

BMJ Open Random sample community-based health surveys: does the effort to reach participants matter?

Antoine Messiah,¹ Grettel Castro,² Pura Rodríguez de la Vega,² Juan M Acuna²

To cite: Messiah A, Castro G, Rodríguez de la Vega P, *et al*. Random sample community-based health surveys: does the effort to reach participants matter?. *BMJ Open* 2014;**4**:e005791. doi:10.1136/bmjopen-2014-005791

► Prepublication history for this paper is available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2014-005791>).

Received 26 May 2014
Revised 5 November 2014
Accepted 6 November 2014



CrossMark

¹INSERM Research Unit U-1178, "Mental Health and Public Health", Hôpital Paul Brousse, Villejuif, France
²Department of Medical and Health Sciences Research, Herbert Wertheim College of Medicine, Florida International University, Miami, Florida, USA

Correspondence to
Dr Antoine Messiah;
antoine.messiah@inserm.fr

ABSTRACT

Objectives: Conducting health surveys with community-based random samples are essential to capture an otherwise unreachable population, but these surveys can be biased if the effort to reach participants is insufficient. This study determines the desirable amount of effort to minimise such bias.

Design: A household-based health survey with random sampling and face-to-face interviews. Up to 11 visits, organised by canvassing rounds, were made to obtain an interview.

Setting: Single-family homes in an underserved and understudied population in North Miami-Dade County, Florida, USA.

Participants: Of a probabilistic sample of 2200 household addresses, 30 corresponded to empty lots, 74 were abandoned houses, 625 households declined to participate and 265 could not be reached and interviewed within 11 attempts. Analyses were performed on the 1206 remaining households.

Primary outcome: Each household was asked if any of their members had been told by a doctor that they had high blood pressure, heart disease including heart attack, cancer, diabetes, anxiety/depression, obesity or asthma. Responses to these questions were analysed by the number of visit attempts needed to obtain the interview.

Results: Return per visit fell below 10% after four attempts, below 5% after six attempts and below 2% after eight attempts. As the effort increased, household size decreased, while household income and the percentage of interviewees active and employed increased; proportion of the seven health conditions decreased, four of which did so significantly: heart disease 20.4–9.2%, high blood pressure 63.5–58.1%, anxiety/depression 24.4–9.2% and obesity 21.8–12.6%. Beyond the fifth attempt, however, cumulative percentages varied by less than 1% and precision varied by less than 0.1%.

Conclusions: In spite of the early and steep drop, sustaining at least five attempts to reach participants is necessary to reduce selection bias.

BACKGROUND

Health surveys conducted on community-based random samples are essential when one wishes to investigate all aspects of health matters. In

Strengths and limitations of this study

- This study tackles an issue rarely addressed in the medical literature: What is the amount of effort to reach participants that is necessary to avoid selection bias in community-based health surveys with random sample?
- The survey was based on a random sample of households in an underserved and understudied population; state-of-the-art methodology is seldom applied in research on non-majority populations.
- In this study, the proportion of four of seven major health outcomes decreased significantly as the effort increased, suggesting that the healthiest households were the most difficult to reach.
- Results also indicated that sustaining at five attempts would be sufficient to avoid effort-related selection bias.
- The survey was conducted by face-to-face interview with a population of North Miami-Dade County, Florida, USA; results might differ in other settings or with other modes of data collection.
- The results might be biased by a 42% non-participation rate; sample and targeted population, however, were similar regarding few variables on which sample and census data could be meaningfully compared.

