Reviewer's report

Title: Analysis of five-year trends in self-reported language preference and issues of item non-response among Hispanic persons in a large cross-sectional health survey: implications for the measurement of an ethnic minority population

Version: 1 Date: 26 January 2010

Reviewer: Maureen O'Leary

Reviewer's report:

Analysis of five-year trends in self-reported language preference and issues of item non-response among Hispanic persons in a large cross-sectional health survey: implications for the measurement of an ethnic minority population.

General comments

This paper examines methodological issues pertaining to the inclusion of the Hispanic population in a large repeat cross-sectional health survey in the United States. It specifically looks at the proportional representativeness of Hispanics in the sample and it also investigates differences in number of call attempts required to complete the survey and completeness of responses by language preference for those Hispanics who took part.

The paper is well written, with an easy clear style of writing; although perhaps it could be more concise.

The results of these analyses seem to indicate that language is not really a barrier to participation in the survey or uptake of the survey or completeness of responses, with those whose preference is to take part in Spanish being easier to contact and having fewer instances of non-response for a greater number of questions. However the study cannot demonstrate causality; it does not demonstrate that providing a Spanish version of the survey increased the completeness of the responses, all it demonstrates is that those with a Spanish language preference were more likely to respond to a greater number of questions. However those English speaking v’s Spanish speaking participants may be characterised by different levels of educational attainment or employment level, this could explain some of the differences.

The use of the term acculturation in the discussion / conclusions is a bit vague, it may be more helpful for the authors to spell out what they mean by this, namely that surveys should consider country of birth and number of years in the US, to determine if these impact on participation in the survey and quality of data collected – discretionary revision.

Title

Ok
Abstract

Conclusions:

Regarding the conclusion that increasing the sample size would result in more precise findings – a minimum of 8500 Hispanic persons were included in the survey over the five years covered in the paper. Therefore the study is already highly powered to detect differences between Hispanics and other ethnic groups. Given this, although I am not a statistician and although this paper does not specify what health outcomes were being measured or the sample sizes required to detect these outcomes, I would question to what degree increasing the sample size would increase precision. Rather, increasing the sample size to reflect the proportion of Hispanics in the general population would increase the representativeness of the sample – Minor essential revision

Background:

The authors state that by 2050 the number of Hispanic persons in the US will double; however it is obvious from table 1 that the overall size of the US population is also increasing, therefore the main issue is not the increasing numbers of Hispanics, rather it is the degree of change in the size of the Hispanic population as a proportion of the overall US population – Minor essential revision.

Regarding the statement: “researchers have reported that Spanish language preference is associated with barriers to access and use of health care services, receiving less-efficient care and participation in risky health behaviour such as binge drinking”; although the paper focuses on the Hispanic population, barriers to accessing health care and participation in risky health behaviours are not just confined to Spanish speaking communities; these issues have been documented among other migrant and minority ethnic groups in many countries. Therefore this study has a wider relevance to migrant groups in general. Although language does play a role, so does overall literacy, educational attainment and socio-economic status, lower levels of which may be a feature of first generation migrants in comparison to the general population. Cultural practices in the country of origin can also have an impact – i.e. higher prevalence of smoking etc.

Regarding the demonstrated differences in health outcomes for the Hispanic population compared to other ethnic populations, could the authors please elaborate on this. I see no mention of it in reference 10 – Major compulsory revision.

The background and rationale for the study are well explained. The overall objectives of the study are explained with greater clarity in the abstract than in the background.

I wonder if the description of the outcomes measured in the study are better placed in the methods section of the paper – discretionary revision.
Methodology

Full methodology is available on the web; however this is quiet detailed and therefore it is not easy to extract the methodological information of relevance to this paper. The paper would therefore benefit from a short explanation of the study methods as they relate to the outcomes discussed in this paper. For instance how were participants selected? Was data collected on country of birth? Regarding number of attempts to complete the survey – does this refer to the number of calls to make the initial contact, calls purely to complete the questionnaire or both? Also when were calls made? How many attempts were made before someone was deemed non-contactable? – Major compulsory revision.

Analyses were restricted to those who defined themselves as Hispanic. Surely a sensible starting point for the paper is to demonstrate a difference between the outcomes for the overall Hispanic sample compared to the total sample and to the other major ethnic groups; demonstrating such a difference provides an overall justification for looking at the impact of language preference among Hispanics as an explanation for this. Within the context of this survey, if Hispanics are easier to contact and provide more complete data than the general sample, then why ask these questions at all? Also it would be interesting to know if those with an English language preference are more similar to the general population or to other ethnic groups, assuming that English language preference is an indicator of having been born in the USA or residing in the USA for many years – discretionary revision.

Even though the authors state that analyses are limited to Hispanics, they then proceed to describe the first analysis, a comparison between Hispanics, African-Americans and whites – minor compulsory revision.

