

5-9-2014

Attitudes and Beliefs Regarding the Collection of Race, Ethnicity and Primary Language Information in Healthcare Settings

Roberto A. Henry

University of Connecticut, robertoahenry@gmail.com

Recommended Citation

Henry, Roberto A., "Attitudes and Beliefs Regarding the Collection of Race, Ethnicity and Primary Language Information in Healthcare Settings" (2014). *Master's Theses*. 578.
https://opencommons.uconn.edu/gs_theses/578

This work is brought to you for free and open access by the University of Connecticut Graduate School at OpenCommons@UConn. It has been accepted for inclusion in Master's Theses by an authorized administrator of OpenCommons@UConn. For more information, please contact opencommons@uconn.edu.

Attitudes and Beliefs Regarding the Collection of Race, Ethnicity and Primary Language Information in Healthcare Settings

Roberto Andrés Henry

B.S., University of Connecticut, 2011

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Public Health

At the

University of Connecticut

2014

APPROVAL PAGE

Masters of Public Health Thesis

Attitudes and Beliefs Regarding the Collection of Race, Ethnicity and Primary Language Information in Healthcare Settings

Presented by

Roberto Andrés Henry, B.S.

Major Advisor _____
David I. Gregorio, Ph.D., M.S.

Associate Advisor _____
Jane A. Ungemack, Dr.P.H.

Associate Advisor _____
Brenda Shipley, M.A.

University of Connecticut

2014

Acknowledgments

This thesis would not have been possible without the support and guidance of a number of individuals.

I would like to thank my mom, my dad, my sister, and my friends for their staunch support through this process, for dealing with my mood swings, and for not hating me for bailing on them to work on my thesis. Most importantly, thank you for helping me unwind and take my mind off of work when I most needed it.

My advisors, Dr. Gregorio and Dr. Ungemack, thank you for your constant feedback, constructive criticism, and guidance. Without it, this project would not be what it has become.

Barb and Laurene, thank you for helping me with the required paperwork and making sure I met all of the deadlines.

Deb and Melanie, I am forever grateful for all of the help you gave me throughout this project. Understanding the IRB process and fine-tuning the data analysis would not have been possible without your willingness to help.

Lawrence and Rebecca, thank you for your input and your suggestions during the development of this project. Our meeting was an integral part of launching the study and fully understanding what needed to be done.

Brenda and Jeanette, I cannot thank you enough for all that you have done to make this project possible. From its inception to the printing of the last page, you have walked this often-rocky path with me, and I could not be prouder of the work that we have accomplished.

Last but not least, thank you Katie for your selflessness and unwavering support. Thank you for always being there with a smile, a pep talk, or a swift kick in the butt when I needed it, and for knowing when I needed which one. Thank you for believing in me, even when I did not.

Table of Contents

ATTITUDES AND BELIEFS REGARDING THE COLLECTION OF RACE, ETHNICITY AND PRIMARY LANGUAGE INFORMATION IN HEALTHCARE SETTINGS

APPROVAL PAGE	ii
ACKNOWLEDGEMENTS	iii
ABSTRACT	v
INTRODUCTION	1
BACKGROUND AND SIGNIFICANCE	3
RESEARCH DESIGN	10
RESULTS	25
DISCUSSION	29
LIMITATIONS	39
RECOMMENDATIONS	42
CONCLUSIONS	48
REFERENCES	51
APPENDICES	
Appendix A: <i>Survey Questionnaire</i>	55
Appendix B: <i>Attitudes & Beliefs Scales</i>	60
Appendix C: <i>Framework for Future Study</i>	61
TABLES	
Table 1: <i>Study Participant Demographics</i>	62
Table 2: <i>Adequacy of Representation by OMB Standard Classifications</i>	64
Table 3: <i>Comfort Self-Reporting REL Based on Setting</i>	65
Table 4: <i>Agreement/Disagreement with Statements About Attitudes</i>	66
Table 5: <i>Agreement/Disagreement with Statements About Beliefs</i>	67
Table 6: <i>T-Tests: Agreement with Negative Attitudes & Beliefs by Race</i>	68

Abstract

Public health researchers, practitioners and policy makers are increasingly trying to uncover, quantify and address health disparities, which are differences in health outcomes among population subgroups (Nepaul et al., 2007). Health disparities are understood to be the consequences of differences in care, the health services infrastructure, and information systems available to persons by virtue of their gender, race, ethnicity, education level, etc. (U.S. Department of Health and Human Services, 2011; Institute of Medicine, 2009). In the United States, we commonly study health disparities occurring across racial and ethnic groups, requiring the collection of race, ethnicity and language data in order to determine the existence and/or the extent to which health disparities are occurring (Carter-Pokras et al., 2002; Thorlby et al., 2011). However, this information is not always easily collected, if collected at all. In order for public health agencies to address health disparities in their communities, they must first know whether disparities exist and whom they affect. Therefore, accurate reporting of race, ethnicity and primary language (REL) data is necessary to properly identify, describe and investigate potential health disparities in the community of concern.

This study explored perceptions of race, ethnicity, and primary language data collection in healthcare settings, and more specifically the attitudes and beliefs that people have when they are asked to self-report this information. The study was conducted in collaboration with the University of Connecticut's Health Disparities Institute. The goal of this study was to develop recommendations to encourage individuals in the state of Connecticut to self-report this information, and increase their level of comfort when requested to do so in a healthcare setting.

Introduction

This study sought to explore perceptions of race, ethnicity, and primary language data collection in healthcare settings, and more specifically the attitudes and beliefs that people have when they are asked to self-report this information. The study was conducted in collaboration with the University of Connecticut's Health Disparities Institute. The goal of this study was to develop recommendations to encourage individuals in the state of Connecticut to self-report this information, and increase their level of comfort when requested to do so in a healthcare setting.

In the field of public health, researchers, practitioners and policy makers are increasingly trying to uncover, quantify and address health disparities, which are differences in health outcomes among population subgroups (Nepaul et al., 2007). Health disparities, reflected in adverse health status/outcomes, are understood to be the consequences of differences in care, the health services infrastructure, and information systems available to persons by virtue of their gender, race, ethnicity, education level, etc. (U.S. Department of Health and Human Services, 2011; Institute of Medicine, 2009). In the United States, we commonly study health disparities occurring across racial and ethnic groups, requiring the collection of race, ethnicity and language data in order to determine the existence and/or the extent to which health disparities are occurring (Carter-Pokras et al., 2002; Thorlby et al., 2011). However, this information is not always easily collected, if collected at all. It has been shown that even when forms given to individuals ask about race, ethnicity and/or language information, they are often left blank or filled inaccurately (Hirschman et al., 2000). In order for public health agencies to address health disparities in their communities, they must first know whether

disparities exist and whom they affect. Therefore, accurate reporting of race, ethnicity and primary language (REL) data is necessary to properly identify, describe and investigate potential health disparities in the community of concern.

The purpose and objective of this exploratory study was to identify and describe the attitudes, beliefs and practices among individuals with regard to self-reporting race, ethnicity and language information in healthcare settings. In so doing, these results can inform future efforts (media/educational campaigns, provider training, etc.) to encourage the self-reporting of this information in a way that respects the concerns and feelings of health care utilizers.

During the course of planning, designing and actually executing the study, we came to realize the true scope of this issue. In addition to the preliminary data that were collected and analyzed, we created a framework and guide for a future study that would adequately address all of the facets of this complex topic. This framework was informed by our conversations with community organizations and partners, as well as through a review of relevant literature.

Background and Significance

The definition of a health disparity is often debated and may vary from country to country, state to state and even organization to organization. However, all include the central belief that a health disparity is an unnecessary, unfair and avoidable difference in health between populations. Declaring a given difference in health as unfair or avoidable is open to interpretation by the agency or organization researching it, and requires some ethical judgments (Carter-Pokras et al., 2002).

Once a health issue or condition has been selected to assess health disparities, one must measure the difference in the health issue of concern among population groups (Carter-Pokras et al., 2002). Frequently, race, ethnicity and primary language (REL) are used to delineate population groups among which to measure potential health disparities in the community (Varcoe et al., 2009). For this reason, it is important for healthcare organizations to collect accurate race, ethnicity and language data at the points of enrollment and service delivery in order to identify health disparities and develop targeted programs to mitigate these disparities (New Zealand Ministry of Health, 2004).

In 2009, 57 million Americans reported speaking a language other than English at home, with 43% speaking English less than “very well” (Shin et al., 2011). This is troublesome for delivering healthcare as language barriers have been shown to not only make it more difficult for providers to establish a relationship with their patients, but it is harder for them to determine symptoms, explain treatments and most importantly educate their patients about their health (Karlner et al., 2005).

By the year 2060, racial and ethnic minorities will make up 57% of the population of the United States (U.S. Census Bureau, 2012a). This is important because research

shows that health disparities not only exist in terms of prevalence and incidence of disease, but they also exist in terms of diagnosis and treatment. For example, African Americans are 33% less likely to have a bypass surgery than are whites (Agency for Healthcare Research and Quality, 2013). With breast cancer, the time between an abnormal screening mammogram and a follow-up diagnostic test is over twice as long for minority women than it is for white women (Agency for Healthcare Research and Quality, 2013).

The Robert Wood Johnson Foundation (RWJF) and America's Health Insurance Plans (AHIP) collaborated on a study that looked into the reasons why insurance companies collect race, ethnicity and language data from their clients. The companies that were investigated reported that the top five reasons for collecting race, ethnicity and language from their clients are:

- 1) To identify enrollees at risk for certain conditions.
- 2) To support culturally and linguistically appropriate communications.
- 3) To base quality improvement efforts to reduce disparities.
- 4) To assess variation in quality measures.
- 5) To develop disease management or other specialized programs.

They also reported that knowing the primary language of their patients allows them to determine whether their materials need to be translated into a given language, or if interpretation services should be offered (AHIP, 2004). Furthermore, collecting race, ethnicity and language data to address health disparities could lead to:

- 1) Increases in compliance with follow-up treatment.
- 2) Increased access to preventive care services.

- 3) Improved patient satisfaction.
- 4) Improved patient/provider communication (AHIP, 2005).

These improvements could greatly impact the prevalence and effect of health disparities in many communities.

