

*Short Report*

**Impact of resurvey non-response on the associations between established risk factors and cardiovascular disease mortality: prospective cohort study**

*Impact of resurvey non-response in a prospective cohort study*

G. David Batty<sup>a</sup>  
Catharine R. Gale<sup>b</sup>

<sup>a</sup>MRC Social & Public Health Sciences Unit, Glasgow, UK

<sup>b</sup>MRC Epidemiology Resource Centre, University of Southampton, Southampton, UK

Word count: 1375 (abstract 231)

Correspondence: David Batty, MRC Social & Public Health Sciences Unit, University of Glasgow,  
4 Lilybank Gardens, Glasgow, UK, G12 8RZ. E. [david-b@sphsu.mrc.ac.uk](mailto:david-b@sphsu.mrc.ac.uk)

## **Abstract**

*Background.* Selection bias in observational epidemiology – the notion that people who participate in a study are fundamentally different to those that do not – is a perennial concern. In cohort studies, a potentially important but little-investigated manifestation of selection bias is the distortion of the exposure-disease relationship according to participation status.

*Methods.* Seven years after the original UK Health and Lifestyle Survey (HALS1), attempts were made to resurvey participants (HALS2). In 6484 men and women who took part in both surveys, we compared the baseline characteristics, mortality experience following the completion of HALS2, and, finally, the baseline risk factor-cardiovascular disease (CVD) mortality gradients in HALS2 non-participants and participants.

*Results.* Resurvey non-participants were, based on data from HALS-1, younger, of lower social class, had a lower prevalence of hypertension or self-reported limiting long-standing illness, but a higher prevalence of psychological distress ( $p\text{-values} \leq 0.027$ ). The risk of death from future CVD was significantly higher in those that did not participate in HALS-2. However, the magnitude of the association between a series of risk factors and CVD mortality was essentially the same in the HALS-2 non-participants and participants ( $p\text{-value for interaction} \geq 0.108$ ).

*Conclusion.* In the present cohort study, non-response at resurvey did not bias the observed associations between baseline risk factors and later cardiovascular disease mortality. Future studies should examine the impact of non-response to baseline surveys on these relationships.

## Introduction

In prospective cohort studies, bias introduced by systematic (i.e., non-random) non-response is a perennial problem, and one that is likely to become more acute with secular declines in survey participation over recent decades.[1,2] Widely referred to as selection bias, this phenomenon has several manifestations. First, it may lead to error in the estimation of risk factor or disease prevalence at study baseline. Second, it may result in inaccuracy in the assessment of disease rates during follow-up. The very nature of non-attendance complicates empirical examination of the impact of these types of bias. However, this has been overcome by investigators using a number of approaches: gathering background information on non-participants from routine data sources (e.g., drivers licence records[3] or mandatory census[4]); comparing the characteristics of study members who took part in all phases of data collection (e.g., medical examination *and* questionnaire administration) versus those who only partially participated (e.g., questionnaire administration);[5,6] or, contrasting the characteristics of participants in a baseline survey according to whether or not they took part in subsequent resurvey(s).[7]

Based on these methods, there is some evidence that risk factor or disease prevalence at study baseline differ according to response status, but the study findings are inconclusive. Thus, in comparison to responders, non-responders may be younger[3,8] or older,[7,9,10] and have greater[11-13] or similar morbid load. There is a much greater degree of consistency as regards smoking[7,12,14] and socio-economic disadvantage[4,8,13-16] both of which appear to be more common in non-participants. Despite this apparently discordant literature, with very few exceptions,[17] a series of studies examining death rates after baseline survey have reported elevated rates of future mortality in non-participants relative to participants.[5,6,18-22]