these surveys, participants are reached independently of any particular health condition or any healthcare system utilisation. However, such surveys often need to implement a complex sampling methodology including pre-stratification and post-stratification, multiple stages, unequal selection probabilities, clusters, resampling, etc, and often a mix of these, to cite the most frequent situations^{1–7}; in turn, survey planning, implementation and analyses are time-consuming and expensive. In addition, these can be further resource consuming because of the numerous visits (face-to-face surveys) or calls (telephone surveys) required before obtaining an interview. It is, therefore, legitimate to seek to minimise this number and consider as 'unreachable' households or individuals that have not been investigated after a number of attempts arbitrarily limited—in

other words, setting a survey effort limit. But in doing so, if it turns out that the distribution of some variables of interest is effort dependent, it would be necessary to maintain the effort at a high level. To the best of our knowledge, studies that have provided answers to this issue are scanty and give inconsistent results.^{2 4 8–20} Furthermore, face-to-face survey costs and durations are strongly linked to the fieldwork due to the canvassing that these imply. Thus, there is a gap of knowledge regarding (1) the health outcomes and frequency estimates that are effort dependent and (2) the adequate setting of effort limit.

In order to fill this gap, we conducted a face-to-face random sample health survey among the general population of North Miami-Dade County, South Florida, for which the number of visit attempts before getting an interview was recorded; we used this number to assess the effort level. In this article, we report frequencies of interviews and also seven major health outcomes, analysed by effort level.

METHODS

Sampling and fieldwork

The survey was the first step of community-oriented health initiatives launched by the Florida International University (FIU) College of Medicine, targeted at an understudied population located in North Miami-Dade County, Florida, where most residents are minority members. The survey's area corresponded to the zip codes of patients covered by the North Jackson Hospital, which belongs to the FIU health system. From this area, survey boundaries were drawn in order to match census tracts or census blocks. This area was considered by the hospital administrators as underserved; census data confirmed that 87–100% of residents belonged to minorities. A list of 2200 addresses was obtained from the Miami-Dade County Public Housing and Community Development Office by random sampling of residential homes located in the area.

A team of 20 interviewers was hired from the same minorities as the targeted communities and trained for 5 weeks to administer the survey. To make residents aware of the survey and facilitate interview acceptance, letters were handed to all selected households by the interviewers explaining the objectives of the survey and its subsequent community-oriented health initiatives. If no household member was present, the letter was slipped under the door. Interviewers wore specific vests as well as official badges. No interview took place at this stage.

Interviews were conducted during the weeks following the letter remittance, between October 2009 and April 2010. Interviewers were pair teamed and canvassed the area in successive waves. A household was deemed unreachable if neither interview nor refusal could be obtained within 11 waves, whether or not in-person contact had been established during the prior waves. The wave number (1–11) served us as the measurement of effort level.

Data collection and analysis

Data were collected via face-to-face interviews with one self-selected household member, acting as informant for the entire household (referred to as 'informants' in the remainder of this article). Questions addressed the household situation regarding demographics, education, financial and housing situation, involvement in neighbourhood activities and neighbourhood connectedness, perception of community facilities and services, health and access to healthcare, and health-related behaviours. In particular, households were asked if any of their members had been told by a doctor that they had high blood pressure, heart disease including heart attack, cancer, diabetes, anxiety/depression, obesity or asthma. For the purpose of this article, we analysed answers to these questions as a function of number of canvass waves needed to obtain the interview, using logistic regression (qualitative variables) or linear regression (quantitative variables) to elicit significant variations. The statistical unit of analysis was the household (randomly selected), except for few demographic characteristics of the informant (self-selected). Analyses were performed using SPSS V.19.0.

RESULTS

Of the 2200 addresses, 30 corresponded to empty lots and 74 were abandoned houses, yielding an updated sample of 2096 households. Of these, 625 (29.8%) declined to participate and 265 (12.6%) could not be reached and interviewed within 11 canvassing waves. At a given wave, interviewers had in-person contact with 12–21% of households who were not interviewed at that wave.

Sociodemographic characteristics of the final sample are shown in the last column of [table 1](#). A majority of informants were African-Americans; education of almost half of them had stopped during or at the end of high school, and 6 of 10 were employed; most of them were females, in their mid-40s. Forty-four per cent of households had an annual income below \$30 000, and 4 of 10 households had no health insurance. Few household sociodemographic characteristics of our sample were also available in the 2010 census tracts and blocks data; comparing them showed no to moderate discrepancies ([table 2](#)). Some, but not all combinations of race and ethnicity were available with comparable categories in the sample and the census.

