score Scale, an instrument that has been developed and validated in a representative sample of the American population, and used in similar projects (37).
.

Study Population

The study population for this analysis included individuals who participated in the original Environmental Risk, Coping, and Hispanic Health Study. In brief, sampling unit consisted of two groups. The first was based on a census of every Hispanic and the second was every 8th non-Hispanic housing unit. The study unit was a random selection of Hispanic subjects between the ages of 25 and 64, and additionally all Hispanics older than 64. Finally, a one in eight random sample of Non-Hispanic subjects above 25 years of age was included. Participants in the epidemiologic survey, those who agreed to be interviewed, were the population subject of the current analysis, and were classified according to their participation in the serum part of the survey as respondents (those who agreed to donate blood for biomarker analysis), and nonrespondents (those who did not donate blood).

Plan of Analysis

The specific aim of this Capstone was to test the hypothesis that individuals reporting low socioeconomic status, poor health habits, and higher disease burden are less likely to participate in serum sample surveys.

Project objectives were

(a) To determine the percentage of non-participation in blood sample collection: Frequencies of participation in both the epidemiological study and serum sample collection study were compared to participation in only the epidemiological study.

(b) To compare the demographic characteristics, risk factors, perceived health status, and certain specificreported diseases of those who donated a blood sample and those who declined participation in the blood sample collection. The frequency of nonresponse, gender (male/female), ethnic group, education (below high school versus completed high school and higher), marital status (married or not married), health insurance status (insured versus not insured), income (lower versus equal to or more than $20,000 per year) were described as proportions. Smoking status and exercise were also described as proportions, and the same was done for the presence or absence of diseases (diabetes, hypertension, respiratory disease, heart attack, stroke and cancer). The perceived health status (good or poor self-rated health) was also described as proportion. Age and the score on the stress scales were analyzed as continuous variables. The groups defined by the dependent variables (response status, "respondent" and "nonrespondent", as defined above) were compared on demographic characteristics, health status, habits, disease profile, and stress scales, using differences of proportions, means, and variances, according to the variable of interest.

(c) To examine the influence of ethnicity on participation in blood collection. Models were compared by ethnicity (Non-Hispanic Whites, Non-Hispanic Blacks, U.S.-born Hispanics, and Foreign-born Hispanics).

(d) To analyze the relationship between participation in serum sample collection and stress. The association between response and stress was examined, , while controlling for the demographic, health status, and disease profile, and was evaluated with logistic regression.

For all analyses, the significance level was set at $p<0.05$. All computer programming and analysis were completed using SPSS version 13.

Results

AIM 1

To determine the percentage of nonparticipation (nonresponse) in the blood collection stage of the survey, defined as response to the initial interview but failure to provide a blood sample for analysis of stress biomarkers.

A total of 2,706 persons were interviewed and answered the survey questionnaire. Of these, 58.4% were females, and the mean age for the population at interview time was 48.87 years. Non-Hispanic Whites and U.S.-born Hispanics formed the majority of participants (34.1%, and 36.1%, respectively), with a lower presence of Foreign-born Hispanics (17.0%), and Non-Hispanic Blacks (12.0%). Almost two-thirds had a high school degree or higher education (62.0%), and more than half of the participants were married at the time of the interview (56.7%). One out of four participants lived in a household with annual income lower than $ 20,000 (25%), and 59.6% had some type of health insurance. Only 18.7% of the participants engaged regularly in moderate physical activity, and almost one out of three (29.5%) were smokers. The most frequently reported disease was hypertension (36.4%), followed by respiratory diseases (18.7%) and diabetes mellitus (16.5%). Only 29.9% described their health as poor. Out of the total 2,706 participants interviewed, 1,271 (47%) declined the invitation to donate blood for the biomarker study, and were, thus, considered nonrespondents. The results are summarized in Table 1.