The cornerstone for collecting race, ethnicity and language information is quality improvement, as is evidenced by the aforementioned uses of these data. Collecting and analyzing this information helps healthcare organizations to gain a better understanding of their patients and customers as it pertains to their behavior, attitudes and beliefs regarding health and treatment (Gazmararian et al., 2012). Language information has been used by insurance plans to determine whether or not to provide their customer service representatives with basic language classes, others have used it to evaluate whether they should contract interpreters on a need-basis or if they should hire permanent interpreters for certain common languages (Gazmararian et al., 2012). Insurance plans have also used race and ethnicity data to establish and target disease management and detection programs. For example, one plan found differences in the severity of diabetes among their African American members. Based on these findings, they started a program to increase diabetes screening among their African American members, and other plans have initiated health programs based on their own findings (Gazmararian et al., 2012).

The AHIP/RWJF study also looked at the challenges insurance plans perceived they faced in the collection of race, ethnicity and language data from their patients.

Among these perceived barriers were:

- 1) The reaction of enrollees to the questions.
- 2) The enrollee's lack of understanding about how the data will be used to

benefit their health.

- 3) A lack of understanding of how it will be safeguarded.
- 4) A lack of standardized race and ethnicity categories, including options for multiple races and ethnicities (AHIP, 2005).

The Institute of Medicine (IOM) has declared that the lack of fundamental race, ethnicity and language data for patients in health care organizations is a major impediment in the movement to eliminate health disparities and improve patient care quality (Hasnain-Wynia et al., 2004).

Although standards for collecting race, ethnicity and language data have been put forth and made mandatory for many organizations, the collection of inaccurate data and patients' unwillingness to provide the data is still an issue. Without accurate demographic data, health information cannot be stratified to detect and address health disparities (Smith et al., 2010).

Among the reported reasons why people are sometimes hesitant to disclose this information are:

- 1) Fear of the information being used against them.
- 2) Being treated poorly based on their race/ethnicity.
- 3) Having a cultural history of discrimination.
- 4) Not seeing a race, ethnicity or language category that they feel represents them.
- 5) Poor phrasing of the questions.
- 6) Not knowing what the information is being used for and its benefits (Smith et al., 2010; Long et al., 2006; Jorgensen et al., 2010; Varcoe et al. 2009).

Further complicating the issue, many staff members do not believe it is important to collect the data and others feel uncomfortable having conversations about race and ethnicity with patients (Wynia et al., 2010).

In 2009 the IOM and the Office of Management and Budget (OMB) created standards for the effective collection of race, ethnicity and language data in healthcare settings (IOM, 2009). Among these standards are:

- 1) The two-question format for race and ethnicity.
- 2) Asking about ethnicity before race.
- 3) Allowing for the selection of more than one race.
- 4) Using the five OMB race categories (black or African American, white, Asian, American Indian or Alaska Native, and Native Hawaiian or Other Pacific Islander).
- 5) Only having Hispanic/Latino ethnicity (Institute of Medicine, 2009).

The OMB categories for race and ethnicity that were adopted by the IOM were developed in 1997 (Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity, 1997). Although the OMB did not make recommendations on the collection of language data, the IOM did state that identifying primary language is essential to addressing health disparities (IOM, 2009).

A 2002 study that looked into the accuracy of administrative race, ethnicity, and language data, or data that are observed or inferred by staff and entered into the patients' charts found that it was highly inaccurate. Using data from the Department of Veterans Affairs, the study compared the administrative data from dental outpatients with their records from the 1999 Large Health Survey of Veteran Enrollees, which had their self-

reported race and ethnicity. The study found that while 77.1% of the time, race data was correctly entered for white patients, non-white patients were more likely to have their race and ethnicity incorrectly listed. Only about 68% of black patients, 61% of Hispanic patients, 54% of Asian patients, and 5% of American Indian patients had been correctly identified in their dental records. These results indicate that when collecting race, ethnicity and language, self-reported information is highly accurate and should be used when possible (Boehmer et al., 2002).

A 2007 audit of Connecticut databases by the Connecticut Department of Public Health found that even with the OMB standards, there is still a lack of standardization among the data collection practices of Connecticut state agencies. Some agencies used “Hispanic” as a race rather than an ethnicity, some combined race and ethnicity into the same question, and different agencies used different race categorizations (Nepaul, 2007). Furthermore, the investigation found that 40.5% of these agencies use observed race, ethnicity and language data rather than self-reported data, which has been shown to be inaccurate (Nepaul, 2007).

In healthcare settings, race, ethnicity and primary language information is being collected to identify and address health disparities by improving the delivery and quality of care. However, for healthcare organizations to properly use this information, it must be accurate. It has been found that the most accurate way to collect this information is when it is self-reported by the patients or customers. Therefore, it is important for these organizations to address the preconceived attitudes and beliefs their patients and customers have as it pertains to the collection of race, ethnicity and language information. This paper aims to understand what these attitudes and beliefs are in order to develop

recommendations for encouraging people to self-report their race and ethnicity when it is requested in healthcare settings.

Research Design

An exploratory, cross-sectional, observational study based on a 10-15 minute web survey (Appendix A) that was distributed to participants via a crowdsourcing platform, was undertaken to examine the attitudes and beliefs that individuals have when asked to self-report race, ethnicity and language information in a healthcare setting, as well as their comfort doing so.

This study aimed to explore and assess the following questions:

- 1) What are the attitudes and feelings expressed by individuals when they are asked to voluntarily self-report their race, ethnicity and language information in a healthcare setting?
- 2) What do people believe is the purpose for asking information about their race, ethnicity and language in healthcare settings?
- 3) Are people comfortable or uncomfortable self-reporting this information, and under what conditions are they the most and least comfortable?
- 4) Do people feel that the OMB standard race and ethnicity classifications represent them?

Additionally, we were interested in using the demographic information that we collected to explore whether there were some differences in these attitudes and beliefs among various demographic groups.

In developing the study, the original focus was to measure the attitudes and beliefs of residents of Connecticut in order to develop targeted recommendations for promoting self-reporting race, ethnicity and language in Connecticut. For this reason, we included a question asking whether they were residents of Connecticut. Originally, we

wanted to recruit enough participants so that we would have sufficient Connecticut residents and we could measure if there were differences between residents of Connecticut and the national population. However, given that this is an issue of national relevance, we were confident that the results of a national sample would be applicable to residents of Connecticut in terms of developing recommendations for campaigns and messaging to promote self-reporting race, ethnicity and language.

Although race, ethnicity and primary language are variables that we believe may have an affect on an individual's attitudes and beliefs related to self-reporting race, ethnicity and language in a health care setting, we were more interested in general attitudes and beliefs regarding the collection of the information and their comfort and willingness in self-reporting it. Therefore, while we hoped to survey a sample that mirrored the racial, ethnic and linguistic composition of the state, it was not our primary objective. We were confident that we could achieve the goal of this study (determine the attitudes and beliefs of individuals regarding the collection of race, ethnicity and language information) without a sample that exactly matched the composition of Connecticut, since we believe that this issue extends beyond Connecticut, and the participant's state of residence would not impact their attitudes and beliefs.

Recruitment

The platform used to solicit responses was CrowdFlower, which is the world's leading crowdsourcing platform (CrowdFlower, 2014). Crowdsourcing allows individuals or organizations to solicit work, services or content from online contributors, usually for a small fee. Crowdsourced work varies widely. It can include data collection, sentiment analysis, content creation, and surveys (CrowdFlower, 2014). The only

inclusion criteria for the survey were that the participants must be at least 18 years of age, since that is the age of majority in most states, and be residents of the United States.

Recruitment of participants occurred through the CrowdFlower platform, which published the survey to its users and allowed anyone who was interested, and fit the criteria, to take the survey. The survey was distributed online in an effort to remove any possible confounding variables that could have affected the results had the survey been distributed in person in a place like a community health center, hospital or other care facility.

Use of a crowdsourcing platform allowed participants to be recruited in a neutral setting. We believed that distributing the survey in a location like a healthcare setting might unintentionally bring patient anxiety into the responses or make participants feel as if their responses could affect the care they would receive. Secondly, by using crowdsourcing for the distribution of the survey, we were able to gain access to a larger population in a shorter amount of time than we would have been able to given the staffing, resource and time constraints that we had. Thirdly, we believe that the online survey prevented participants from associating the questions on the survey, and their responses, with the care they might receive from their provider. Given that clinic and hospital waiting rooms can produce anxiety and stress for patients, we wanted to avoid that having an effect on the participants' responses. While using crowdsourcing for academic research is fairly new and uncommon, it has been shown to produce equally reliable data and be highly efficient (Behrend et al., 2011).

The web survey was created using Research Electronic Data Capture (REDCap), which is an application hosted by the University of Connecticut Health Center that is

used to develop secure online surveys and manage databases (Harris et al., 2009). The survey was developed and pilot-tested by the researchers. However, data analysts and others with advanced survey development experience were consulted throughout the process to ensure that the survey tool was well designed. All survey responses were collected and managed using REDCap, and once all data were collected, it was exported to SAS for analysis.

Since questions about race, ethnicity and language have the possibility of having very complex responses and reasoning, we wanted to supplement the quantitative survey data with qualitative data. However, since focus groups and structured interviews can be very resource and time intensive, in order to collect this qualitative data, we included a few open-ended questions in the web survey. In order to incentivize participation in the study, individuals who completed the survey were paid \$0.50 each. Most of the research in this area has been from the perspective of providers and insurance companies; this study sought to gain the perspective of participants on the issue of race, ethnicity and language data collection in healthcare settings.

Ideally, we wanted to stratify the data using different groupings such as race, ethnicity, primary language, education level, etc., to determine if differences existed across various subgroups, given the time and personnel constraints, we were not able to target our survey to include the necessary number of people in each group. We were confident that we could survey at least 100, if not more, adults, and analyze the results at the individual level. A total of 104 individuals were enrolled in the study. The sample size, while being large enough to achieve statistical significance, was also of a manageable size for the research staff.

Questionnaire

The demographic variables that were measured were age, race, ethnicity, primary language spoken, education level, and health insurance status. Age was broken down into ranges: 18-25, 26-35, 36-45, 46-55, 56-65, and 66+. We did not set out to find differences among age groups, but we felt that it would be useful to have these data points so that the dataset could be used in the future to see if age was a potential factor in the attitudes and beliefs regarding race, ethnicity and language information, possibly due to past histories and memories of racial discrimination among older individuals.