The overall description of the different analyses is very clear.

Results

Consistency in reporting of results, some figures are rounded up to whole numbers, some reported to one decimal place – minor compulsory revision.

Paragraph 1:

The authors initially report the increasing numbers of Spanish language surveys between 2003 and 2007. Perhaps this has been reported elsewhere already; however it would be interesting to know if the proportion of Hispanics consenting to take part increased with the introduction of the Spanish version of the survey in 2003 – major compulsory revision.

The authors state that the number of Spanish-language surveys increased by almost 34% between 2003 and 2007. By my calculations the number of surveys increased by 51%; I think the authors used the 2007 total number of surveys as the denominator in their calculation instead of the 2003 number. They appear to have made this error for all of the results reported in paragraph 1 – major
The authors report the proportional increase in the numbers of Spanish language surveys and in the overall numbers of Hispanics participating in the survey compared to the increase in the size of the Hispanic population. Although this in itself is interesting, it does not account for changes in the overall size of the US population and in the overall sample size for the survey. Even though the numbers of Hispanics being surveyed has increased by about 14% over the 5-years, and even though the overall sample size has increased, there has been no proportional increase in the numbers of Hispanics in the survey. Hispanics still account for approximately 7% of the sample and are therefore under-represented in the sample overall, despite the fact that they make up an increasing proportion of the overall population. This is not emphasised in the results, the reader is left to work this out for themselves from Table 1 – major compulsory revision.

Paragraph 2: The authors compared the characteristics of those Hispanic participants who responded in English to those who responded in Spanish by age and gender. Although these are important comparisons, to my mind a comparison of being born in the USA or not, educational attainment and employment status are equally important. It is possible that preference for use of Spanish language is associated with being a first generation migrant. Migrant communities in general are characterised by lower levels of educational attainment, lower employment status and lower socioeconomic status than the general population. These are likely to be associated with number of call attempts to complete the survey as first generation migrants may have different patterns of employment. Lower levels of literacy and educational attainment may impact on the response rate to specific questions. The authors themselves have previously reported differences in educational attainment and income by language preference – major compulsory revision.

No comparison appears to have been made between the non-response to questions and number of call attempts to complete the survey among the Hispanic respondents overall and those among the African-American respondents, the white respondents and the overall survey population. The results of these analyses by language preference would be more meaningful if they were couched within the context of there being a difference between this sub-group, the other ethnic subgroups and the overall study population – discretionary revision.

Paragraph 3: Although the differences in non-response by language preference are interesting, what do these differences mean? It is difficult to draw conclusions for the individual questions. Probably what is of interest is the overall completeness of the questionnaire. Were those with Spanish language preference more likely to complete more of the questionnaire than those with English language preference? From the analysis presented, you could say yes, but it is only a difference of one question. Is this really meaningful? – discretionary revision.
Paragraph 4: The reported result of a difference in the number of call attempts for completion of the survey by language preference would have been strengthened by conducting an analysis that controlled for the effect of educational attainment, employment level and if possible being born in the USA or not – major compulsory revision.

Discussion:

“The analysis sheds light on how the BRFSS is capturing a representative sample of an ethnic minority population by adapting to a cultural need within that population and conducting Spanish language surveys” – it is difficult to draw this conclusion without a comparison to the proportion of Spanish participants before the introduction of the Spanish language questionnaire. Spanish speakers as a proportion of all Hispanics has remained more or less constant over the 5 years reported in the survey. Furthermore Hispanics as a proportion of the overall sample has also remained more or less constant and they are still somewhat under-represented in the survey. Interestingly the proportional increase in African Americans participating in the survey over the 5-year period was even greater than for Hispanics, indicating that any increase in participation may be due to overall strategies to target minority ethnic groups, rather than the provision of Spanish language versions of the survey – major compulsory revision.

“The Spanish language survey was able to increase the amount of health care data collected, which is necessary to study disparate outcomes for the Hispanic population in the United States.” – this implies causality and I don’t think you prove that here, it could well be that those with a Spanish language preference have different characteristics which would impact on the amount of time they have to answer questions and to participate in the survey. If you compared the completeness of the data before the introduction of the Spanish version of the survey to that collected after its introduction, then you could draw this conclusion – major compulsory revision.

“The findings from our work have shown that language preference, one measure of acculturation, affects not only health outcomes but also survey outcomes such as item non-response and number of contact attempts for survey completion.” – How have the authors demonstrated differences in health outcomes in this paper? The authors demonstrate that those with an English language preference are more difficult to contact and provide less information – is this what they expected? What does this mean from the point of view of conducting surveys among the Hispanic community? – major compulsory revision

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: Yes, but I do not feel adequately qualified to assess the
statistics.

Declaration of competing interests:
I declare that I have no competing interests