Table 1.- Characteristics of study subjects

Variable	Frequency n (%)
Demographic Characteristics	
Gender	
Male	1125 (41.6%)
Female	1581 (58.4%)
Age	
Years [Mean (sd)]	48.87 (+/- 16.0)
Ethnicity	
Non-Hispanic White	924 (34.1%)
US-born Hispanic	976 (36.1%)
Foreign-born Hispanic	459 (17.0%)
Non-Hispanic Black	325 (12.0%)
Socioeconomic Status	
Education	
Less than high school	1019 (37.7%)
Completed high school and higher	1679 (62.0%)
Marital Status	
Married	1533 (56.7%)
Not married	1172 (43.3%)
Annual Family Income	
Less than 20,000 / year	677 (25.0%)
20,000 or higher / year	1733 (64.0%)
Health Insurance	
Insured	1613 (59.6%)
Not Insured	1093 (40.4%)
Health Status and Habits	
Smoking Status	
Smoker	798 (29.5%)
Nonsmoker	560 (20.7%)
Physical Activity	
Moderate physical activity	506 (18.7%)
Not physically active	2200 (81.3%)
Self-rated Health	
Good health	1897 (70.1%)
Poor health	809 (29.9%)
Disease Profile	
History of diabetes mellitus	446 (16.5%)
History of hypertension	985 (36.4%)
History of respiratory disease	506 (18.7%)
History of cancer	263 (9.7%)
History of coronary disease	169 (6.2%)
History of cerebrovascular disease	139 (5.1%)
Response Status	
Respondents	1435 (53.0%)
Nonrespondents	1271 (47%)

AIM 2

To compare the demographic, health risks and behaviors, and disease profile of those who donated a blood sample and those who declined participation (nonrespondents) in blood sample collection.

Table 2 shows the comparison between respondents and nonrespondents. The main demographic differences were in gender and age. Nonresponse was associated with male gender and younger age. In comparison with Non-Hispanic Whites, the response was lower for the other ethnic groups. None of the socioeconomic variables evaluated (income, health insurance, education, and marital status) was associated with nonresponse. Being a smoker was associated with nonresponse, while being in poor health was associated with a lower frequency of nonresponse, and the participation in physical activity did not show any relationship with the response status. In general, the presence of any disease was associated with a lower probability of being nonrespondent.

Table 3 shows a logistic model for nonresponse in the total sample. The only significant demographic variable associated with nonresponse was age at interview (with decreasing probability of being nonrespondent as age increases). There was interaction between age and ethnicity. With the exception of Non-Hispanic Blacks, for all other ethnic groups advanced age was related to lower likelihood of nonresponse. The only association with a socioeconomic factor was with family income. If annual household income was lower than $20,000, there was lower probability of being nonrespondent. Smoking was associated with nonresponse, and of the evaluated self-reported diseases, only hypertension had a negative association with nonresponse.

Table 2.- Characteristics of study subjects, according to response status, and relationship with Nonresponse

Variable	Respondents n = 1435	Nonrespondents n = 1271	Univariate OR (95% CI)
Demographics			
Gender			
Male	550 (50.7%)	555 (49.3%)	1.18 (1.01, 1.18)
Female	865 (54.7%)	716 (45.3%)	Reference
Age			
Years [Mean (sd)]	51.75 (15.97)	45.62 (15.4)	4.94, 7.31 *
Ethnicity			
US-born Hispanic	531 (54.4%)	445 (45.6%)	1.22 (1.02, 1.47)
Foreign-born Hispanic	211 (46.0%)	248 (54.0%)	1.58 (1.25, 1.99)
Non-Hispanic Black	149 (45.8%)	176 (54.2%)	1.59 (1.22, 2.07)
Non-Hispanic White	530 (57.4%)	394 (42.6%)	Reference
Socioeconomic Status			
Education			
Less than high school	533 (52.3%)	486 (47.7%)	1.05 (0.89, 1.23)
High school and higher	897 (53.4%)	782 (46.6%)	Reference
Marital Status			
Not married	615 (52.5%)	557 (47.5%)	1.04 (0.89, 1.21)
Married	819 (53.4%)	714 (46.6%)	Reference
Annual Family Income			
Less than $20,000 / year	385 (56.9%)	292 (43.1%)	0.85 (0.70, 1.02)
$20,000 or higher / year	914 (52.7%)	819 (47.3%)	Reference
Health Insurance			
Not insured	568 (52.0%)	525 (48.0%)	1.07 (0.92, 1.26)
Insured	867 (53.8%)	746 (46.2%)	Reference
Health Status and Habits			
Smoking Status			
Smoker	382 (47.9%)	416 (52.1%)	1.80 (1.44, 2.26)
Nonsmoker	349 (62.3%)	211 (37.7%)	Reference
Physical Activity			
Not physically active	1182 (53.7%)	1018 (46.3%)	0.86 (0.71, 1.05)
Moderate physical activity	253 (50.0%)	253 (50.0%)	Reference
Self-rated health			
Poor health	453 (56.0%)	356 (44.0%)	0.84 (0.71, 1.00)
Good health	982 (51.8%)	915 (48.2%)	Reference
Disease Profile			
Diabetes	282 (63.2%)	164 (36.8%)	0.61 (0.49, 0.75)
Hypertension	593 (60.2%)	392 (39.8%)	0.63 (0.54, 0.74)
Respiratory disease	304 (60.1%)	202 (39.9%)	0.70 (0.57, 0.86)
Cancer	156 (59.3%)	107 (40.7%)	0.75 (0.58, 0.98)
Coronary disease	115 (68.0%)	54 (32.0%)	0.51 (0.36, 0.72)
Cerebrovascular disease	84 (60.4%)	55 (39.6%)	0.73 (0.51, 1.05)