Race and ethnicity categories were measured using the OMB standard classifications. Although research demonstrates that many people do not believe that the OMB standard classifications of race and ethnicity properly represent them, we chose to use them in this survey for two reasons. The first being that we believed that most of the participants would be accustomed to these categories, and we would be able to measure how they typically respond to these two questions. Secondly, by using the OMB categories, and then asking them the following two visual analog scale questions: “How well do these options for ethnicity represent you?” and “How well do these options for race represent you?”, it allowed us to collect a measurable value of the extent to which individuals felt that the commonly used standards for race and ethnicity reflected their perceptions of their race and ethnicity. Education level was measured based on the respondents’ highest achieved education level. The insurance question was asked because Medicaid patients make up a large portion of the patient base at Federally Qualified Health Centers, which are mandated to collect data on the race, ethnicity and primary language of the patients they serve (Rothkopf et al., 2011; Health Resources and

Services Administration, 2012). We believed that these patients would be more accustomed to seeing these questions, and therefore may have different attitudes towards them than would the uninsured or privately insured. This demographic information was collected using multiple-choice questions on the survey instrument.

In addition to collecting the participants' demographic information, the survey collected quantitative and some qualitative data about their attitudes, beliefs and feelings towards the collection of race, ethnicity and language information in healthcare settings. Two "yes/no" questions - "Have you ever been asked for your race and/or ethnicity in a healthcare setting?" and "Have you ever been asked for your primary language in a healthcare setting?" - were used to measure whether this information is being requested as frequently as the insurance industry claims. These questions also provided a baseline for determining whether people had refused to provide this information after being asked for it. Two other "yes/no" questions - "Have you ever chosen not to answer questions about your race/ethnicity in a healthcare setting?" and "Have you ever chosen not to answer questions about primary language in a healthcare setting?" - were asked to measure whether there were people who had ever refused to self-report this information. These two questions were followed by open-ended questions asking why they had chosen to or not to self-report this information. These questions were used to gather qualitative data to better understand why people chose to provide race, ethnicity and language information, or chose not to. The survey then asked participants: "What do you believe this information about your race and ethnicity (primary language) is used for?" We included these two open-ended questions because from our review of the literature, a common theme we found was that patients and customers did not understand what the

information was used for. These questions allowed us to test the veracity of this idea, as well as see if there was an association between those who provide their race, ethnicity, and language information and what they believe the information is used for. Respondents were then asked a second series of “yes/no” questions: “Has anyone, in a healthcare setting, ever explained to you the reasons for collecting your race and ethnicity information?” and “If yes, did the explanation change your mind about providing the information about your race and ethnicity?” We believed that being told what the information was being collected for by the entity requesting said information, would have an impact on their attitudes and beliefs to self-reporting their information.

The survey then presented a series of questions and statements using visual analog scales. Visual analog scales with a range of 0-100 were used rather than typical 1-5 scales in order to get a more accurate and sensitive representation of their sentiments as it allows for a wider range of responses. The slider also made it more interesting and interactive for the participants, as there was a series of 19 of these statements. The first series of visual analog scale questions asked participants to indicate their comfort level (ranging from “Very Uncomfortable” to “Very Comfortable”) in four situations where they may be asked to self-report their race, ethnicity and language. These settings were in person, over the phone, in writing or on a form in a waiting room, and online. These questions were asked to gather data on what the best setting would be for healthcare organizations to request this information from their patients or customers.

The next set of 15 visual analog scale questions asked participants to rate their level of agreement or disagreement with a series of statements that assessed their feelings, attitudes and beliefs about the collection of race, ethnicity and language information.

More specifically, these statements addressed their beliefs about the necessity of collecting the information, what they believed the information is used for, and their reactions to being asked for this information. Presenting these statements, and asking the participants whether or not they agreed with them, allowed us to measure whether or not they believe healthcare organizations need this information, whether or not they fear it will be misused, and whether or not they feel providing their race, ethnicity and language information will be beneficial to them.

The combination of question types and the content of the questions were chosen carefully to collect comprehensive and accurate data from the subjects to properly explore the issue at hand. The quantitative and qualitative data that were analyzed for this study were collected from the subjects through the survey that was administered online. More data points were collected than were analyzed and presented in this paper. Again, the survey can be found in Appendix A. The online data collection period for the study lasted from 3:45pm to 8:00pm on March 2nd, 2014.

Human Subjects

Since this study involved surveying humans, there was the potential for some risk to the subjects. However, given the nature of the interactions, the risk was minimal if at all existent. The largest concern was causing distress to the subjects based on the subject matter of some of the questions that were asked. The survey that was administered asked questions regarding personal beliefs and feelings about race and ethnicity, and had the potential to bring up traumatic experiences related to racism and/or discrimination for some subjects. In such cases, participation could have been distressing. However, we believe that because race and ethnicity are issues that are commonly discussed, and we

did not ask about specific instances, the potential to cause distress was greatly minimized. While the information that was collected was not sensitive in nature, all efforts were made to ensure the confidentiality of the data and the anonymity of the participants. No personally identifiable information was collected, all data were stored securely and only the researchers had access to the data. All of the survey results were collected electronically and stored on the student researcher's computer in a password protected file and on a flash drive which was also password protected.

Although there were potential, yet very minimal risks to the participants, we believe that the benefits of this research study far outweighed the potential harm that may have been caused. As mentioned previously, health care organizations and health insurance companies routinely collect race, ethnicity and language information from their patients and customers in order to identify and address health disparities. Having this information allows them to create programs targeted at certain groups, determine whether to offer materials in various languages, etc. Furthermore, it has been shown that self-reported race, ethnicity, and language information is the most accurate. This research sought to explore the attitudes and beliefs of individuals regarding self-reporting race, ethnicity and language in healthcare settings, as well as their comfort actually reporting it. Ultimately, by being able to encourage individuals to provide this information when requested by their providers or insurance plans, health disparities can be identified and addressed, in turn improving the quality of care provided.

Prior to beginning any recruitment or data collection, exempt approval was sought and received from the University of Connecticut Health Center's Institutional Review Board ensuring that the study that was conducted was ethical.

Analysis

The data collected from the survey questionnaire were analyzed to answer the aforementioned research questions. Because this study is exploratory, the analysis was primarily concerned with determining whether there were any patterns in the responses, and not necessarily whether any differences or associations between predetermined groups existed. Given the small sample size (n=104), relationships between most demographic groups and their attitudes and beliefs were not analyzed for statistical significance. Only the relationships between negative attitudes and beliefs and race were analyzed for statistical significance. However, the data were analyzed using descriptive statistics to tabulate frequencies percentages of individuals who held certain positive or negative attitudes and beliefs towards the collection of race, ethnicity and language information, as well as their willingness to self-report this information and their comfort doing so in particular settings.

Frequency distributions and means reports were generated to analyze the individual variables and questions. In order to measure how well the visual analog scale statements captured the respondents' attitudes and beliefs, Cronbach's alpha coefficients were calculated for both the attitudes and the beliefs. T-tests were then performed to determine whether there were statistically significant differences between white respondents and minority respondents in terms of their attitudes and beliefs towards the collection of race, ethnicity and language information. Chi-square tests were performed to measure the association between having negative or positive feelings toward the collection of race, ethnicity and language information, and whether or not they have ever chosen not to self-report this information, and other demographic variables. However,

given the small-expected cell counts, these results were not presented.

Given the structure and format of the survey questionnaire, certain variables had to be recoded after all the data were collected in order to properly analyze them for the purpose of this study.

Question 4 (“How well do these options for ethnicity represent you?”) and Question 6 (“How well do these options for race represent you?”), which were visual analog scales, were recoded into new categorical variables with two options: “poorly” and “well.” Originally, the variables provided interval data, with a number between 0-100 indicating the extent to which the options represented the respondents’ own perceptions of their race and ethnicity. By recoding the variable into ordinal values, it allowed us to create definitive boundaries and run frequency analyses of who felt poorly or well represented by the options. All values between 0-49 were defined as “poorly” and values between 50-100 were defined as “well.” The scale was recoded in this way to avoid declaring an arbitrary neutral point. We believed that the few participants that chose the value 50 were most likely leaning towards “well” because the mean values for both variables were about 50.

Questions 21-24 (“How comfortable do you feel providing information about your race, ethnicity and primary language: in person, over the phone, in writing or on a form in a waiting room, online?”) were recoded as new variables similarly to Questions 4 and 6. These were also originally visual analog scale questions that provided a numerical value between 0-100 that quantified the extent to which respondents were comfortable, or uncomfortable, providing this information in the given settings. By recoding the variable as a categorical value rather than an interval value, it allowed us to present the

respondents' comfort levels in a more descriptive manner. For these four variables, all values between 0-49 were defined as "uncomfortable" and all values between 50-100 were defined as "comfortable". The scale was recoded in this way to avoid declaring an arbitrary neutral point. We believed that the few participants that chose the value 50 were most likely leaning towards "comfortable" because the mean values for all four of these variables were above 50.

The final 15 questions (Questions 25-39) were also visual analog scale questions with interval values from 0-100 indicating their agreement, or disagreement with a series of statements about their attitudes and beliefs about the collection of race, ethnicity and language information. These were recoded as new variables, similarly to the other recoded variables. New categorical variables with the values "disagree" and "agree" were created for all 15 of these variables. These 15 variables were recoded not only to present them in a more descriptive format, but also to be able to easily use them for chi-square tests. For variables that had a mean value above 50, all values between 0-49 were defined as "disagree" and all values between 50-100 were defined as "agree". For variables that had a mean value below 50, all values between 0-50 were defined as "disagree" and all values between 51-100 were defined as "agree". The scale was recoded in this way to avoid declaring an arbitrary neutral point. We believed that the few participants that chose the value 50 were most likely leaning in a given direction based on the mean value of for that particular question.

Given the small sample size, the race variable needed to be recoded in order to analyze the results based on race. A new dichotomous race variable was created where all who had identified as white were labeled as "white" and all others were labeled as

“minority.”

Cronbach’s alpha coefficients were calculated for a scale of attitude statements and a scale of belief statements in order to measure how well they actually measured the participant’s attitudes and beliefs. Since some of the statements in the survey had been worded with a positive tone and others with a negative tone, before calculating the alpha coefficient, the variables had to be recoded into new variables to make sure the responses were relatable. For consistency, all statements were recoded to be “negative” statements, therefore agreement with a statement indicated that the respondent held a given negative attitude or belief towards the collection of race, ethnicity and language information in healthcare settings. The following statements were recoded for this purpose:

- “I trust that my race and ethnicity information will not be misused.”
- “My race and ethnicity are relevant to my medical treatment.”
- “I think information about my race and ethnicity will be helpful for doctors.”

Given that these variables used a numeric scale from 0-100, they were recoded to depict a negative statement by subtracting each respondent’s value from 100. The new values indicated the participants’ agreement or disagreement with negative versions of the original statements.