Although epidemiologists commonly examine differences between respondents and non-respondents in levels of risk factors, and, where possible, mortality, few studies have examined a

third manifestation of selection bias: whether survey participation status affects estimates of the association between risk factors and disease. In a cohort study of cancer incidence, the magnitude of the association between cigarette smoking and carcinoma of the lung,[5] and between body mass index and colon cancer,[3,5] was essentially the same according to survey participation status. In two studies that assessed risk of cardiovascular disease (CVD) mortality according to participation status, the strength of the association between body mass index and death from CVD was similar in non-responders and responders in a older US female population.[3] In a cohort of men and women from Finland, similar results were reported when socio-economic disadvantage was the exposure of interest.[4]

To our knowledge, the impact of non-response on the association of other well established[23] – measured blood pressure, alcohol intake, physical activity and common mental disorder with CVD mortality – has yet to be examined. This is the primary purpose of the present analyses.

## **Methods**

### *Baseline examination (HALS-1)*

The UK Health and Lifestyle Survey (HALS-1) was conducted in 1984/5. In a household-based random sample of 9003 adults aged 18 and over (77.5% of target population), interviews were administered and physical measurements were made in participants' homes.[24] The socio-economic profile of this sample is almost identical to that seen in the 1981 UK census data for both men (38.9% non-manual in HALS1 vs. 39% in the census) and women (59.2% non-manual in HALS1 vs. 60% in the census);[25] a similar level of agreement is apparent for age and ethnicity.[24] During the home-visits, enquiries were made about employment history, smoking habits, alcohol consumption, engagement in exercise activities such as keep fit, sports, jogging, swimming, cycling, or dancing in the past fortnight, long-standing illness or disability, and experience of heart attack, angina or a stroke, including treatments. Using standard protocols, blood

pressure, height and weight were collected on a further home-visit (7268 responded; 81% of those initially interviewed), and the respondents were invited to complete and mail back the 30-item General Health Questionnaire (GHQ-30)[26] which provides an assessment of psychiatric status (6317 responded; 70% of those interviewed). In total, 98% of those who took part in HALS1 were subsequently 'flagged' for mortality using the Office for National Statistics' NHS Central Register. Death certificates were obtained for those who died and causes of death were coded according to the International Classification of Diseases (9th revision).

### *Resurvey (HALS2)*

In 1991/2, approximately seven years after the first survey, a resurvey (HALS2) was carried out in order to describe changes that had occurred in health and lifestyle among the original respondents.[27] Of the 9003 men and women who took part in HALS1, 718 (8%) had died before the start of HALS2 in 1991, leaving 8285 as potential participants. Of these, 6484 (72% of study members) had complete data on levels of risk factors at baseline (social class, blood pressure, height and weight, self-reported illness, smoking habits, alcohol consumption, physical activity and GHQ-30 score) together with data on mortality, so could be included in our analyses.

CVD was coded according to the 9th revision of the International Classification of Diseases[28] (ICD-9 codes 390–434,436–448.30). Analyses are based on deaths that occurred over a maximum of 14 years between the start of the HALS2 survey in September 1991 and 30th May 2005. We used Cox proportional hazards regression models[29] to compute hazard ratios with accompanying 95% confidence intervals for the association of risk factors with CVD mortality.

## **Results**

Compared with those who participated in HALS-2, non-participants in the resurvey were, based on data from HALS-1, younger, of lower social class, had a lower prevalence of hypertension or self-