*95% CI difference of means.

Table 3.- Logistic Model of the Relationship Between Nonresponse, Demographics, Socioeconomic Status, and Health

Variable	n = 2706	Exp (ß)
Constant		2.60
Demographics		
Gender		
Female	1581 (58.4%)	Reference
Male	1125 (41.6%)	1.08 (0.84, 1.38)
Ethnicity		
Non-Hispanic White	924 (34.1%)	Reference
US-born Hispanic	976 (36.1%)	0.94 (0.62, 1.41)
Foreign-born Hispanic	459 (17.0%)	0.82 (0.54, 1.26)
Non-Hispanic Black	325 (12.0%)	1.12 (0.65, 1.93)
Age by Ethnicity		
Age by Non-Hispanic White		0.60 (0.36, 0.99)
Age by US-born Hispanic		0.23 (0.10, 0.55)
Age by Foreign-born Hispanic		0.61 (0.20, 1.87)
Age by Non-Hispanic Black		1.63 (0.59, 4.53)
Socioeconomic Status		
Education		
Completed high school and higher	1679 (62.0%)	Reference
Less than high school	1019 (37.7%)	1.02 (0.79, 1.33)
Annual Family Income		
$20,000 or higher / year	1733 (64.0%)	Reference
Less than $20,000 / year	677 (25.0%)	0.74 (0.55, 0.98)
Health Status and Habits		
Smoking Status		
Nonsmoker	560 (20.7%)	Reference
Smoker	798 (29.5%)	1.45 (1.12, 1.89)
Self-rated health		
Good health	1897 (70.1%)	Reference
Poor health	809 (29.9%)	1.22 (0.93, 1.60)
Perceived Stress		0.98 (0.96, 0.99)
Disease Profile		
Diabetes	446 (16.5%)	0.87 (0.61, 1.25)
Hypertension	985 (36.4%)	0.72 (0.54, 0.96)
Respiratory disease	506 (18.7%)	0.77 (0.57, 1.05)
Cancer	263 (9.7%)	1.08 (0.72, 1.60)
Coronary disease	169 (6.2%)	0.70 (0.40, 1.21)
Cerebrovascular disease	139 (5.1%)	0.90 (0.50, 1.61)

AIM 3

To examine the influence of ethnicity on nonresponse to the serum survey.

Separate logistic models of nonresponse were constructed for each ethnicity (Table 4). The only factor related across all ethnic groups with nonresponse was age. No additional factor showed a significant association with nonresponse for Non-Hispanic Whites and Foreign-born Hispanics. For those described as U.S.-born Hispanics, in addition to age, to be a smoker and report poor health were associated with nonresponse, while a low income was associated with less likelihood of being nonrespondent. Besides age, the model for Non-Hispanic Blacks found a significant association with hypertension, since reporting this disease decreased the likelihood of nonresponse.

AIM 4

To analyze the relationship between stress and nonresponse to the serum survey.

Perceived stress score was included in the model for the all sample, and in the separate logistic models of nonresponse constructed for each ethnicity (Tables 3-4). The stress score did not show association with response status.