After this, all of the attitude and belief statement variables were recoded into new variables for the Cronbach’s alpha calculation. The original variables were on a scale of 0-100, however, in an effort to prevent having a scale that was too large and cumbersome, the responses for each statement were divided by a factor of 10 to reduce the scale to 0-10. A breakdown of the attitude and belief statements that were used for

the Cronbach's alpha coefficients can be found in Appendix B.

The five statements that were categorized as "attitudes" were grouped together under this heading because they deal with how the respondents feel when providing information about their race, ethnicity and language in healthcare settings. The six statements that were categorized as "beliefs" were labeled as such because they address what participants believe is the purpose of collecting this information in healthcare settings. The two belief statements that directly referred to race, ethnicity and language data collection for insurance purposes ("Insurance premiums are affected by the individual's race and ethnicity" and "My insurance company should know my race and ethnicity") were not included in the "beliefs" scale. An inter-item analysis was performed for the beliefs statements and the intercorrelation matrix suggested that these two items fell outside of the scale as they had a low correlation with the other items ($r=0.09$, $r=0.15$).

After the Cronbach's alpha coefficients were calculated, summary scores were created for the "attitudes" scale and the "beliefs" scale. These summary scores were calculated by taking the mean of the scores for the five attitude statements that were used for the "attitudes" Cronbach's alpha coefficient, and then multiplying those means by the total amount of attitude statements, 5. The same was done to create a summary score for the belief statements. However in this case, the mean values for the responses to the belief statements were multiplied by 6, instead of 5. The mean was used rather than simply adding all of the values because some participants did not respond to certain statements, therefore by using the mean it allowed for a score to be created without penalizing those who did not respond to a statement.

The summary scores that were created were used to perform an independent t-test. This t-test was used to determine if there was a statistically significant difference between the level of agreement, or disagreement, with the negative attitude statements based on race (white vs. minority). A similar independent t-test was also performed to determine whether a statistically significant difference existed between the level of agreement, or disagreement, with the negative belief statements based on race (white vs. minority).

Analysis of all quantitative data was performed using SAS.

Results

One hundred and four participants responded to the survey and their data were analyzed. The respondents were demographically diverse. The respondents were 71.15% white, 15.38% Asian, 10.58% black or African American, 2.88% American Indian or Native American, 1.92% some other race, and 0.96% Native Hawaiian or Pacific Islander. Fewer than one out of four participants (18.27%) were residents of Connecticut. Only 11.54% reported being Hispanic. The participants were also varied in regards to their age, level of education, and insurance status. The full demographic breakdown of the study participants can be found in Table 1.

One of the purposes of the survey was to measure the extent to which individuals felt that the OMB's standard classifications of race and ethnicity, which have been adopted at the federal level and by the IOM, truly represented their perceptions of their race and ethnicity. Respondents were asked to select their ethnicity: "Hispanic or Latino", "Not Hispanic or Latino", "Rather not Respond", and their race: "American Indian or Native American", "Asian", "black or African American", "Native Hawaiian or Pacific Islander", "white", "Other", "Rather not Respond." Following each of these questions, they were asked to select the extent to which these options for ethnicity and race represented them.

As seen in Table 2, 91.09% of the 104 respondents felt that the OMB categories for race represented them well, but only 66.67% felt that the OMB standard options for ethnicity represented them well. The average value on the scale for ethnicity was 55.85 ($s=31.73$). While the mean value for the race question was 80.07 ($s=23.02$), indicating that on average individuals felt that the standard options for race represented them better

than the did the ethnicity options.

When asked if anyone in a healthcare setting had ever explained the reasons for collecting their race and ethnicity information, 93.27% of respondents indicated that, in fact, they had never been told why said information was being collected. Six of seven individuals who responded that they had received an explanation for the collection of their race and ethnicity information stated that the explanation had changed their mind about providing the information. Moreover, only 12.75% of participants said that they had ever chosen not to answer questions about their race or ethnicity in a healthcare setting. Correspondingly, very few (6.80%) respondents claimed that they had ever decided not to answer questions about their primary language in a healthcare setting.

Participants were asked about their levels of comfort or discomfort self-reporting their race, ethnicity and primary language in a series of different settings. While some respondents indicated some discomfort providing this information in these settings, the majority stated that they were comfortable regardless of the setting (Table 3). The proportions of respondents who were comfortable or uncomfortable about providing their race, ethnicity and language information ranged from 5.88% in person to 13.73% over the phone.

On average, the respondents did not agree with the negative attitude statements that were presented in the survey. The average values, which can be found in Table 4, on the visual analog scale attributed to the statements expressing worry that their race or ethnicity will be used against them ($\bar{x}=38.93$, $s=26.22$), fear of providing their race or ethnicity to their doctor ($\bar{x}=27.47$, $s=24.87$), annoyance when faced with questions about their race and ethnicity ($\bar{x}=41.92$, $s=27.98$), and discomfort when answering questions

about their race or ethnicity ($\bar{x}=31.82$, $s=25.37$), were within the “disagree” range.

Similarly, on average, participants were more likely to trust that their race and ethnicity information will not be misused ($\bar{x}=68.67$, $s=22.20$).

The average quantifications of agreement and disagreement with the belief statements regarding the collection of race, ethnicity and language information (Table 5) appeared relatively more varied than the attitude statements. Although only slightly, participants were more likely to agree that their insurance company should know their race and ethnicity ($\bar{x}=58.64$, $s=25.26$). Likewise, participants only marginally agreed that their race and ethnicity were relevant to their medical treatment ($\bar{x}=58.55$, $s=25.78$). When presented with the statement: “Insurance premiums are affected by the individual’s race and ethnicity,” the average rating of agreement was essentially undecided ($\bar{x}=49.68$, $s=22.57$). On average, participants believed that their primary language has no connection to the quality of their health ($\bar{x}=66.15$, $s=28.42$). The average responses for the statements “Race and ethnicity affect the quality and type of health care given” and “I am unsure about the reasons why the doctors need to know my race and ethnicity” were also undecided ($\bar{x}=48.05$, $s=24.29$; $\bar{x}=50.84$, $s=27.37$).

The Cronbach’s alpha coefficient of internal consistency for the five-statement attitudes inventory was .86. The Cronbach’s alpha coefficient for the six-statement beliefs inventory was .75.

As can be seen in Table 6, based on the results of the independent t-tests, minority participants were significantly more likely to hold negative attitudes regarding the collection of race, ethnicity and primary language information in healthcare settings than white participants ($\bar{x}=22.97$, $s=9.51$ versus $\bar{x}=14.67$, $s=9.46$), $t(54.37)=4.01$, $p=0.0002$.

On the other hand, minority participants ($\bar{x}=28.27, s=8.87$) did not report significantly different levels of agreement with negative beliefs regarding the collection of race, ethnicity and primary language information in healthcare settings than white participants ($\bar{x}=26.52, s=10.89$), $t(66.49)=0.85, p=0.3997$.

Discussion

Given the exploratory nature of this study, formal hypotheses were not formulated prior to initiation. The data gathered from this study provide us with some insight regarding how individuals feel about the collection of their race, ethnicity and language information, and self-reporting it in healthcare settings. This study also provided information about how the public actually views the widely adopted race and ethnicity standards that were created by the OMB.

The current standards for reporting and collecting race and ethnicity data have been in effect since 1997. However, there has been some concern from organizations that these standards don't really work for the changing American population. Additionally, insurance agencies and other healthcare organizations have stated that while they are collecting race and ethnicity data from their customers and patients, the information is not always accurate. For this reason, we sought to measure the extent to which individuals felt that the OMB standards for race and ethnicity actually described their perception of their race and ethnicity. Given the doubts of the insurance companies and organizations regarding the accuracy of the data being collected, and questions about the relevance of the almost 17-year-old classifications, we believed that a large proportion of our participants would state that they did not feel that the OMB standards represented them. However, based on our results, this was not the case. The majority (91.09%) of respondents reported that the OMB classifications for race represented them well. While these results are not necessarily valid for the general population, it indicates that the current racial classifications might be more accurate than one would believe. The responses to the question asking how well the OMB standard ethnicity options

represented the participants were much more varied. Two-thirds of respondents stated that the ethnicity options represented them well. These data suggest that the ethnicity options may be less acceptable. Having only two options for ethnicity creates a dilemma for people who are not Hispanic and consider their ethnicity to be more than just “Not Hispanic or Latino”.

Based on our review of the literature, as well as our own experiences, we suspected that most people had never been given an explanation as to why their race and ethnicity information was being collected. Our suspicions were confirmed, as 93.27% of participants reported that their insurance company or healthcare provider had never explained why they were collecting information about their race and ethnicity. Given that one of the main reasons why people are hesitant to self-report their race and ethnicity is not knowing what the information is being used for, it is important to make note of this statistic. Granted, individuals could know what the information is used for without having been told by their provider, or may not care to ask. However, providers and insurance companies cannot assume that their customers and patients will know why they are being asked for this information. With so few participants having ever received an explanation of the reasons for collecting this information, not receiving an explanation could reduce the chance of people self-reporting their race and ethnicity. Data from this study showed that when people receive an explanation for why these data are collected, the majority of people change their mind about providing race and ethnicity information. Although they were not asked in what direction that change occurred, whether it led them to provide it or to withhold it, it is a clear indication of the importance of that explanation.

As mentioned earlier, the ultimate purpose of this study was to determine effective ways to craft messaging and programs to promote the importance of self-reporting race, ethnicity and language information in healthcare settings. Therefore, we asked participants if they had ever chosen not to answer questions about their race, ethnicity and language. However, these two questions were not ideal, as they only asked if they had ever in their lifetime refused to provide this information. These questions would in essence group an individual who refused once with an individual who has refused every time it was asked. Very few (12.75%) respondents reported ever refusing to answer questions about their race and ethnicity in healthcare settings. This would suggest that non-response in terms of race and ethnicity information is not a common issue. The participants who reported ever choosing not to answer questions about their race and ethnicity followed two common themes in terms of their reasoning. Those who had chosen not to answer these questions in healthcare settings believed that the information was not relevant or important in receiving treatment. A few of the respondents also stated that they did not provide this information because they did not want to be judged based on their race.

Similarly, only 6.8% of respondents reported ever having chosen not to answer questions about their primary language. This is almost half the rate of refusal for race and ethnicity questions. This difference would suggest that for patients it is less troublesome to report their primary language, or that patients have a better understanding of the purpose for collecting their primary language. When asked why they had ever chosen not to respond to language questions, the responses followed two themes. The first being that the question had never been asked because they spoke English, and the

second being that they did not believe it was important information.