The frequencies of interviews obtained at each wave dropped dramatically as the number of wave increased, from 381 at the first wave to 8 at the 11th wave ([figure 1](#)). Based on a final N=1206, the return per wave fell below 10% after four rounds, below 5% after six rounds and below 2% after eight rounds. Based on the number of households remaining to be interviewed after each wave (1471 at wave 1, 1090 at wave 2, 833 at wave 3, etc), the relative return per wave fell below 20% after four waves and below 10% after eight waves ([figure 2](#)).

Table 1 Sample sociodemographic characteristics and day and time of interview by canvassing round; p value of β coefficient (qualitative variable: logistic regression; quantitative variable: linear regression)

	Canvassing rounds						p Value	Total N=1206
	Round 1 N=381	Round 2 N=257	Round 3 N=193	Rounds 4–5 N=193	Rounds 6–7 N=94	Rounds 8–11 N=88		
Household informant								
Age (mean)	47	47	47	46	47	46	0.302	47
Gender: female (%)	60	58	62	55	60	67	0.631	60
Married/living with someone (%)	48	44	47	47	47	44	0.801	46
Education: high school or less (%)	47	50	48	48	50	38	0.488	48
African-American (%)	55	53	54	52	45	55	0.277	53
Employed (%)	50	59	62	63	68	71	0.000	59
Household								
Annual income: <\$ 30 000 (%)	52	36	44	44	44	34	0.036	44
Household size (mean)	3.6	3.5	3.6	3.6	3.2	3.0	0.011	3.5
Children 0–11 years old (%)	37	34	40	39	34	32	0.757	37
Absence of health insurance (%)	40	38	40	36	44	32	0.408	39
Day and time of interview								
Weekday daytime (before 17:00; %)	79	78	72	83	83	89		79
Weekday evening (17:00 or after; %)	8	7	11	4	3	10	0.064	8
Weekend (%)	13	15	17	13	14	1		13

For subsequent analyses, we merged rounds 4–5, 6–7 and 8–11 in order to study survey variables by effort-to-reach using adequate statistical power and precision. As the effort increased, household size (number of persons per household) decreased, while household income, and percentage of informants active and employed, increased (table 1). Other sociodemographic characteristics were stable across levels of effort. Proportion of all household health conditions diminished as the effort increased (figure 3); five of seven health conditions showed statistically significant trends (table 3); four of which were still significant after adjustment for household size, household income and employment status. Three health conditions varied

substantially: heart diseases (from 20.4% to 9.2%), anxiety/depression (from 24.4% to 9.2%) and obesity (from 21.8% to 12.6%).

The cumulative effect of subsequent attempts on the proportion of conditions and their SEs showed that neither the proportions nor their precision varied substantially ($\leq 1\%$ and $\leq 0.1\%$, respectively) beyond the fifth attempt (table 4).

DISCUSSION

To the best of our knowledge, the survey practitioner is left with very little guidance when it comes to making a decision on how much effort and resources should be invested in order to reach and interview potential participants. In spite of attempts to use standardised terminology to designate the various sources of survey non-response,²¹ the academic literature uses inhomogeneous wording for non-response due to reaching difficulties: ‘not at home’, ‘inaccessible’, ‘non-contact’ or ‘failure to deliver’.^{4 5 12} In addition, there is no indication regarding the amount of effort beyond which a household could qualify under one of these terms. In turn, little is known about non-response as a function of reaching effort.