Table 4.- Logistic Models of Nonresponse by Ethnic Group

Characteristic	Non-Hispanic Whites	US-Born Hispanic	Foreign-born Hispanics	Non-Hispanic Blacks
Constant	3.24	2.36	1.14	6.47
Demographics				
Gender				
Male	0.94	1.35	1.56	0.48
	0.64, 1.37	0.88, 2.06	0.65, 3.73	0.19, 1.18
Female	Reference			
Age				
Years	0.97	0.97	0.98	0.97
	0.96, 0.98	0.95, 0.98	0.95, 1.01	0.94, 1.00
Socioeconomic Status				
Education				
Less than high school	1.08	0.95	0.73	2.25
	0.70, 1.67	0.62, 1.45	0.31, 1.75	0.91, 5.55
Completed high school	Reference			
Annual Family Income				
Less than $20,000 / year	0.79	0.54	1.16	0.73
	0.50, 1.25	0.33, 0.90	0.49, 2.77	0.31, 1.73
$20,000 or higher / year	Reference			
Health Status and Habits				
Smoking Status				
Smoker	1.13	1.77	2.60	1.02
	0.74, 1.71	1.15, 2.71	1.13, 5.95	0.41, 2.55
Nonsmoker	Reference			
Self-rated health				
Poor health	1.06	1.64	0.76	0.86
	0.69, 1.62	1.03, 2.60	0.29, 2.04	0.35, 2.07
Good health	Reference			
Perceived Stress				
Perceived Stress Score	0.98	0.97	1.00	1.00
	(0.95, 1.01)	(0.94, 1.00)	(0.94, 1.00)	(0.95, 1.06)
Disease Profile				
Diabetes	1.02	0.73	0.73	1.36
	0.60, 1.76	0.40, 1.35	0.14, 3.81	0.45, 4.14
Hypertension	0.82	0.85	1.16	0.28
	0.54, 1.25	0.50, 1.45	0.39, 3.42	0.11, 0.69
Respiratory disease	0.84	0.58	0.89	0.98
	0.55, 1.28	0.32, 1.07	0.22, 3.46	0.39, 2.47
Cancer	1.00	1.33	2.02	0.87
	0.58, 1.73	0.63, 2.78	0.31, 13.09	0.22, 3.43
Coronary disease	0.85	0.45	1.43	0.53
	0.41, 1.77	0.14, 1.38	0.04, 46.7	0.11, 2.54
Cerebrovascular disease	0.67	1.04	0.70	2.28
	0.28, 1.57	0.36, 3.02	0.03, 12.5	0.49, 10.5

Discussion

In a recent review of participation rates in general surveys and epidemiologic surveys, Galea et al. showed that participation is steadily declining (1). In surveys that include an interview and a limited physical examination or the collection of blood samples, usually described as "health surveys", the participation rate in the examination and serum surveys varied between 76% for NHANES 2003-2004 (32), and 66% for NYC-HANES (31). Our data show that the participation rate in the Environmental Risk, Coping, and Hispanic Health Study was 53%, which was lower than that described in similar surveys.

In the current project, we found differences between individuals who were participants and those who were not participants in demographic characteristics (males, younger people, and ethnic minorities were more frequently nonrespondents), health habits (with smoking being associated with nonresponse), and health status (nonrespondents had a lower frequency of specific diseases). The lower frequency of response in males and minorities agrees with some of the expected differences under the "social disengagement" hypothesis (4-14), while better health status and fewer diseases in nonrespondents does not agree with that explanation. As most of the differences between respondents and nonrespondents are based on the analysis of epidemiologic surveys that do not include health examinations or the provision of biologic samples, it is probable that the "social disengagement" and "worried healthy" models have limited applicability to this kind of survey.

The presence of demographic differences between participants and nonparticipants in the serum survey was expected, based on the information from many general epidemiologic surveys (4-14), where nonrespondents also tended to have a higher burden of disease. It is unclear whether the same happens in surveys that include a physical examination and biological samples. Information from the NHANES II (35) partially supports the existence of demographic and socioeconomic differences, but the differences are limited to employment status, less interest in discussing health problems, and age. Also, the evaluation of the converted refusal of the physical examination in NHANES (36, 37) showed that only smoking habit status was different between full respondents and converted refusals. Multiple factors could be related with lack of participation in the second stage (physical examination and serum sampling) of health surveys; however, as not many projects involved an initial interview with a later physical examination and/or biological sampling, there is very limited information in this area. In our population, the differences between participants who agreed to be part of the serum survey and those who did not, have both similarities and differences with what is reported in the literature. Some nonrespondent characteristics are similar to the converted refusals from the NHANES 2003-2004 (a high frequency of smokers) (36, 37), but differ from the NHANES II (32), which found nonrespondents to be older than respondents. Our