Insurance companies and healthcare providers can request race, ethnicity and language information in a multitude of settings and locations. In order to gauge whether certain settings versus others made individuals more or less comfortable with reporting their race, ethnicity and primary language, participants were asked to rate their level of comfort (or discomfort) in providing this information. There was no setting in which participants felt uncomfortable providing their race, ethnicity, and language to a healthcare organization. Although reporting this information over the phone received the lowest mean score ($\bar{x}=74.33$, $s=23.08$), it was still well within the comfortable range. These results indicate that the participants feel relatively comfortable providing information about their race, ethnicity and language in a variety of settings.

Somewhat surprisingly, most participants tended to disagree with the negative attitude statements that they were presented. Very few respondents expressed feeling fear when they were asked by a provider to provide their race and ethnicity. Similarly, the majority of individuals did not feel uncomfortable when asked questions about their race and ethnicity. Following this trend, only 37.62% of participants worried that their race and ethnicity information would be used against them. Interestingly, most of the participants expressed that they trusted that their race and ethnicity information would not be misused. The fact that only 14% of respondents did not trust that their race and ethnicity information would be used properly, yet 37.62%% voiced some worry that their race or ethnicity information will be used against them points to three possibilities. The first possibility is that some individuals have an inherent trust their information will be used properly, but still worry that it might not. The second possibility is that “misuse”

and “used against me” were interpreted as being different enough situations that a person could worry about one situation occurring yet trust that the other would not happen. The third possibility is that the wording of the statements was confusing to survey takers and caused this discrepancy in the results. Slightly less than half of the respondents, 40%, found questions about their race and ethnicity to be annoying. With a fairly large proportion of individuals finding questions about race and ethnicity to be annoying, it could lead to nonresponse, or inaccurate responses. Additionally, finding these questions annoying could be an indication of not knowing the purpose of healthcare organizations collecting this information. If this were the case, by educating healthcare consumers about the reasons for collecting this data, it might reduce the annoyance and produce more accurate and consistent self-reporting of race, ethnicity and language information. In general, individuals held positive attitudes towards the collection of race and ethnicity data in healthcare settings. Participants did not express fear, worry, discomfort or annoyance when this information was requested, and more importantly, they trusted that it would be used properly.

Unlike the response towards the attitude statements, the statements dealing with the participants’ beliefs regarding the use and purpose for collecting race, ethnicity and language data elicited more varied responses. Interestingly, only 59% of participants believed that race and ethnicity has some connection to health. One would assume that the rate of disagreement would be much higher considering the existence of genetic illnesses and frequent discussion of conditions affecting certain populations more frequently. In contrast, 73.27% of individuals believed that there was no connection between a person’s primary language and the quality of their health. This would indicate

that the majority of people do not see it as important to provide their primary language, even though one of the primary reasons it is collected is to ensure that the healthcare organizations can provide care and information in a linguistically understandable method. However, the fact that 72% of respondents felt this way is less surprising when paired with the fact that 98% of the participants spoke English as their primary language.

When participants were asked to rate their agreement with the statement: “My insurance company should know my race and ethnicity,” 74.75% agreed with it. Conversely, 60% of respondents believed that their race and ethnicity affect their insurance premiums. This appears to be counterintuitive, unless those that believed that insurance premiums were affected by race and ethnicity thought that their own race and ethnicity would be beneficial in terms of having a lower premium. Otherwise, it would not make sense for an individual to believe that their insurance company should know their race and ethnicity even though it might negatively impact their premium rates. Given that these two variables are not correlated ($r=-0.02$, $p=0.85$), this issue is a great candidate for further qualitative research. It would be beneficial to analyze this further in an interview or focus group, to see why people believe that insurance companies should know their race and ethnicity, even though they believe it will affect their insurance premiums.

As became fairly apparent, the majority of participants felt that providing information about their race, ethnicity and language in healthcare settings is beneficial to the care and treatment they receive. An overwhelming majority of participants (91.09%) expressed a belief that having race and ethnicity data on their patients was helpful to doctors. Similarly, a majority of participants (71.29%) agreed that their race and

ethnicity were relevant to their medical treatment. This shows that people do understand, or at least expect that there is a valid purpose for collecting this information, and that it will be used in some way to improve the quality of their healthcare. However, one should keep in mind that 93.27% of respondents reported never having received an explanation, in a healthcare setting, for why this information was being collected. Although it seems as if a majority of individuals believe there are beneficial reasons for collecting this data, it is important that providers, insurance companies, and other healthcare organizations reinforce these beliefs by explaining in some way what this information is being used for.

The importance of explaining why healthcare organizations collect information about race, ethnicity and language is indicated by the responses to the statement: “I am unsure about the reasons why the doctors need to know my race or ethnicity.” Even though 91.09% believe it is helpful to doctors and 71.29% believe that it is relevant to their treatment, 58% of participants were unsure why doctors needed to know it. While it is clear that most individuals think that there is some benefit to self-reporting this information to their healthcare providers, it is evident that most do not know exactly why it is beneficial. If providers, insurance companies, and other healthcare organizations provided explanations when they requested this information, it would not only give patients a better understanding of what the data are used for, but it might encourage them to be more truthful and forthcoming with their responses and to contribute said information more consistently.

When presented with the statement: “race and ethnicity affect the quality and type of health care given,” the response was very mixed. The statement itself was fairly

ambiguous, as it did not specify whether the effect on the quality and type of care given was positive or negative. About half (50.51%) of respondents agreed with the statement. Simply knowing that about half of the participants believe that their race and ethnicity affects how they will be treated medically raises more questions. These results do not give an indication of whether these individuals believe that if a doctor knows their race or ethnicity, they will receive better care, or if they believe that doctors would use this information to provide worse care based on racial or ethnic prejudices. If follow-up interviews or focus groups were to be conducted, this is an issue that should be explored further. This should be done to determine what the perceived effects on quality and type of health care are, and if differences exist between the individuals who believe that knowing their race or ethnicity would help doctors provide better care and those one believe it would cause them to receive worse care.

The Cronbach's alpha coefficients of internal consistency for the attitude and belief statements confirmed that the variables within each category were correlated enough that the designated groupings were appropriate to create separate attitude and belief scales.

Based on the results of the t-test that was conducted on the differences in negative attitudes towards the collection of race, ethnicity and language information in healthcare settings based on race, a significant difference does exist between minorities and whites. Minority respondents were more likely than white respondents to have negative attitudes (fear, worry, annoyance, distrust, discomfort) when asked to report their race, ethnicity and language information. This supports the idea that populations with histories of racial discrimination have an apprehension to providing race, ethnicity and language

information. However, there was no difference in terms of their beliefs about the utility and purpose of this information in healthcare settings. White and minority participants were equally likely to have negative beliefs about the reasons why healthcare organizations collect this information. This could be because while minorities appear to have more negative attitudes about the collection of race, ethnicity and language information, they have an understanding of the purpose that it serves in the healthcare field.

The results of the survey provided important preliminary data. We were able to begin to piece together a better understanding about the collection of race, ethnicity and language data. This analysis raised questions that were not originally thought of in the development of the study. The major finding from the survey results was that a majority of individuals do believe that there is a beneficial purpose for healthcare organizations collecting race, ethnicity, and primary language data. However, it was apparent that most individuals were unsure of what that reason was, which is most likely due to the fact that an overwhelming majority had never been told what the information was being collected and used for. Contrary to expectations, most participants were satisfied with the OMB standard classifications of race and ethnicity; although fewer felt that the ethnicity classifications represented them well than felt that the race classifications did. While this is not necessarily representative of the United States population as a whole, it did contradict the initial expectation that people did not feel that the OMB standards were relevant to their perceptions of their race and ethnicity. However, this could be due to the fact that people have become accustomed to these standards, as they have been adopted extensively. This initial analysis provided a platform from which to develop more

thorough and targeted studies to further analyze and address the issue of collecting and self-reporting race, ethnicity and primary language in healthcare settings.

Limitations

Everyone who undertakes a research project sets out to develop a study that addresses all potential shortcomings, and is without limitations. Unfortunately, we do not live in an ideal world, and even the most detailed and well-planned study has constraints. Although efforts were made to minimize study limitations, as with any research venture, limitations in this study were inevitable. The limitations that presented themselves through the development of this study, and impacted the generalizability and quality of this study stemmed from resource, time and personnel constraints.

The main limitation in this study was the generalizability of our results. We chose to use the crowdsourcing platform, CrowdFlower, to recruit participants in order to access a large number of people in a short amount of time, and to distribute the survey online to unknown individuals in order to remove the possibility of our presence or the setting affecting the responses. However, while our recruitment method did serve this purpose, it likely hurt the generalizability of our results. People who participate in online crowdsourcing are likely to be more computer literate and educated than the general population. As evidenced by our results, the distribution of education levels was fairly varied. However, the crowdsourced sample was more educated than the population of the United States. In our sample, 44.23% had a Bachelor's degree or higher, while only 29.1% of the American population has a Bachelor's degree or higher (U.S. Census Bureau, 2012b). Although we do not know whether education, computer literacy, or access to the Internet has an effect on people's attitudes and beliefs regarding the collection of race, ethnicity, and language information, it negatively impacts the external validity. Our sample was also significantly younger than the population of the United

States. Of our participants, 73.08% were 45 years old or younger. Conversely, only 36.4% of the population of the United States is under the age of 45 (U.S. Census Bureau, 2012b).

Moreover, because we could not specify that participants be only from the state of Connecticut, and we did not recruit enough participants to generate a large enough sample of Connecticut residents among our subjects, we cannot claim that our results can be generalized to Connecticut residents specifically. Nevertheless, we are confident that the patterns and themes we uncovered are applicable to Connecticut and can be used to develop recommendations for action in Connecticut. This study was exploratory in nature, therefore generalizability was not as much of a concern as was determining what patterns existed.

Furthermore, because the sample size was small ($n=104$), it was not possible to conduct analysis to determine whether statistically significant differences and associations existed between most demographic groups with respect to self-reporting race, ethnicity and language. Although this was not necessary for the exploratory purposes of the study, it would have greatly strengthened our results had it not been for the resource, time, and personnel constraints, as well as the lack of prior research studies on this topic that would have allowed for us to design an explanatory study rather than an exploratory one.