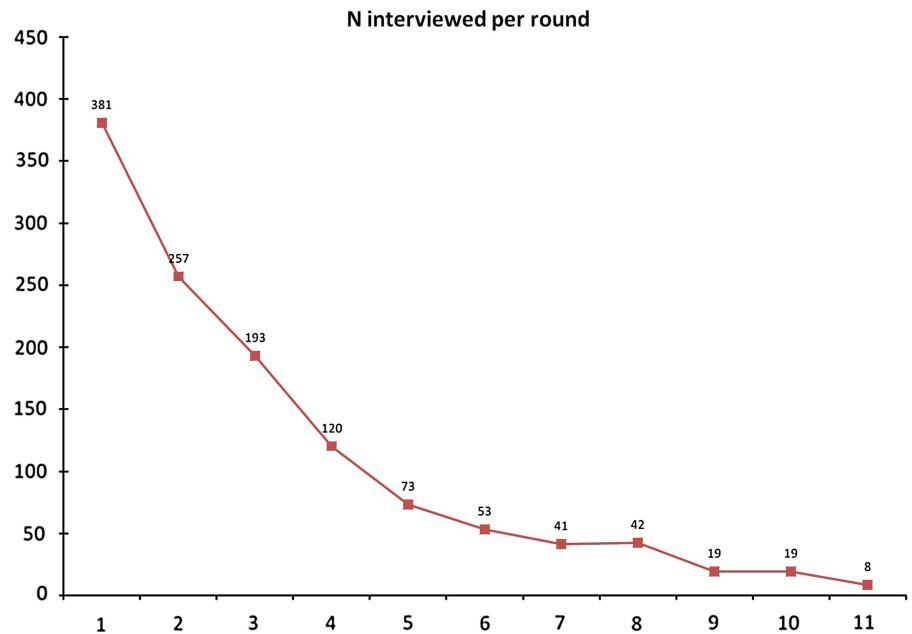
Nevertheless, since the inferential paradigm of random sampling demands full completion to guarantee that survey estimates are unbiased, survey researchers are prescribed to maximise response rates.^{2 15 22} How high the response rate should be, however, has been answered so diversely (from 50% to 85%)²² that survey practitioners are likely to be confused. Furthermore, empirical evidence has shown that response rate is a weak predictor of non-response bias.^{8 11 22–26} The amount of effort that lowers non-response bias under an acceptable level is therefore unknown, but indicators of representativeness

Table 2 Sociodemographic characteristics of households: final sample (point estimate and 95% CI) versus 2010 census tracts and blocks covering the sample area

Variable	Sample (95% CI)	Census (2010)
Speaks Spanish, %	25 (23 to 28)	27
Speaks English, %	68 (66 to 71)	68
Household average size	3.5 (3.4 to 3.6)	3.3
Hispanic households, %	32 (29 to 35)	33
White non-Hispanic households, %	3 (2 to 4)	6
African-American household, %	53 (50 to 56)	NA
Black (non-Hispanic and Hispanic) household, %	NA	74
Other race/ethnicity households, %	12 (10 to 14)	6

NA, not applicable.

Figure 1 Number of completed interviews (vertical axis) by canvassing round (horizontal axis).



that can inform fieldwork decisions are now under development.²⁴ The issue is aggravated by the hypothesis that measurement bias might increase with effort to reach participants: late respondents, who are less available and/or more reluctant to take the survey—but will eventually take it—might be more likely to provide answers filled with measurement errors.¹⁵

Maximum number of attempts reported to contact potential participants vary considerably: from four¹⁹ to unlimited^{13 18} and there is no typical value in between.^{2 4 12 15 17 20} Reports, however, consistently show a dramatic decrease in completion rate from one attempt to the next; first attempts have yielded 30–70% of the final sample, and more than 80% were completed by the fourth attempt.^{2 4 12 13 15 18 19} Our results are in

line with these reports, suggesting that effort to reach and interview participants could be stopped earlier.

Hence, analyses of key variable estimates as a function of collection effort are all the more important. In our health survey, frequencies of most of the medical conditions diminished significantly as the effort increased, while sociodemographic characteristics stayed stable—except for household size, which is expectedly correlated with contact rate,⁷ employment status and income. This result suggests that non-participation at early stages is related to the survey topic, with healthier households participating at later stages. In contrast to the abundant literature on survey non-response, few health surveys compared late or hard-to-reach respondents to earlier or easier-to-reach respondents.^{10 18 27–29} Some found results

Figure 2 Percentage of completed interviews (vertical axis) by canvassing round (horizontal axis), with adjustment at each round for the total of household still available for interview.