findings also agree with the genetic consent component of NHANES (39), where the participation rate was usually lower for ethnic minorities, compared with Non-Hispanic Whites. However, the differences in participation according to ethnic group were not maintained after a multivariate analysis, while the relationship with smoking status was maintained, and an annual family income lower than $20,000 also appeared to be related to nonresponse. Of special interest is the fact that the only health-related variable associated with the response status was the presence of hypertension, while the other diseases and self-rated health did not have a relationship with the decision to participate in the serum survey. There could be multiple reasons why we were not able to find significant differences between those who were respondents and those who were nonrespondents to the serum portion of the survey. One explanation is that respondents and nonrespondents to the serum survey were not very different from one another, and that the main differences between respondents and nonrespondents in epidemiologic surveys are between those who decide to participate in one way or another (and be included in the first step of the process, the interview), and those who were nonrespondent to the initial invitation. Our findings are in agreement with this explanation.

In our separate analysis of the differences between respondents and nonrespondents across ethnic groups, the only common factor for all groups of nonrespondents was a younger age. No other significant differences were found for Non-Hispanic Whites and Foreign-born Hispanics. For Non-Hispanic Blacks the absence of hypertension was related to nonresponse. Additional differences between respondents and nonrespondents were found in U.S.-born Hispanics, including annual household income (higher income associated with nonresponse), smoking habit, and poor perceived health (associated with nonresponse). The intra-ethnic group differences in respondents and nonrespondents agree with what has previously been reported for Non-Hispanic Blacks and Foreign-born Hispanics. Participants and nonparticipants of these ethnic groups did not have great differences between them (19-27), which is similar to what we found here. We still lack an explanation for our findings in minority groups, but our findings do agree with other studies that have found low response in Non-Hispanic Blacks, probably as a symptom of the persistent distrust of medical research due to negative historic experiences, such as the Tuskegee Project (21). For Non-Hispanic Whites the findings are different from what would be expected, based on the experience from general epidemiologic surveys that have provided the foundation for the development of the "worried healthier" explanation for participation and response. Another interesting finding was that the presence of differences between respondents and nonrespondents in the U.S.-born Hispanic group, closely resembles what has been described as the typical profile of the nonrespondent. Again, most of the information about nonresponse and the characteristics of nonrespondents has come from surveys without a physical examination or serum sample component, a factor that limits the comparison to our findings. Previous efforts to compare the characteristics of nonrespondents and respondents in health examination surveys have not included ethnicity-specific analysis, with the exception of one report form the ARIC project (25), but the only comparison in the ARIC analysis was between Non-Hispanic Blacks and Non-Hispanic Whites.

Since multiple factors, beyond the demographic, socioeconomic, and health factors of the participant or the target population, are related to the response rate in surveys (47), we can not exclude the role of "structural survey" factors, including the type of incentives offered, the time window offered in which to donate the samples, and the different locations available for donating the samples. These additional factors were not included in our analysis, which was focused on socioeconomic and health factors at the individual level, not with the survey results or the instrument themselves.

Our analysis of the differences between respondents and nonrespondents to the serum survey part of a health examination study showed different participation rates across ethnic groups, but did not show significant differences between those respondents to the interview that agreed to donate blood for biological measures (the serum survey) and those who did not atten. The only consistent difference is the age, with older persons having less likelihood of being nonrespondents. Yhis was seen in all ethnic groups, except Non-Hispanic Blacks. The degree of stress did not have a relationship with the response status.

Conclusion

In conclusion, we found that participation in the serum collection part of this Texas City study was lower than in similar surveys. Response rates varied by ethnicity. We found lower response rates in minority groups, a finding that agrees with the current literature in the field. We did not find great differences between respondents and nonrespondents, but we have validated previous findings that smoking habit, male gender, and perceived poor health are characteristics related to a higher likelihood of nonresponse. Unexpectedly, the association with age was in the opposite direction than that seen in previous studies, with nonresponse being associated with younger age. The only factors that were maintained in a multivariate analysis were age, family income, and smoking habit. We also found an interaction between age and ethnicity, with advanced age being related to less likelihood of nonresponse in all ethnicities except in Non-Hispanic Blacks, where the direction was the opposite. A comparison of respondents and nonrespondents across ethnic groups showed that the typical description of the participant as "worried healthy" or "worried well" was valid only for the U.S.-born Hispanics, in whom the differences between respondents and nonrespondents included age, smoking habit, poor health, and income (low income was associated with less likelihood of being nonrespondent), while for the remaining three ethnic groups no significant differences, besides age, were found between respondents and nonrespondents. Finally, the score on the perceived stress scales did not have any association with response, for either the entire sample or across ethnic groups.