Moreover, this study relied upon self-reported data about attitudes and beliefs, which cannot be verified, and inherently present limitations. Given the nature of the questions and the complete anonymity of the survey, one would hope for honest and accurate responses, but the potential for bias still exists. In this case, recall bias and

exaggeration were the most likely biases to be introduced into the results, although the likelihood is small. With self-reported data, one must take the responses at face value and expect that with a large enough sample size, flawed data points will not have much of an effect on the final results. However, because of the small sample size used for this study, there is a greater possibility that inaccurate responses affected our results. However, we attempted to mitigate this by asking some questions multiple times in different ways and creating scales.

Finally, although a lot of time and effort was placed into developing and fine-tuning the survey instrument, it is not a perfect questionnaire and did introduce some limitations into the results. After analyzing the data, new questions and issues emerged that were not addressed in the survey tool. For example, some of the belief statements were ambiguous in the sense that they were not explicitly positive or negative, therefore in analyzing the data, one could not conclusively determine what they were agreeing or disagreeing with. In the future, some modifications and additions should be made to the survey in order to address this deficiency.

The issue of race, ethnicity, and language data collection is a very rich, and complex topic that requires more than quantitative measurement. Our survey tool did not allow us to go further into the issue and understand why individuals have certain attitudes or beliefs towards the issue, or if they had other feelings towards it that we had not considered. For this reason, using structured interviews and/or focus groups to further study this topic and collect more complete and expansive qualitative data would greatly improve this study.

Recommendations

A Framework for a Future Study

When originally conceived, we believed that the survey along with follow-up structured interviews would allow us to adequately study the issue of race, ethnicity and primary language information in healthcare settings. However, due to time, resource and personnel constraints, personal interviews were removed from the research design. Still, we believed that a well-crafted survey that included open-ended questions would allow us to collect quality data as it pertains to the attitudes and beliefs associated with the collection of race, ethnicity and primary language information. Nevertheless, as the planning and design of the project progressed, we came to realize that a more comprehensive and multi-faceted approach was necessary to collect high quality data on this complex issue. Through our planning sessions, research, and meetings with community stakeholders, we began to form a plan for the ideal study to address our question. The following are my recommendations for conducting a future study given the appropriate time, resources and personnel.

The best way to study this issue would be a multi-faceted approach. Using a mixture of online surveys, paper-based surveys, structured interviews and focus groups would provide the best and most comprehensive data. The issue of race, ethnicity and language data collection potentially affects people very differently. The reactions and feelings that people have to being asked these questions can be hard to capture through just a survey with predetermined choices, which is why adding focus groups and interviews would be beneficial. A visual representation of this framework can be found in Appendix C.

Focus Groups

The importance of qualitative data when it comes to researching this issue has been stressed throughout this paper. Focus groups would be an excellent way to gather data from community members in a less structured, but equally useful, manner. Although there should be predetermined questions used for the focus groups, they would allow for participants to bring up ideas, concerns, and/or feelings surrounding the issue that had not occurred to the researchers. In a group setting, participants might have thoughts spurred from another participant's comments that would not necessarily have come up from a direct question in a survey or interview.

For this reason, focus groups would be very beneficial, as they could generate more, in-depth qualitative data to inform the development of the survey and interview tools. However, they are time and resource intensive. Depending on the number and size of the groups, the focus groups could take up a lot of time, and transcribing and analyzing the comments would also take a lot of time. Finding a space to host the focus groups is another consideration, as it must be safe and accessible for the participants, as well as large enough to hold the groups. Effort must also be put into recruiting participants for the focus groups. However, researchers could achieve this by posting fliers in different community buildings, or by contacting community-based groups and organizations that already meet regularly and asking to take over one of their meetings. Involvement in focus groups would most likely need to be incentivized. Furthermore, when possible, personally identifiable information should not be collected during the focus groups in order to provide anonymity to the participants.

Paper-Based Survey

We recommend taking three approaches for recruiting participants for the paper-based survey. Recruitment should occur at the point of encounter (e.g., at a clinic, community health center, hospital, etc.), at the point of enrollment (e.g., at a patient navigator or access to care office), and general recruitment outside of a healthcare setting (e.g., in community based organizations, religious organizations, library, etc.). These three settings should be used for recruitment because the study is seeking to determine the attitudes and beliefs held by people in regard to the collection of race, ethnicity and language information in healthcare settings. Therefore, distributing a survey in the settings where they are being asked to self-report this information, and in a control environment, would help to determine whether the environment plays a role in their attitudes.

The inclusion of a paper-based survey is beneficial because it would allow the researchers to measure whether being in these settings that may cause some anxiety (point of encounter and point of enrollment), where their race, ethnicity and language information is being collected, has an effect on their attitudes and beliefs towards self-reporting. Furthermore, by using a paper survey, the researchers would have greater access to underserved populations and individuals who may not have access to the Internet.

The paper-based survey should be shorter than the web survey, and should not include as many, if any, open-ended questions. This is because the survey will most likely be distributed in waiting rooms, while participants are waiting for an appointment so not only will they not have the time to complete a long, involved survey, but they will

most likely be preoccupied or distracted and their responses will be less reliable.

Although Likert-style questions could not replace the type of data that open-ended questions would provide, they would allow for the collection of more sensitive pieces of data.

Web Survey

A survey similar to the one used in this study should be distributed online in addition to the paper-based survey. Using a web survey would serve multiple purposes for this study. First of all, distributing the survey online would allow the use of a longer survey, with more open-ended questions. If participants are taking the survey online, they will most likely have more time and be more attentive than they would be if they were completing it in a waiting room. The web survey could also be distributed to a much wider audience than the paper-based survey.

Distributing a paper-based survey is reliant on what populations you have access to, and being granted permission to recruit survey participants could be difficult. The online survey, especially if distributed through a crowdsourcing platform, allows researchers to gain access to a much larger recruitment base, and collect responses in a relatively short amount of time. Furthermore, the online distribution of the survey eliminates the potential feeling of anxiety that individuals might feel if they are in the waiting room of a health center. Along these lines, asking questions in a healthcare setting about feelings towards providing certain information in a healthcare setting could be conflicting for potential participants. This is because they may be suspicious of the researchers' intentions, and be less truthful in their responses, believing that they may affect the care they are about to receive. Employing an online survey such as the one

used for this study will allow for gathering more data in a very efficient fashion, while at the same time removing the potential of anxiety or suspicion affecting the responses. However, this would likely introduce other biases, which should be accounted for accordingly.

If a specific geographic population is of interest, a mail survey to a random sample of people in that target area could be used in lieu of a web survey.

Structured Interviews

As has been mentioned, the survey questionnaire is an effective tool in gathering preliminary data. However, with a survey, it is difficult to gather in-depth qualitative data, which is necessary to truly investigate this issue. Through the development of this study and during the data analysis, there were many follow-up questions that presented themselves but could not be properly addressed with survey questions. In dealing with attitudes, beliefs and feelings towards an issue, it can be hard to effectively categorize them into answer choices for survey questions. For this reason, follow-up structured interviews would be greatly beneficial in delving deeper and getting behind the survey responses. The data collected from the interviews could then be analyzed and themes could be better categorized. In this case, the survey responses would not be anonymous, and respondents who are interested in participating in the follow-up interviews would have to provide some contact information on their survey.

For example, while our survey tool captured the proportion of respondents that believed the OMB race and ethnicity classifications represented them well, this information would be much more useful if the reasons why they felt this way were known and could be compared. If the ultimate goal of a study of this type were to

promote self-reporting one's race, ethnicity and language in healthcare settings, collecting this in-depth qualitative data would be ideal in developing appropriate and effective messaging.

Structured interviews would require more time and financial resources to conduct than the surveys would. Depending on how talkative the subjects are the interviews could take a long time. Also, the subjects might need to be incentivized to participate in the follow-up interviews, which would add a financial cost. Finally, in order to make the most of the interviews, the person, or people, conducting them should have some training or experience in interviewing. Otherwise, they could make the subjects uncomfortable, or do an ineffective job of drawing information from them, which would defeat the purpose of including interviews as a part of the study.

A study that combined all four of these approaches would be the ideal method to comprehensively investigate the issue of race, ethnicity, and primary language information collection in healthcare settings. The combination of quantitative and qualitative data would help generate stronger conclusions, and the results could be used to develop effective campaigns and messaging to promote self-reporting of race, ethnicity and language information in healthcare settings.

Conclusion

It was clear that most individuals in this study appreciate the importance and relevance of race, ethnicity, and language information as it relates to medical treatment and care. However, it was also evident that although many believed it was beneficial, they were unsure about why the information was needed. This speaks to the need for healthcare organizations to communicate to their patients and customers the reasons for collecting these data and what they are used for. By doing so, they will not only put people at ease who might be apprehensive about reporting this information, but they will increase the likelihood that they will get accurate and complete responses

This study, and future studies that may stem from it, have many implications for both public health practice and policy. Collecting accurate race, ethnicity and primary language information is critical to identifying and addressing health disparities within communities. This study aimed to explore the attitudes and beliefs that individuals hold regarding the collection of race, ethnicity and language information in healthcare settings. The data collected from this study can be used to begin to develop recommendations for campaigns or other efforts to promote and encourage self-reporting of race, ethnicity and language, as well as address fears or apprehensions some may have with providing this information. Based on the findings, any messaging that is developed must address the fact that healthcare consumers are not being told the reasons for collecting this information. Just as importantly, promotional campaigns must keep in mind that minorities hold more negative attitudes towards the collection of race, ethnicity and language information than whites. Messaging should be targeted to these communities to address these negative attitudes and assure them that they have no need to worry, or be

afraid of providing said information.

If targeted campaigns can be created and implemented effectively, self-reporting of race, ethnicity and language will increase in healthcare settings. This will allow healthcare providers, health insurance companies, and other healthcare organizations to better understand their patients and customers. With this information, healthcare organizations can begin to develop services and programs to improve the quality of the care they provide to their patients and begin to address health disparities.

The study also investigated the relevance of the current Office of Management and Budget race and ethnicity classification standards. A majority of respondents believed that the race and ethnicity standards represented them well. While an overwhelming majority (91.09%) felt well represented by the race classifications, a much smaller proportion (66.67%) felt represented by the ethnicity standards. This should be further explored because the current ethnicity classifications may need to be reworked in order to capture data more accurately.

An innovation of this study was the use of crowdsourcing to conduct the research. As mentioned previously, although crowdsourcing is suited for academic research, it is a relatively new and uncommon practice. Therefore, there was some uncertainty regarding our sample, and whether we would be able to recruit a diverse enough sample to validate our results and make them more generalizable. While, there was wide variability across all of the demographic variables that were collected, the sample was younger and more educated than the general population of the United States (see Table 1). Despite the fact that this study was not meant to be a review of the efficacy of using crowdsourcing for academic research, the results that were generated do speak to its practicality and value.