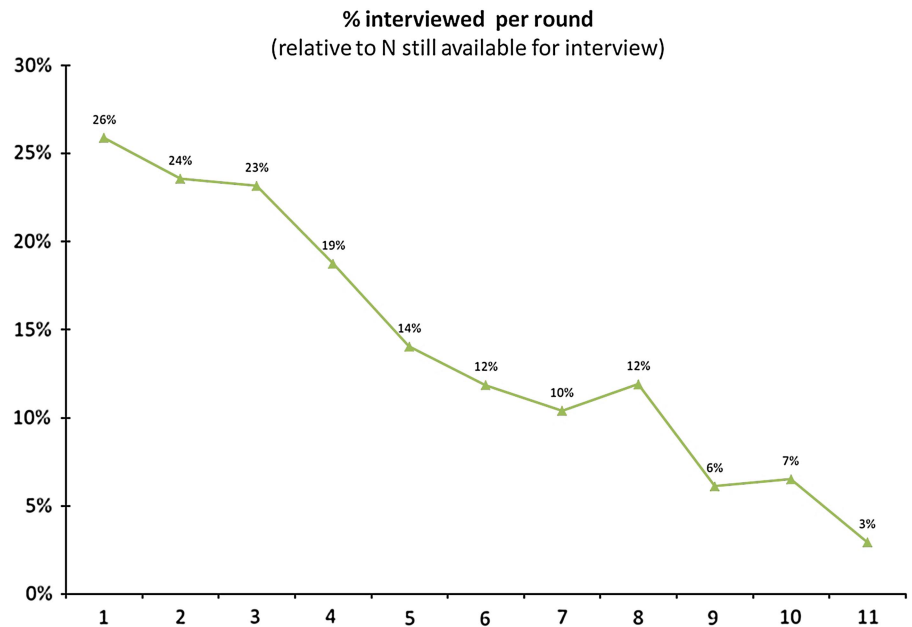
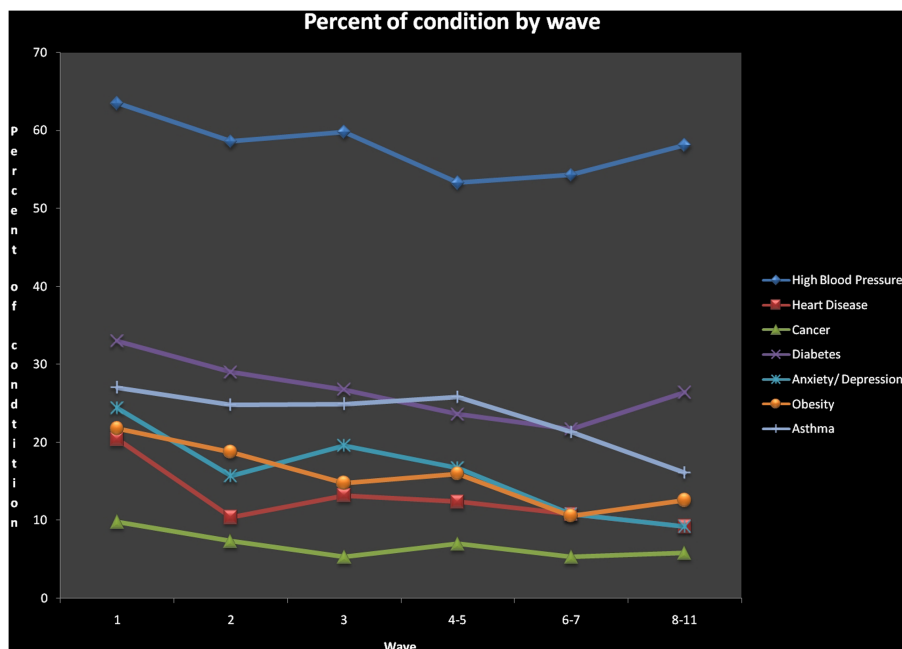


Figure 3 Reported conditions (percentage) by canvassing wave.



in line with ours,^{18 27} although to a lesser extent, but others did not or gave equivocal results.^{10 28 29} Since the literature on this subject is scanty, replication of similar analyses is needed.