**Table 1: Baseline risk factors according to participation in the resurvey (HALS2) (n=6484)**

	Participated in resurvey		P value for difference
	Yes (n=4590)	No (n=1894)	
Sex – n (%)			0.859
Male	2054 (44.7)	843 (44.5)	
Female	2536 (55.3)	1051 (55.5)	
Age, years – n (%)			<0.001
≤32	1023 (22.3)	695 (36.7)	
33-44	1338 (29.2)	445 (23.5)	
45-60	1283 (28.0)	407 (21.8)	
>60	946 (20.6)	347 (18.3)	
Social class – n (%)			<0.001
I (high)	264 (5.8)	93 (4.9)	
II	1109 (24.2)	408 (21.5)	
III	603 (13.1)	280 (14.8)	
IV	1613 (35.1)	633 (33.4)	
V	745 (16.2)	322 (17.0)	
Unclassified/no occupation	221 (4.8)	117 (6.2)	
	35 (0.8)	41 (2.2)	
Hypertensive – n (%)			0.004
Yes	1790 (39.0)	666 (35.2)	
No	2800 (61.0)	1228 (64.8)	
Obese – n (%)			0.949
Yes	424 (9.2)	174 (9.2)	
No	4166 (90.8)	1720 (90.8)	
Current or former smoker – n (%)			0.253
Yes	2868 (62.5)	1212 (64.0)	
No	1722 (37.5)	682 (36.0)	
Heavy drinker – n (%)			0.297
Yes	1925 (41.9)	821 (43.3)	
No	2665 (58.1)	1073 (56.7)	
Physical activity in past fortnight – n (%)			0.732
Yes	2108 (45.9)	861 (45.5)	
No	2482 (54.1)	1033 (54.5)	
History of cardiovascular disease – n (%)			0.925
Yes	301 (6.6)	123 (6.5)	
No	4289 (93.4)	1771 (93.5)	
Limiting long-standing illness – n (%)			0.027
Yes	1464 (31.9)	551 (29.1)	
No	3126 (68.1)	1343 (70.9)	
Psychological distress – n (%)			0.006
Yes	980 (21.4)	464 (24.5)	
No	3610 (78.6)	1430 (75.5)	

Hypertension was defined as blood pressure  $\geq 130/85$  mmHg and/or use of anti-hypertensive drugs; obesity as a body mass index  $>30$  kg/m<sup>2</sup> according to WHO definitions;[30] and heavy drinking as a reported weekly alcohol consumption  $>21$  units for men and  $>14$  for women.[31] A score of 5 or more on the GHQ-30 was used to denote the presence of psychological distress.

reported limiting long-standing illness, but a higher prevalence of psychological distress, as indicated by a score of 6 or more on the GHQ (table 1).

Compared with those who participated in HALS-2, risk of death from all-causes and from cardiovascular disease was significantly higher in the non-participants: age- and sex-adjusted hazard ratios (95% CI) for all-cause and cardiovascular mortality were 1.39 (1.24, 1.57) and 1.28 (1.07, 1.55), respectively. Notably, these associations were little attenuated by further adjustment for the risk factors described in table 1: fully-adjusted hazard ratios (95% CI) for all-cause and cardiovascular mortality were 1.36 (1.21, 1.53) and 1.26 (1.04, 1.53), respectively.

In table 2 we show the relation between each of the baseline risk factors and mortality from cardiovascular disease in the participants compared with the non-participants. Among the participants, all the baseline risk factors were associated with an increased risk of death from cardiovascular disease in age- and/or sex-adjusted analyses with statistical significance apparent in most analyses. Some of these relationships were attenuated after mutual adjustment. Similar results were apparent among the non-participants such that no interaction term was statistically significant. A similar pattern of results was evident when total mortality was the outcome of interest (results available on request).

## **Discussion**

In the present study, risk factor and health indices at baseline differed according to participation at resurvey some seven years later. That non-participants at resurvey herein were, at baseline, younger, of lower socioeconomic status, and had elevated minor psychiatric morbidity, accords with most,[3,4,8,11-13] but not all,[7,9,10,14] studies. Despite these non-participants also having a lower prevalence of hypertension or self-reported limiting long-standing illness, in keeping with a series

**Table 2: Hazard ratios (95% confidence intervals) for the relation of baseline risk factors (HALS1) with cardiovascular disease mortality according to participation in resurvey (HALS2) (n=6484)**