Our findings do not support either the "worried healthy" explanation of response, neither the "social disengagement" explanation of nonresponse to the serum collection portion of a health examination survey. They are in agreement with data from similar health surveys, and do not support the suspicion that there might be a "response bias" in this kind of study. Our findings do not apply to the first part of the survey, the interview, and they still need to be replicated in different settings.

References

1. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol 2007; 17: 643-653.
2. Austin MA, Criqui MH, Barret-Connor E, Holdbrook MJ. The effect of response bias on the odds ratio. Am J Epidemiol 1981; 114: 137-143.
3. Stang A. Nonresponse research, an underdeveloped field in epidemiology. Eur J Epidemiol 2003; 18: 929-931.
4. Comstock GW, Helsing KJ. Characteristics of respondents and nonrespondents to a questionnaire for estimating community mood. Am J Epidemiol 1973; 97: 233-239.
5. Slymen DJ, Drew JA, Wright BL, Elder JP, Williams SJ. Compliance with a 12 month assessment in an elderly cohort participating in a preventive intervention study: The San Diego Medicare Preventive Health Project. Int J Epidemiol 1992; 21: 701-706.
6. Tell GS, Fried LP, Hermanson B, Manolio TA, Newman AB, Borhani NO. Recruitment of adults 65 years and older as participants in the Cardiovascular Health Study. Ann Epidemiol 1993; 3: 358-366.
7. Shahar E, Folsom AR, Jackson R. The effect of nonresponse on prevalence estimates for a referent population: Insights from a population-based cohort study. Ann Epidemiol 1996; 6: 498-506.
8. Zaslavsky AM, Zaborski LB, Cleary PD. Factors affecting response rated to the Consumer Assessment of Health Plans Study Survey. Medical Care 2002; 40: 485-499.
9. Chou P, Kuo H-S, Chen C-H, Lin H-C. Characteristics of nonparticipants and reasons for non-participation in a population survey in Kin-Hu, Kinmen. Eur J Epidemiol 1997; 13: 195-200.
10. van der Akker M, Buntinx F, Metsemakers JFM, Knottnerus JA. Morbidity in responders and non-responders in a register-based population survey. Family Practice 1998; 15: 261-263.
11. Launer LJ, Wind AW, Deeg DJ. Nonresponse pattern in a community-based cross-sectional study of cognitive function in the elderly. Am J Epidemiol 1994; 139: 803-812.
12. Korkeila K, Suominen S, Ahvenainen J, Ojanlatva A, Rautava P Helenius H, Koskenvou M. Non-response and related factors in a nation-wide health survey. Eur J Epidemiol 2001; 17: 991-999.
13. Eagan TML, Eide GE, Gulsvik A, Bakke PS. Nonresponse in a community cohort study. Predictors and consequences for exposure-disease association. J Clin Epidemiol 2002; 55: 775-781.
14. Richiardi L, Boffetta P, Merletti F. Analysis of nonresponse bias in a population-based case-control study on lung cancer. J Clin Epidemiol 2002; 55: 1033-1040.
15. Markides KS, Dickson HD, Pappas C. Characteristics of dropouts in longitudinal research on aging: a study of Mexican-Americans and Anglos. Exp Aging Res 1982; 8: 163-167.