While percentages of conditions varied from one round to the next, the same was not true for cumulative percentages, which represent the percentages that would have been obtained if the survey was stopped at each given round. Beyond the fifth attempt, percentages varied by less than 1%. Likewise, SEs of these percentages decreased by less than 0.1% beyond the fifth wave, suggesting that maintaining the effort up to five rounds might be enough.

In the absence of a sufficient body of knowledge, the health survey practitioner might err on the safe side and take our results for granted. Indeed, there is a rationale for such results: less healthy households

are likely to be interviewed at early stages of a home-based survey, since their members are more likely to be present at home and more inclined to take a survey that resonates with their current preoccupations. To some extent, such 'unhealthy at-home' effect can be seen as the reciprocal of the healthy worker effect.³⁰ Indeed, in our survey and in several large surveys of the British population, being employed and being hard to reach were found to be correlated.²⁰ In these British surveys, the participants who were easy-to-reach during daytime on weekdays were found to be less healthy than other participants; this supports an 'unhealthy at-home' effect. In our survey, however, increasing the reaching effort did not correspond to increasing evenings or weekend interviews: a non-significant trend ($p=0.064$) was found in the opposite direction.

Table 3 Percentage of reported condition by canvassing round, and unadjusted and adjusted OR per round (reference category: round 1; adjustment for household size; further adjustment on employment status and/or income if they remained significantly associated with the condition once household size was accounted for)

Reported condition	Canvassing rounds						Unadjusted OR	p Value	Adjusted OR	p Value
	Round 1 N=381	Round 2 N=257	Round 3 N=193	Rounds 4-5 N=193	Rounds 6-7 N=94	Rounds 8-11 N=88				
High blood pressure	64	59	60	53	54	58	0.93	0.045	0.92	0.025
Heart disease	20	10	13	12	11	9	0.84	0.002	0.87	0.019
Cancer	10	7	5	7	5	6	0.87	0.069	0.88	0.110
Diabetes	33	29	27	24	22	26	0.90	0.011	0.91	0.088
Anxiety/depression	24	16	20	17	11	9	0.83	<0.001	0.82	0.003
Obesity	22	19	15	16	11	13	0.86	0.002	0.86	0.003
Asthma	27	25	25	26	21	16	0.92	0.062	0.94	0.161

Table 4 Cumulative effect of successive attempts: cumulated percentage of reported condition by canvassing round and its SE

Round	Condition						
	High blood pressure	Heart disease	Cancer	Diabetes	Anxiety/depression	Obesity	Asthma
1	64±2.5	20±2.1	10±1.5	33±2.4	24±2.2	22±2.1	27±2.3
2	62±1.9	16±1.5	9±1.1	31±1.8	21±1.6	21±1.6	26±1.7
3	61±1.7	16±1.3	8±0.9	30±1.6	21±1.4	19±1.4	26±1.5
4–5	60±1.5	15±1.1	8±0.8	29±1.4	20±1.2	19±1.2	26±1.4
6–7	59±1.5	15±1.1	8±0.8	28±1.3	19±1.2	18±1.1	25±1.3
8–11	59±1.4	14±1.0	8±0.8	28±1.3	18±1.1	18±1.1	25±1.2

Our results might be due, however, to the aforementioned measurement bias. Since we dealt with self-reported conditions, our trends might reflect reporting trends rather than actual condition trends. Since late respondents are less available and/or more reluctant to take the survey, they might be inclined to provide answers which alleviate the survey burden—hence depicting the household health as better than it actually is. Although we cannot rule out such bias, it is unlikely to create our results entirely: (1) trends were dramatic for some conditions, much less so for others, with no rationale for such variations under a ‘declaration bias’ hypothesis; (2) for four of seven health conditions, the lowest frequencies corresponded to rounds 4–5 or rounds 6–7 rather than rounds 8–11. Likewise, our result might reflect trends in undiagnosed diseases, especially since the survey was conducted in an underserved population. This hypothesis, however, is not supported by the trends of sociodemographic variables: late responders were more likely to be employed and to have higher income than their earlier counterparts; health insurance (or absence thereof) was fairly constant from one wave to the next.