	Participants in resurvey	Non-participants in resurvey	P-value for interaction
Number of death/number at risk	399/4590	157/1894	
<b>Age (per SD)</b>			
Sex-adjusted	7.62 (6.53, 8.88)	5.95 (4.82, 7.32)	0.557
Fully-adjusted <sup>a</sup>	5.99 (4.99, 6.99)	4.73 (3.74, 6.00)	0.441
<b>Female</b>			
Age-adjusted	0.56 (0.46, 0.68)	0.59 (0.43, 0.81)	0.417
Fully-adjusted	0.62 (0.50, 0.78)	0.72 (0.51, 1.02)	0.883
<b>Manual social class</b>			
Age & sex-adjusted	1.54 (1.25, 1.89)	1.46 (1.05, 2.04)	0.899
Fully adjusted	1.40 (1.14, 1.73)	1.37 (0.98, 1.93)	0.724
<b>Hypertensive</b>			
Age & sex-adjusted	2.14 (1.67, 2.75)	3.01 (1.97, 4.60)	0.331
Fully adjusted	1.97 (1.53, 2.54)	3.06 (2.00, 4.69)	0.744
<b>Obese</b>			
Age & sex-adjusted	1.82 (1.39, 2.38)	1.20 (0.76, 1.90)	0.108
Fully adjusted	1.43 (1.09, 1.88)	1.03 (0.65, 1.63)	0.285
<b>Current smoker vs never smoked</b>			
Age & sex-adjusted	1.58 (1.22, 2.06)	1.38 (0.92, 2.07)	0.430
Fully adjusted	1.55 (1.19, 2.03)	1.20 (0.79, 1.82)	0.126
<b>Ex smoker vs never smoked</b>			
Age & sex-adjusted	1.24 (0.96, 1.61)	1.25 (0.80, 1.93)	0.557
Fully adjusted	1.20 (0.92, 1.56)	1.16 (0.74, 1.82)	0.708
<b>Heavy drinker</b>			
Age & sex-adjusted	1.26 (1.03, 1.54)	0.95 (0.69, 1.31)	0.123
Fully adjusted	1.20 (0.98, 1.47)	0.87 (0.63, 1.20)	0.140
<b>No exercise in past fortnight</b>			
Age & sex-adjusted	1.82 (1.41, 2.34)	1.43 (0.94, 2.17)	0.696
Fully adjusted	1.49 (1.14, 1.93)	1.26 (0.83, 1.91)	0.221
<b>History of cardiovascular disease</b>			
Age & sex-adjusted	2.36 (1.87, 2.97)	1.94 (1.29, 2.91)	0.487
Fully adjusted	1.86 (1.46, 2.38)	1.57 (1.02, 2.41)	0.390
<b>Limiting long-standing illness</b>			
Age & sex-adjusted	1.68 (1.38, 2.05)	1.42 (1.03, 1.95)	0.328
Fully adjusted	1.39 (1.13, 1.72)	1.16 (0.80, 1.62)	0.296
<b>Psychological distress</b>			
Age & sex-adjusted	1.38 (1.09, 1.73)	1.49 (1.05, 2.10)	0.590
Fully adjusted	1.07 (0.86, 1.34)	1.50 (1.05, 2.16)	0.815

<sup>a</sup>Fully adjusted is adjustment for all other risk factors in the table. Referent categories are as for table 1, unless otherwise stated.

of other cohorts,[5,6,18,20-22,32] we observed an elevated risk of CVD and all-cause mortality at follow-up in resurvey non-participants relative to participants. Notably, however, this selection bias did not have an impact on the association of established and emerging risk factors with future total and CVD mortality. These results support those from the very few other studies on this topic, where obesity[3] and socio-economic disadvantage[4] were similarly predictive of CVD mortality according to survey response. While we examined the predictive value of a greater range of risk factors than was possible in these reports, a limitation of our study is that we were not able to directly explore the impact of non-response to the original (baseline) survey on the associations between risk factors and CVD deaths. However, as indicated, census comparison suggests that the men and women who took part in HALS1 were highly representative of the UK population.

In conclusion, in this prospective cohort study which has a typical level of participant non-response at resurvey, and in which there were expected risk factors and mortality differences according to participation status, there was no evidence that this selection bias actually modified the risk factor—CVD mortality association. This suggests that, in the present cohort study at least, non-response at resurvey does not bias the observed associations between baseline risk factors and later cardiovascular disease mortality.