16. Wild TC, Cunningham J, Adlaf E. Nonresponse in a follow-up to a representative telephone survey of adult drinkers. J Stud Alcohol 2001; 62: 257-261.
17. Brennan M, Hoek J. The behavior of respondents, nonrespondents, and refusers across mail surveys. Public Opin Quarterly 1992; 56: 530-535.
18. Mond JM, Rodgers B, Hay PJ, Owen C, Beaumont PJ. Nonresponse bias in a general population survey of eating disorderer behavior. Int J Eat Disord 2004; 36: 89-98.
19. Durant RW, Davis RB, St. George DMM, Williams IC, Blumenthal C, Corbie-Smith GM. Participation in research studies: factors associated with failing to meet minority recruitment goals. Ann Epidemiol 2007; 17: 634-642.
20. Link MW, Mokdad AH, Stackhouse HF, Flowers NT. Race, ethnicity, and linguistic isolation as determinants of participation in public health surveillance surveys. Prev Chronic Dis [serial online] 2006 Jan. Accessed on 3/20/2008. Available at http://www.cdc.gov/pcd/issues/2006/jan/05_0055.htm.
21. Corbie-Smith G, Thomas SB, Williams MV, Moody-Ayers S. Attitudes and beliefs of African Americans toward participation in medical research. J Gen Intern Med 1999; 14: 537-546.
22. Shavers VL, Lynch CF, Burmeister LF. Racial differences in factors that influence the willingness to participate in medical research studies. Ann Epidemiol 2002; 12: 248-256.
23. Vernon SW, Roberts RE, Lee ES. Ethnic status and participation in longitudinal health surveys. Am J Epidemiol 1984; 119: 99-113.
24. Psaty BM, Cheadle A, Koepsell TD, Diehr P, Wickizer T, Curry S, VonKorff M, Perrin EB, Pearson DC, Wagner EH. Race- and ethnic-specific characteristics of participants lost to follow-up in a telephone cohort. Am J Epidemiol 1994; 140: 161-171.
25. Jackson R, Chambless LE, Yang K, Byrne T, Watson R, Folsom A, Shahar E, Kalsbeek W. Differences between respondents and nonrespondents in a multicenter community-based study vary by gender and ethnicity. J Clin Epidemiol 1996; 49: 1441-1446.
26. Sweeney C, Edwards SL, Baumgartner KB, Herrick JS, Palmer LE, Murtaugh MA, Stroup A, Slattery ML. Recruiting Hispanic women for a population-based study: validity of surname search and characteristics of nonparticipants. Am J Epidemiol 2007; 166: 1210-1219.
27. Ashing-Giva KT, Padilla GV, Tejero JS, Kim J. Breast cancer survivorship in a multiethnic simple, challenges in recruitment and measurement. Cancer 2004; 101: 450-465.
28. Oropesa RS, Landale NS. Nonresponse in follow-back surveys of ethnic minority groups: an analysis of the Puerto Rican maternal and infant health study. Maternal Child Health J 2002; 6: 49-58.
29. Veenstra MY, Friesema IHM, Zwietering PJ, Garretsen HF, Kanottnerus JA, Lemmens PH. Lower prevalence of heart disease but higher mortality risk during follow-up was found among nonrespondents to a cohort study. J Clin Epidemiol 2006; 59: 412-420.

30. The Hispanic Health and Nutrition Examination Survey (HHANES), Description. Accessed on04/10/2008. Available from http://www.cdc.gov/nchs/products/elec_prods/subject/hhanes.htm. Accessed on 04/10/2008.
31. Thorpe LE, Gwynn RC, Mandel-Ricci J, Roberts S, Tsoi B, Berman L, Porter K, Ostchega Y, Curtain LR, Montaquila J, Mohadjer L, Frieden TR. Study design and participation rates of the New York City Health and Nutrition Examination Survey, 2004. Prev Chron Dis 2006. Accessed on 4/10/2008, available from http://www.cdc.gov/pcd/issues/2006/jul/05_0177.htm.
32. NHANES, Response rates. Accessed on 4/10/2008, available from http://www.cdc.gov/nchs/about/major/nhanes/nhanes_cps_totals.htm.
33. Woteki CE. The Hispanic Health and Nutrition Examination Survey (HHANES 1982-1984): background and introduction. Am J Clin Nutr 1990; 51: 897S-901S.
34. Hankinson SE, Willett WC, Manson JE, Colditz GA, Hunter DJ, Spiegelman D, Barbieri RL, Speizer FE. Plasma sex steroid hormone levels and risk of breast cancer in postmenopausal women. J Natl Cancer Inst 1998; 90: 1292-1299.
35. Forthofer RN. Investigation of nonresponse bias in NHANES II. Am J Epidemiol 1983; 117: 507-515.
36. Chong Y, Carroll M, Burt V, Montalvan P. NHANES converted refusals: are they different from willing respondents in cardiovascular risk factors? Paper presented at the annual meeting of the American Association for Public Opinion Association, Fontainebleu Resort, Miami Beach, Fl. Accessed on 4/10/2008, available from http://www.allacademic.com/meta/p16695_index.html.
37. Chong Y, Carroll M, Montalvan P. NHANES converted refusals: are they different from willing respondents in socio-demographic composition? Paper presented at the annual meeting of the American Association for Public Opinion Association, Fontainebleu Resort, Miami Beach, Fl. Accessed on 4/10/2008, available from http://www.allacademic.com/meta/p16696_index.html.
38. McQuillan GM, Kruszon-Moran D, Kottiri BJ, Curtin LR, Lucas JW, Kington RS. Racial and ethnic differences in the seroprevalence of 6 infectious diseases in the United States, data from the NHANES III, 1988-1994. Am J Public Health 2004; 94: 1952-1958.
39. McQullan GM, Porter KS, Agelli M, Kington R. Consent for genetic research in a general population: the NHANES experience. Genet Med 2003; 5: 35-42.
40. Rowland ML, Forthofer RN. Adjusting for nonresponse bias in a health examination survey. Pub Health Rep 1993; 108: 380-386.
41. Groves RM. Nonresponse rates and nonresponse bias in household surveys. Publ Opinion Quart 2006; 70: 646-675.
42. Stang A, Jockel KH. Studies with low response proportions may be less biased than studies with high high response proportions. Am J Epidemiol 2004; 159: 204-210.
43. Kreiger N, Nishri ED. The effect of nonresponse on estimation of the relative risk in a case-control study. Ann Epidemiol 1997; 7: 194-199.