Based on this experience, given the speed, minimal expense and quality of the data collected, one can be confident in saying that the use of crowdsourcing for academic research is a viable option, especially for resource and time-constrained studies that are hoping to collect large amounts of data.

This exploratory study gathered meaningful preliminary data regarding the attitudes and beliefs of individuals as it pertains to the collection of race, ethnicity and language information in healthcare settings. The results that were collected have established a starting point for future, more in-depth research on specific aspects of this very broad issue. These results, and future analysis of the data, will be used to begin conversations with local community based organizations on how they can work together with healthcare organizations to promote self-reporting of race, ethnicity and primary language information to improve the health of their community members.

References

1. Agency for Healthcare Research and Quality. (2013). *Addressing Racial and Ethnic Disparities in Health Care*. Rockville, MD
2. America's Health Insurance Plans. (2004). *Health Insurance Plans Address Disparities in Care: Highlights of a 2004 AHIP/RWJF Quantitative Survey, Collection and Use of Data on Race and Ethnicity*.
3. America's Health Insurance Plans. (2005). *Tools to Address Disparities in Health: Data as Building Blocks for Change*. Rep. Washington D.C.: America's Health Insurance Plans.
4. Behrend, Tara, David Sharek, Adam Meade, and Eric Wiebe. (2011). The Viability of Crowdsourcing for Survey Research. *Behavior Research Methods* 43(3): 800-813.
5. Boehmer, Ulrike, Nancy R. Kressin, Dan R. Berlowitz, Cindy L. Christiansen, Lewis E. Kazis, and Judith A. Jones. (2002). Self-Reported vs Administrative Race/Ethnicity Data and Study Results. *American Journal of Public Health* 92 (9): 1471-2.
6. Carter-Pokras, Olivia, and Claudia Baquet. (2002). What Is a 'Health Disparity'? *Public Health Reports* 117.5: 426-34. Web. 20 Sept. 2013.
7. CrowdFlower. (2014). <<https://crowdfunder.com/index>>. 27 Feb. 2014.
8. Gazmararian, Julie, Rita Carreón, Nicole Olson, and Barbara Lardy. (2012). Exploring Health Plan Perspectives in Collecting and Using Data on Race, Ethnicity, and Language. *American Journal of Managed Care*, 18(7), 254-261.
9. Harris, Paul A., Robert Taylor, Robert Thielke, Jonathon Payne, Nathaniel Gonzalez, & Jose G. Conde. (2009). *Research Electronic Data Capture (REDCap) - A*

Metadata-Driven Methodology and Workflow Process for Providing Translational Research Informatics Support (4.8.19 ed.) Vanderbilt University.

10. Hasnain-Wynia, Romana, Debra Pierce, and Mary A. Pittman. (2004). *Who, When, and How: The Current State of Race, Ethnicity, and Primary Language Data Collection in Hospitals*. New York: The Commonwealth Fund.
11. Health Resources and Services Administration. (2012). Health Center Data. <<http://bphc.hrsa.gov/healthcenterdatastatistics/index.html>>. 14 March 2014.
12. Hirschman, Charles, Richard Alba, and Reynolds Farley. (2000). The Meaning and Measurement of Race in the U.S. Census: Glimpses into the Future. *Demography* 37.3: 381-93. Web. 20 Sept. 2013.
13. Institute of Medicine. (2009). *Race, Ethnicity and Language Data: Standardization for Health Care Quality Improvement*. Rockville, MD: U.S. Department of Health and Human Services.
14. Jorgensen, Selena, Ruth Thorlby, Robin M. Weinick, and John Z. Ayanian. (2010). Responses of Massachusetts Hospitals to a State Mandate to Collect Race, Ethnicity and Language Data from Patients: A Qualitative Study. *BMC Health Services Research* 10.352. Web. 20 Sept. 2013.
15. Karliner, Leah S., Eliseo J. Pérez-Stable, and Ginny Gildengorin. (2004). The Language Divide: the Importance of Training in the Use of Interpreters for Outpatient Practice. *Journal of General Internal Medicine*, 19(2), 175-183.
16. Long, Judith A., Melissa I. Bamba, Bruce Ling, and Judy A. Shea. (2006). Missing Race/Ethnicity Data in Veterans Health Administration Based Disparities Research: A Systematic Review. *Journal of Health Care for the Poor and*

- Underserved* 17.1:128-40. Web. 19 Sept. 2013.
17. Nepaul, Ava N., Margaret M. Hynes, and Alison Stratton. (2007). *The Collection of Race, Ethnicity, and Other Sociodemographic Data in Connecticut Department of Public Health Databases*. Hartford, CT: Connecticut Department of Public Health.
 18. New Zealand Ministry of Health. (2004). *Ethnicity Data Protocols for the Health and Disability Sector*. Wellington, N.Z.: New Zealand Ministry of Health. 32.
 19. Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity. *62 Federal Register* 210 (October 30, 1997). p. 58782.
 20. Rothkopf, Jennifer, Katie Brookler, Sandeep Wadhwa, Michael Sajovetz. (2011). Medicaid Patients Seen at Federally Qualified Health Centers Use Hospital Services Less Than Those Seen by Private Providers. *Health Affairs (Project Hope)*, 30(7), 1335. doi:10.1377/hlthaff.2011.0066
 21. Shin, Hyon B., and Jennifer M. Ortman. (2011). Language Projections: 2010 to 2020. *Federal Forecasters Conference*. Washington, DC. 1-12.
 22. Smith, Felicia D., Meghan Woo, and S. B. Austin. (2010). 'I Didn't Feel Like Any of Those Things Were Me': Results of a Qualitative Pilot Study of Race/Ethnicity Survey Items With Minority Ethnic Adolescents in the USA. *Ethnicity & Health* 15.6: 621-38. Web. 21 Sept. 2013.
 23. Thorlby, Ruth, Selena Jorgensen, Bruce Siegel, and John Z. Ayanian. (2011). How Health Care Organizations Are Using Data on Patients' Race and Ethnicity to Improve Quality of Care. *The Milbank Quarterly* 89.2: 226-55. Web. 22 Sept. 2013.
 24. U.S. Census Bureau. (2012a). *U.S. Census Bureau Projections Show a Slower Growing, Older, More Diverse Nation a Half Century From Now*. Web. 1 Apr. 2014.

25. U.S. Census Bureau (2012b). American Community Survey, 2012 American Community Survey 1-Year Estimates, Table S0201. Generated by Roberto Henry. Using American FactFinder. <<http://factfinder2.census.gov>>. 21 April 2014.
26. U.S. Department of Health and Human Services. (2011). *HHS Action Plan to Reduce Racial and Ethnic Disparities: A Nation Free of Disparities in Health and Health Care*. Washington, D.C.: U.S. Department of Health and Human Services.
27. Varcoe, Colleen, Annette J. Browne, Sabrina Wong, and Victoria L. Smye. (2009). Harms and Benefits: Collecting Ethnicity Data in a Clinical Context. *Social Science & Medicine* 68.9: 1659-666. Print.
28. Wynia, Matthew K., Susan L. Ivey, and Romana Hasnain-Wynia. (2010). Collection of Data on Patients' Race and Ethnic Group by Physician Practices. *New England Journal of Medicine* 362.9: 846-50. Web. 20 Sept. 2013.

Appendix A

Survey Questionnaire
Confidential

Page 1 of 5

Race, Ethnicity and Primary Language Information Collection in Healthcare Settings

Dear Participant,

I am a 2nd year Master of Public Health student at the University Of Connecticut Health Center doing research as part of my thesis, which is required for graduation. The principal investigator for this study is my advisor, Dr. David Gregorio

The title of this study is: Attitudes and Beliefs Related to the Collection of Race, Ethnicity and Primary Language Information in Healthcare Settings. I am studying the feelings of patients associated with questions related to their race, ethnicity and primary language in health care settings, as well as their level of comfort providing this information when is it requested by their health care providers and/or insurance plans.

You are invited to participate in this study. Your participation is voluntary, and personally identifiable information will not be collected. If you choose to participate, please complete the following survey. While complete surveys will provide better data, you may skip questions that you do not feel comfortable answering. Your response to these questions should take between 10 and 15 minutes to complete. Completion of this survey signifies your consent to participate in the study.

Upon completing the survey, enter the validation code in order to receive payment of \$0.50.

You may withdraw from the survey at any time.

I appreciate your participation in this research study. For any questions regarding the study, please feel free to e-mail me at rhenry@mph.uchc.edu.

Thank you.

Sincerely,

Roberto A. Henry

David Gregorio, Ph.D., M.S.

What is your age?

- 18-25
- 26-35
- 36-45
- 46-55
- 56-65
- 66+

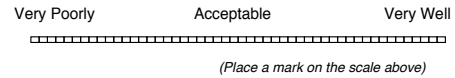
Are you a resident of the state of Connecticut?

- Yes
- No

What is your ethnicity?

- Hispanic or Latino (A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race.)
- Not Hispanic or Latino
- Rather Not Respond

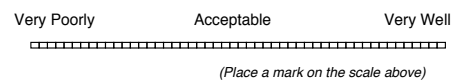
How well do these options for ethnicity represent you?



What is your race?(Select all that apply)

- American Indian or Native American
- Asian
- Black or African American
- Native Hawaiian or Pacific Islander
- White
- Other
- Rather Not Respond

How well do these options for race represent you?



Is English your primary language?

- Yes
- No

If not, what is your primary language?

What is the highest educational level you have achieved?

- Less than a High School Diploma
- High School Diploma/GED
- Some College/Technical School
- Associates Degree
- Bachelors Degree
- Graduate/Professional Degree

What is your current health insurance status?(Select all that apply)

- Medicaid
- Medicare
- Private Insurance
- No Insurance

Have you ever been asked for your race and/or ethnicity in a healthcare setting (insurance enrollment, hospital or clinic registration, etc...)?

- Yes
- No

Have you ever been asked for your primary language in a healthcare setting (insurance enrollment, hospital or clinic registration, etc...)?

- Yes
- No

Have you ever chosen not to answer questions about your race/ethnicity in a healthcare setting?

- Yes
- No

Why or why not?

What do you believe this information about your race and ethnicity is used for?

Have you ever chosen not to answer questions about your primary language in a healthcare setting?

- Yes
- No

Why or why not?

What do you think the information about your primary language is used for?

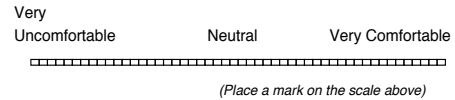
Has anyone, in a healthcare setting, ever explained to you the reasons for collecting your race and ethnicity information?