### **What is already known**

- Investigators working on cohort studies typically attempt to explore selection bias by comparing the baseline characteristics and mortality experience of participants and non-participants.
- In general, the risk factor profile and mortality experience of non-participants is less favourable.
- However, very little is known about whether these differentials translate into substantially different risk factors—cardiovascular disease (CVD) associations in survey participants and non-participants.

### **What this paper adds**

- In a study which had a typical level of non-response to a resurvey, and in which there were expected differentials in both risk factors and mortality according to resurvey participation, there was in fact no evidence that this selection bias modified the risk factor—CVD mortality association.
- This suggests that, in the present study, non-response at resurvey does not bias the observed associations between baseline risk factors and later cardiovascular disease mortality.
- Future studies should examine the impact of non-response to baseline surveys on the associations between risk factors and CVD.

**Acknowledgements:** The Medical Research Council (MRC) Social and Public Health Sciences Unit receives funding from the MRC and the Chief Scientist Office at the Scottish Government Health Directorates. David Batty is a Wellcome Trust Fellow (WBS U.1300.00.006.00012.01). The MRC supports Catharine Gale.

**Conflict of interests:** None.

**Licence for Publication Statement:** The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence (or non exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd to permit this article (if accepted) to be published in JMG and any other BMJ PGL products and sublicences such use and exploit all subsidiary rights, as set out in our licence (<http://jmg.bmj.com/ifora/licence.pdf>)

## References

- (1) Morton LM, Cahill J, Hartge P. Reporting participation in epidemiologic studies: a survey of practice. *Am J Epidemiol* 2006; 163(3):197-203.
- (2) Hartge P. Raising response rates: getting to yes. *Epidemiology* 1999; 10(2):105-107.
- (3) Bisgard KM, Folsom AR, Hong CP, Sellers TA. Mortality and cancer rates in nonrespondents to a prospective study of older women: 5-year follow-up. *Am J Epidemiol* 1994; 139(10):990-1000.
- (4) Harald K, Salomaa V, Jousilahti P, Koskinen S, Vartiainen E. Non-participation and mortality in different socioeconomic groups: the FINRISK population surveys in 1972-92. *J Epidemiol Community Health* 2007; 61(5):449-454.
- (5) Heilbrun LK, Nomura A, Stemmermann GN. The effects of non-response in a prospective study of cancer: 15-year follow-up. *Int J Epidemiol* 1991; 20(2):328-338.
- (6) Heilbrun LK, Nomura A, Stemmermann GN. The effects of nonresponse in a prospective study of cancer. *Am J Epidemiol* 1982; 116(2):353-363.
- (7) Thomas MC, Walker M, Lennon LT, Thomson AG, Lampe FC, Shaper AG et al. Non-attendance at re-examination 20 years after screening in the British Regional Heart Study. *J Public Health Med* 2002; 24(4):285-291.
- (8) Eagan TM, Eide GE, Gulsvik A, Bakke PS. Nonresponse in a community cohort study: predictors and consequences for exposure-disease associations. *J Clin Epidemiol* 2002; 55(8):775-781.
- (9) Barchielli A, Balzi D. Nine-year follow-up of a survey on smoking habits in Florence (Italy): higher mortality among non-responders. *Int J Epidemiol* 2002; 31(5):1038-1042.
- (10) Chou P, Kuo HS, Chen CH, Lin HC. Characteristics of non-participants and reasons for non-participation in a population survey in Kin-Hu, Kinmen. *Eur J Epidemiol* 1997; 13(2):195-200.
- (11) Veenstra MY, Friesema IH, Zwietering PJ, Garretsen HF, Knottnerus JA, Lemmens PH. Lower prevalence of heart disease but higher mortality risk during follow-up was found among nonrespondents to a cohort study. *J Clin Epidemiol* 2006; 59(4):412-420.
- (12) Criqui MH, Barrett-Connor E, Austin M. Differences between respondents and non-respondents in a population-based cardiovascular disease study. *Am J Epidemiol* 1978; 108(5):367-372.
- (13) Drivsholm T, Eplov LF, Davidsen M, Jorgensen T, Ibsen H, Hollnagel H et al. Representativeness in population-based studies: a detailed description of non-response in a Danish cohort study. *Scand J Public Health* 2006; 34(6):623-631.
- (14) Shahar E, Folsom AR, Jackson R. The effect of nonresponse on prevalence estimates for a referent population: insights from a population-based cohort study. Atherosclerosis Risk in Communities (ARIC) Study Investigators. *Ann Epidemiol* 1996; 6(6):498-506.
- (15) Bergstrand R, Vedin A, Wilhelmsson C, Wilhelmsen L. Bias due to non-participation and heterogenous sub-groups in population surveys. *J Chronic Dis* 1983; 36(10):725-728.
- (16) Boshuizen HC, Viet AL, Picavet HS, Botterweck A, van Loon AJ. Non-response in a survey of cardiovascular risk factors in the Dutch population: determinants and resulting biases. *Public Health* 2006; 120(4):297-308.
- (17) Paul O, Lepper M, Whelan W, et al. A longitudinal study of coronary heart disease. *Circulation* 1963; 28:20.
- (18) Dawber TR, Kannel WB, Lyell LP. An approach to longitudinal studies in a community: the Framingham Study. *Ann N Y Acad Sci* 1963; 107:539-556.
- (19) Doll R, Bradford Hill A. Mortality in relation to smoking: ten years' observations of British doctors. *Br Med J* 1964; 1(5395):1399-1410.