Another limitation is the 42.4% non-response rate (29.8% refusals, 12.6% non-interviewed at the 11th wave). Both rates are in the low end of the range of recent health surveys.^{7 8 10 31 32} Yet, if non-respondents and respondents differ systematically by health status, our results might be biased. However, these differences cannot induce a bias for the association between survey effort and health status unless variables on which non-respondents differ from respondents are correlated to both survey effort and health status, and are so in a way that distorts measures that were used to elicit the association.^{12 15 22} Although similarities between the sample and the targeted population support minimal bias, such a bias cannot be excluded. Indeed, variables for which we could perform valid comparisons between our sample and the census data were scanty, mainly because our sample was probabilistic at the household level only and collected data mostly at that level, while available census data are essentially based on individual characteristics. For the same reasons, valid weighting procedures to reduce bias could not be applied to our situation.

CONCLUSIONS

In spite of the steep drop in return after a limited amount of survey effort, it is important to sustain such effort in order to alleviate selection bias. Pilot surveys might help to determine the amount of effort beyond which results are unlikely to be affected.

Acknowledgements The authors thank the community members who graciously gave their time for our interviewers to complete the survey. They also acknowledge Florida International University Herbert Wertheim College of Medicine's Division of Research and Information and Data Coordinating Centre members for their services. The authors would like to thank Marcia H Varela for her help with earlier versions of the manuscript.

Contributors AM was responsible for the analysis design, conceptualisation of ideas, drafting of the manuscript and interpretation of the findings; GC and PRdIV supervised the field activities, participated in quality assurance in data collection, contributed to data management and data analysis of the study, and helped in the preparation of the Results section of the manuscript. JMA designed and conducted the original benchmark survey, was part of the discussion group for the research idea and has been a part of the review of the manuscript. All authors contributed to the drafting and revising of the manuscript and approved the final manuscript.

Funding This research was supported by the National Institute of Health (NIH) Grant 2P20MD002288 through its Center for Research on U.S. Latino HIV/AIDS and Drug Abuse (CRUSADA).

Competing interests None.

Patient consent Obtained.

Ethics approval The study was approved by Florida International University Institutional Review Board.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

REFERENCES

1. De Leeuw ED. *Data quality in mail, telephone and face to face surveys*. Amsterdam: TT-Publikaties, 1992.
2. Groves RM, Biemer PP, Lyberg LE, et al. *Telephone survey methodology*. New York: John Wiley & Sons, 1988.
3. Kalton G. *Introduction to survey sampling*. London: Sage Publications, 1983.
4. Cochran WG. *Sampling techniques*. New York: John Wiley & Sons, 1977.
5. Särndal CE, Swensson B, Wretman J. *Model assisted survey sampling*. Paris: Springer Verlag, 1992.
6. Lee ES, Forthofer RN, Lorimer RJ. *Analysing complex survey data*. London: Sage Publications, 1989.