- Yes
- No

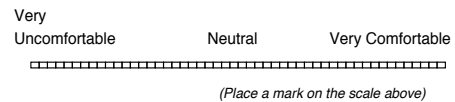
If yes, did the explanation change your mind about providing the information about your race and ethnicity?

- Yes
- No

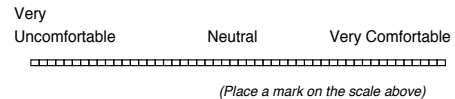
How comfortable do you feel providing information about your race, ethnicity and primary language IN PERSON?



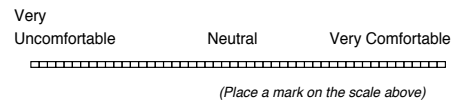
How comfortable do you feel providing information about your race, ethnicity and primary language OVER THE PHONE?



How comfortable do you feel providing information about your race, ethnicity and primary language IN WRITING or ON A FORM IN A WAITING ROOM?



How comfortable do you feel providing information about your race, ethnicity and primary language ONLINE?



Rate your agreement or disagreement with the following statements.

I am unsure about the reasons why the doctors need to know my race or ethnicity.

Strongly Disagree Undecided Strongly Agree

=====

(Place a mark on the scale above)

I worry that information about my race or ethnicity will be used against me.

Strongly Disagree Undecided Strongly Agree

=====

(Place a mark on the scale above)

I am afraid to provide my race and ethnicity when the doctor's office asks.

Strongly Disagree Undecided Strongly Agree

=====

(Place a mark on the scale above)

I trust that my race and ethnicity information will not be misused.

Strongly Disagree Undecided Strongly Agree

=====

(Place a mark on the scale above)

My race and ethnicity are relevant to my medical treatment.

Strongly Disagree Undecided Strongly Agree

=====

(Place a mark on the scale above)

I find questions about my race and ethnicity in a healthcare setting to be annoying.

Strongly Disagree Undecided Strongly Agree

=====

(Place a mark on the scale above)

I think information about my race and ethnicity will be helpful for the doctors.

Strongly Disagree Undecided Strongly Agree

=====

(Place a mark on the scale above)

Questions about my race and ethnicity on medical forms make me uncomfortable.

Strongly Disagree Undecided Strongly Agree

=====

(Place a mark on the scale above)

Appendix B

Attitude and Belief Scales for Cronbach's Alpha and Summary Scores

Attitudes Statements:

- “I worry that information about my race or ethnicity will be used against me.”
- “I am afraid to provide my race and ethnicity when the doctor's office asks.”
- “I find questions about my race and ethnicity in a healthcare setting to be annoying.”
- “Questions about my race and ethnicity on medical forms make me uncomfortable.”
- “I trust that my race and ethnicity information will not be misused.”
 - Responses to this statement were recoded to indicate their agreement/disagreement with the inverse of the original statement, their distrust.

Belief Statements:

- “Race and ethnicity have no connection to health conditions.”
- “Primary language has no connection to the quality of a person's health.”
- “There is no reason why a healthcare provider needs information about race and ethnicity.”
- “Asking for information about race and ethnicity is an invasion of one's privacy.”
- “My race and ethnicity are relevant to my medical treatment.”
 - Responses to this statement were recoded to indicate their agreement/disagreement with the inverse of the original statement, that this information is irrelevant to their treatment.
- “I think information about my race and ethnicity will be helpful for doctors.”
 - Responses to this statement were recoded to indicate their agreement/disagreement with the inverse of the original statement, thinking that the information is not helpful to doctors.

Appendix C

Framework for Future Study

		General	
	Point of Enrollment	Point of Encounter	
	<i>Anxiety Producing Situations</i>		
Population Demographics	Insured and Uninsured	Insured and Uninsured	Insured and Uninsured
	Age, Race, Ethnicity, SES	Age, Race, Ethnicity, Language, SES	Age, Race, Ethnicity, Language, SES
Data Collection Method	Paper Based Survey	Paper Based Survey	Online Survey
	Community Health Centers Patient Navigators	Community Health Centers Hospitals	Crowdsourcing Platforms Social Media Listservs
Sampling/Data Sources			Paper Based Survey Community Based Organizations Professional Organizations Religious Organizations Public Libraries
			Structured Interviews Follow-up with survey participants
			Focus Groups Community Based Organizations Professional Organizations Religious Organizations

Tables

Table 1
Study Participant Demographics

	Frequency	Percentage
N	104	100%
Age		
18-25	18	17.31%
26-35	26	25.00%
36-45	32	30.77%
46-55	14	13.46%
56-65	12	11.54%
66+	2	1.92%
CT Resident		
Yes	19	18.27%
No	84	80.77%
Missing	1	0.96%
Ethnicity		
Hispanic	12	11.54%
Not Hispanic	87	83.65%
Rather Not Respond	4	3.85%
Missing	1	0.96%
Race *		
American Indian or Native American	3	2.88%
Asian	16	15.38%
Black or African American	11	10.58%
Native Hawaiian or Pacific Islander	1	0.96%
White	74	71.15%
Other	2	1.92%
Rather Not Respond	2	1.92%
Missing	1	0.96%
English=First Language		
Yes	98	94.23%
No	5	4.81%
Missing	1	0.96%

(Table Continues)

	Frequency	Percentage
Education		
Less Than High School Diploma	3	2.88%
High School Diploma/GED	24	23.08%
Some College/Technical School	23	22.12%
Associates Degree	8	7.69%
Bachelors Degree	34	32.69%
Graduate/Professional Degree	12	11.54%
Insurance Status**		
Medicaid	9	8.65%
Medicare	13	12.50%
Private Insurance	70	67.31%
No Insurance	23	22.12%

Notes. *=Percentages do not add up to 100% because participants were asked to select all races that apply.

**=Percentages do not add up to 100% because some participants had multiple insurances.

Table 2

Adequacy of Representation by OMB Standard Classifications: Frequencies and Univariate Analysis

	Well	Poorly	N	Mean*
Ethnicity	66 (66.67%)	33 (33.33%)	99	55.85 (<i>s</i> =31.73)
Race	92 (91.09%)	9 (8.91%)	101	80.07 (<i>s</i> =23.02)

Note: *Scale of 0-100 for perception of how well standards represent respondents. 0=Very Poorly, 100= Very Well.

Table 3
Comfort Self-Reporting REL Based on Setting

	Comfortable	Uncomfortable	N	Mean*
In Person	96 (94.12%)	6 (5.88%)	102	80.21 (<i>s</i> =20.97)
Over the Phone	88 (86.27%)	14 (13.73%)	102	74.33 (<i>s</i> =23.08)
On a Form/In a Waiting Room	92 (89.32%)	11 (10.68%)	103	80.30 (<i>s</i> =21.71)
Online	93 (91.17%)	9 (8.82%)	102	79.80 (<i>s</i> =22.26)

Note: *Scale of 0-100 for level of comfort self-reporting. 0=Very Uncomfortable, 100= Very Comfortable.

Table 4
Agreement/Disagreement with Statements About Attitudes Towards Collection of REL

	Agree	Disagree	N	Mean*
Negative				
<i>Worry</i> info about race/ethnicity will be used against me	38 (37.62%)	63 (62.38%)	101	38.93 (s=26.22)
<i>Afraid</i> to provide my race/ethnicity when asked by provider	20 (20.00%)	80 (80.00%)	100	27.47 (s=24.87)
Questions about my race/ethnicity in healthcare settings are <i>annoying</i>	40 (40.00%)	60 (60.00%)	100	41.92 (s=27.98)
Questions about race/ethnicity on medical forms make me <i>uncomfortable</i>	22 (21.78%)	79 (78.22%)	101	31.82 (s=25.37)
Positive				
I <i>trust</i> that my race/ethnicity info will not be misused	86 (86.00%)	14 (14.00%)	100	68.67 (s=22.20)

Note: *Scale of 0-100 for agreement with statements. 0=Strongly Disagree, 100= Strongly Agree.

Table 5
Agreement/Disagreement with Statements About Beliefs Regarding Collection of REL

	Agree	Disagree	N	Mean*
Negative				
Race/ethnicity have <u>no connection</u> to health	41 (41.00%)	59 (59.00%)	100	46.04 (s=27.43)
Language has <u>no connection</u> to quality of health	74 (73.27%)	27 (26.73%)	101	66.15 (s=28.42)
Race/ethnicity affect <u>insurance premiums</u>	60 (60.00%)	40 (40.00%)	100	49.68 (s=22.57)
There is <u>no reason</u> why providers need race/ethnicity info	42 (41.58%)	59 (58.42%)	101	46.24 (s=26.27)
Asking for info about race/ethnicity is an <u>invasion of privacy</u>	30 (30.00%)	70 (70.00%)	100	38.41 (s=26.03)
Positive				
My insurance company <u>should know</u> my race/ethnicity	74 (74.75%)	25 (25.25%)	99	58.64 (s=25.26)
My race/ethnicity are <u>relevant</u> to my treatment	72 (71.29%)	29 (28.71%)	101	58.56 (s=25.78)
Info about my race/ethnicity is <u>helpful</u> for doctors	92 (91.09%)	9 (8.91%)	101	68.41 (s=19.38)
Ambiguous				
Race/ethnicity <u>affect the quality/type</u> of care given	50 (50.51%)	49 (49.49%)	99	48.05 (s=24.29)
I am <u>unsure</u> about why doctors need to know my race/ethnicity	58 (58.00%)	42 (42.00%)	100	50.84 (s=27.37)

Note: *Scale of 0-100 for agreement with statements. 0=Strongly Disagree, 100= Strongly Agree.

Table 6

T-Tests: Agreement with Negative Attitudes & Beliefs Towards REL Collection by Race

	White			Minority			<i>t</i>	<i>df</i>	<i>p</i>
	<i>n</i>	<i>Mean*</i>	<i>SD</i>	<i>n</i>	<i>Mean*</i>	<i>SD</i>			
Negative Attitudes (Summary Score)	71	14.6662	9.4611	30	22.9686	9.5128	4.01	54.372	.0002
Negative Beliefs (Summary Score)	71	26.5169	10.8881	30	28.2733	8.8726	0.85	66.494	.3997

Note: *Scale of 0-50 for agreement with negative attitude statements. 0=Strongly Disagree, 50= Strongly Agree. Scale of 0-60 for agreement with negative belief statements. 0=Strongly Disagree, 60= Strongly Agree.