- (20) Thomas MC, Walker M, Lennon LT, Thomson AG, Lampe FC, Shaper AG et al. Non-attendance at re-examination 20 years after screening in the British Regional Heart Study. *J Public Health Med* 2002; 24(4):285-291.
- (21) Jousilahti P, Salomaa V, Kuulasmaa K, Niemela M, Vartiainen E. Total and cause specific mortality among participants and non-participants of population based health surveys: a comprehensive follow up of 54 372 Finnish men and women. *J Epidemiol Community Health* 2005; 59(4):310-315.
- (22) Paul O, Lepper MH, Phelan WH, Dupertus GW, MacMillan A, McKean H et al. A longitudinal study of coronary heart disease. *Circulation* 1963; 28:20-31.
- (23) Yusuf S, Reddy S, Ounpuu S, Anand S. Global burden of cardiovascular diseases: part I: general considerations, the epidemiologic transition, risk factors, and impact of urbanization. *Circulation* 2001; 104(22):2746-2753.
- (24) Cox B, Blaxter M, Buckle A, Fenner M, Golding J, Gore M et al. The health and lifestyle survey. London: Health Promotion Research Trust; 1987.
- (25) Anon. Census 1981, economic activity, Great Britain. London: HMSO; 1984.
- (26) Goldberg D. The detection of psychiatric illness by questionnaire. London: Oxford University Press; 1972.
- (27) Cox B, Huppert F, Whichelow M. The Health and Lifestyle Survey: seven years on. Aldershot: Dartmouth; 1993.
- (28) Anon. Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death (ninth revision). Geneva: WHO; 1977.
- (29) Cox DR. Regression models and life-tables. *J R Stat Soc [Ser B]* 1972; 34:187-220.
- (30) World Health Organisation. Physical status: the use and interpretation of anthropometry: report of a WHO expert committee. Who Tech. Rep. Ser. Geneva: WHO; 1995.
- (31) Royal College of Physicians. A great and growing evil: the medical consequences of alcohol abuse. London: Royal College of Physicians; 1987.
- (32) Doll R, Bradford Hill A. Mortality in relation to smoking: ten years' observations of British doctors. *Br Med J* 1964; 1(5395):1399-1410.