7. Van der Heyden J, Demarest S, Van Herck K, *et al.* Association between variables used in the field substitution and post-stratification adjustment in the Belgian health interview survey and non-response. *Int J Public Health* 2014;59:197–206.
8. Davern M. Nonresponse rates are a problematic indicator of nonresponse bias in survey research. *Health Serv Res* 2013;48:905–12.
9. Joseph CL, Saltzgaber J, Havstad SL, *et al.* Comparison of early-, late-, and non-participants in a school-based asthma management program for urban high school students. *Trials* 2011;12:141.
10. Mannetje A, Eng A, Douwes J, *et al.* Determinants of non-response in an occupational exposure and health survey in New Zealand. *Aust N Z J Public Health* 2011;35:256–3.
11. Davern M, McAlpine D, Beebe TJ, *et al.* Are lower response rates hazardous to your health survey? An analysis of three state telephone health surveys. *Health Serv Res* 2010;45(5 Pt 1):1324–44.
12. Groves RM, Fowler FJ Jr, Couper MP, *et al.* *Survey methodology*. 2nd edn. Hoboken, NJ: John Wiley & Sons, 2009.
13. Aday LA, Cornelius LJ. *Designing and conducting health surveys: a comprehensive guide*. 3rd edn. San Francisco, CA: Jossey-Bass, 2006.
14. Holle R, Hochadel M, Reitmeir P, *et al.*; for the KORA Group. Prolonged recruitment efforts in health surveys: effects on response, costs, and potential bias. *Epidemiology* 2006;17:639–43.
15. Olson K. Survey participation, nonresponse bias, measurement error bias, and total bias. *Public Opin Q* 2006;70:737–58.
16. Keeter S, Miller C, Kohut A, *et al.* Consequences of reducing nonresponse in a national telephone survey. *Public Opin Q* 2000;64:125–48.
17. Firdion JM. The call rank and partner presence effects in a telephone survey. In: Bozon M, Léridon H, eds. *Sexuality and the social sciences: a French survey on sexual behaviour*. Dartmouth: Aldershot, 1996:107–40.
18. Mishra SI, Dooley D, Catalano R, *et al.* Telephone health surveys: potential bias from noncompletion. *Am J Public Health* 1993;83:94–9.
19. Tennant A, Badley EM. A confidence interval approach to investigating non-response bias and monitoring response to postal questionnaires. *J Epidemiol Community Health* 1991;45:81–5.
20. Hall J, Brown V, Nicolaas G, *et al.* Extended field efforts to reduce the risk of non-response bias: have the effects changed over time? Can weighting achieve the same effects? *Bull Sociol Methodol Méthodol Sociol* 2013;117:5–25.
21. AAPOR (American Association for Public Opinion Research). Standard Definitions Final Dispositions of Case Codes and Outcome Rates for Surveys [Internet]. 2011 [cited 15 October 2014]. http://www.aapor.org/AM/Template.cfm?Section=Standard_Definitions2&Template=/CM/ContentDisplay.cfm&ContentID=3156
22. Groves RM. Nonresponse rates and nonresponse bias in household surveys. *Public Opin Q* 2006;70:646–75.
23. Groves RM, Peytcheva E. The impact of nonresponse rates on nonresponse bias: a meta-analysis. *Public Opin Q* 2008;72:167–89.
24. Shlomo N, Skinner C, Schouten B. Estimation of an indicator of the representativeness of survey response. *J Stat Plan Inference* 2012;142:201–1.
25. Stang A, Jockel KH. Studies with low response proportions may be less biased than studies with high response proportions. *Am J Epidemiol* 2004;159:204–10.
26. Voigt LF, Boudreau DM, Weiss NS, *et al.* Studies with low response proportions may be less biased than studies with high response proportions. *Am J Epidemiol* 2005;161:401–2; author reply 402–3.
27. Kotaniemi JT, Hassi J, Kataja M, *et al.* Does non-responder bias have a significant effect on the results in a postal questionnaire study? *Eur J Epidemiol* 2001;17:809–17.
28. Korkeila K, Suominen S, Ahvenainen J, *et al.* Non-response and related factors in a nation-wide health survey. *Eur J Epidemiol* 2001;17:991–9.
29. Etter JF, Perneger TV. Analysis of non-response bias in a mailed health survey. *J Clin Epidemiol* 1997;50:1123–8.
30. Li CY, Sung F-C. A review of the healthy worker effect in occupational epidemiology. *Occup Med* 1999;49:225–9.
31. Mindell J, Aresu M, Becares L, *et al.* Representativeness of participants in a cross-sectional health survey by time of day and day of week of data collection. *Eur J Public Health* 2012;22:364–9.
32. Van Loon AJM, Tjihuis M, Picavet HSJ, *et al.* Survey non-response in the Netherlands: effects on prevalence estimates and associations. *Ann Epidemiol* 2003;13:105–10.