44. Brogger J, Bakke P, Eide GE, Gulsvik A. Contribution of follow-up of nonresponders to prevalence and risk estimates: a Norwegian respiratory health survey. Am J Epidemiol 2003; 157: 558-566.
45. Rockwood K, Stolee P, Robertson D, Shillinton ER. Reponse bias in a health status survey of elderly people. Age Ageing 1989; 18: 177-182.
46. Mazor KM, Clauser BE, Field T, Yood RA, Gurvitz JH. A demonstration of the impact of response bias on the results of patient satisfaction surveys. Health Serv Res 2002; 37: 1403-1417.
47. Groves RM, Cialdini RB, Couper MP. Understanding the decision to participate in a survey. Publ Opinion Quart 1992; 56: 475-495.
48. Cohen S, Karmack T, Mermelsteinm R. A global mesure of perceived stress. J Health Soc Behav 1983; 24: 385-396.

Vita

Carlos H. Martinez, M.D., was born in Bogota, Colombia, on April 9, 1963, to Mr. Gamaliel Martinez and Mrs. Maria R. Martinez. He finished medical school at the National University of Colombia, in Bogota, Colombia, being awarded an M.D. degree. After his graduation he did his Social Service in an underserved area of his country for six months returning to his University as a Resident in Internal Medicine for three years. After his completion of the Internal Medicine Residency Dr. Martinez served as Assistant Professor, Chief Resident in his Residency Program. He then moved to the National Hospital for Thoracic Diseases, Hospital Santa Clara, Universidad El Bosque, in Bogota, Colombia, where he received two years of training in Pulmonary Medicine. He had additional training in Critical Care, and then returned to Hospital Santa Clara, where he was appointed as Assistant Professor and Director of the Medical Intensive Care Unit. He advanced to the position of Associate Professor and was promoted to Chief of the Office of Medical Education. In 2001, he moved to the University Hospital of Fundacion Santa Fe de Bogota, in Bogota, as Director of the Medical Intensive Care Unit, and Assistant Professor of Pulmonary Medicine, at the Universidad El Bosque.

Dr. Martinez moved to the United States in 2005, to start his training all over again. He was admitted to the Combined Residency Program of Internal Medicine and Occupational Medicine, in the Departments of Internal Medicine and Preventive Medicine and Community Health, at the University of Texas Medical Branch in Galveston Texas. He has completed the three-year Internal Medicine Program and the requisites for his Candidacy for the degree of Master of Public Health, in the Graduate School of Biomedical Sciences of the University of Texas Medical Branch at Galveston.

He will start his Occupational Medicine Practicum Year in July 2008. Dr. Martinez has been admitted to a Fellowship in Pulmonary and Critical Care Medicine at the University of Michigan in Ann Arbor, Michigan, that will start in July 2009.

Dr Martinez has published more than 20 original articles in peer review publications in Colombia and the United States, and has presented more than 10 oral or original poster presentations in meeting of different societies of Pulmonary and Critical Care Medicine, Preventive Medicine, and Occupational Medicine.

Permanent address: 118 Dolphin Avenue, Galveston, TX, 77550

This capstone was typed by the author.