Non-respondence to postal survey is related to excess mortality: A register-based analysis of the Health and Social Support (HeSSup) Cohort Study, Finland

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2011-000657
Article Type:	Research
Date Submitted by the Author:	25-Nov-2011
Complete List of Authors:	Suominen, Sakari; University of Turku, Department of Public Health Koskenvuo, Karoliina; The Social Insurance Institution of Finland, Sillanmaki, Lauri; University of Helsinki, Hjelt Institute, Department of Public Health Vahtera, Jussi; University of Turku, Department of Public Health Korkeila, Katariina; City of Raisio, Kivimaki, Mika; London Medical School, Department of Epidemiology and Public Health, Mattila, Kari; University of Tampere, Center of General Practice Virtanen, Pekka; Tampere School of Public Health, University of Tampere Sumanen, Markku; University of Tampere, Rautava, Paivi; University of Turku, Department of Public Health Koskenvuo, Markku; University of Helsinki, Hjelt Institute, Department of Public Health
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Research methods
Keywords:	EPIDEMIOLOGY, PUBLIC HEALTH, SOCIAL MEDICINE

Non-respondence to postal survey is related to excess mortality: A register-based analysis of the Health and Social Support (HeSSup) Cohort Study, Finland

Sakari Suominen contributed substantially to acquisition of data, the design of the study and to the interpretation of the data, wrote the first draft of the ms and revised it several times critcally for important intellectual content. He has approved the final version of the ms.

Karoliina Koskenvuo contributed substantially to acquisition of data, the design of the study, carried out the first data analyses, contributed substantially to the interpretation of the data and revised the ms critically several times for important intellectual content. She has approved the final version of the ms.

Lauri Sillanmäki contributed substantially to acquisition of data, the design of the study, carried out data analyses, contributed substantially to the interpretation of the data and revised the ms critically several times for important intellectual content. He has approved the final version of the ms.

Jussi Vahtera contributed substantially to the design of the study and to the interpretation of the data and revised the ms critically several times for important intellectual content. He has approved the final version of the ms.

Katariina Korkeila contributed substantially to acquisition of data, the design of the study and to the interpretation of the data and revised the ms critically several times for important intellectual content. She has approved the final version of the ms.

Mika Kivimäki contributed substantially to design of the study and to the interpretation of the data and revised the ms critically several times for important intellectual content. He has approved the final version of the ms.

Kari J. Mattila contributed substantially to acquisition of data, the design of the study and to the interpretation of the data and revised the ms critically several times for important intellectual content. He has approved the final version of the ms.

Pekka J. Virtanen contributed substantially to the interpretation of the data and revised the ms critically several times for important intellectual content. He has approved the final version of the ms.

Markku Sumanen contributed substantially to the interpretation of the data and revised the ms critically several times for important intellectual content. He has approved the final version of the ms.

Päivi Rautava contributed substantially to acquisition of data, the interpretation of the data and revised the ms critically several times for important intellectual content. She has approved the final version of the ms.

Markku Koskenvuo contributed substantially to acquisition of data, to the design of the study and to the interpretation of the data and revised the ms critically several times for important intellectual content. He has approved the final version of the ms.

ABSTRACT

Objective

To examine difference in mortality between postal survey respondents and non-respondents.

Design

A prospective cohort study with baseline survey in 1998 and comprehensive linkage to national mortality registers until 2005, the Health and Social Support (HeSSup) study.

Setting

A population based postal survey of the working aged population in Finland in 1998

Participants

The original random sample comprised 64,797 working-aged individuals in Finland (20-24, 30-34, 40-44, 50-54 years of age; 32,059 women and 32,716 men), yielding 25,898 (40.0%) responses in the baseline postal survey in 1998.

Primary outcome measure

Registry based primary causes of death encoded with the International Classification of Diseases (ICD-10)

Results

In women, hazard ratio for total mortality was 1.75 (95% confidence interval 1.40 - 2.19) times higher among the non-respondents compared to the respondents. In men, non-response was associated with a 1.41 fold (1.21 - 1.65) excess risk of total mortality. Non-response associated in certain age groups with deaths due to diseases in women and with deaths due to external causes in men. The most prominent excess mortality was seen for total mortality for both genders and for mortality due to external causes among men.

Conclusions

Postal surveys can still be considered a valid method for population based health research, but they result in slight underestimation of illness prevalence, especially illness involving mental health in men.

Word Count of Abstract 220

Keywords

Follow-up Studies, Non-respondents, Mortality, Health Surveys, Registries

Article summary

Article focus

Women and individuals from upper social strata tend to participate more actively in postal health surveys

What this exactly means in terms of health selection among respondents is unclear

Postal health surveys are believed to produce underestimates of illness prevalence

Key message

Total mortality was significantly higher among non-respondents compared to respondents during a 7 year follow-up among a total Finnish nationwide sample in working age comprising almost 65 000 individuals

The excess mortality observed was moderate. Among men it was explained by external causes, whereas among women it was due to diseases and was only seen in the age group 50 – 54 years

Postal surveys can still be considered a valid method for population based health research, but they result in slight underestimation of illness prevalence, especially illness involving mental health in men

Strengths and limitations of the study

The linkage to mortality data was successful for virtually all individuals of the original sample comprising nearly 65 000 individuals. The sample size secures the reliability of the conclusions drawn. Furthermore, the registry data on mortality in Finland can be considered as reliable. To the best of the authors' knowledge a corresponding study based on a as large a sample as in this study has not previously been carried out.

Inaccuracy of the final diagnosis of death is possible but nevertheless the data can be considered as reliable as registry data based on clinical work generally can become. Moreover, there is no support for the view that inaccuracy of diagnosis of death would systematically be associated with the response status of the study members.

INTRODUCTION

Word count without Abstract, Tables and References 1791

Response rates in health related postal surveys are declining in the Western world. In the Nordic countries the situation has been somewhat better compared to the rest of Europe and the U.S.A., but recently even there declining trends have been observed [1, 2]. Previous studies have suggested that women, older persons and persons from upper social strata tend to participate more actively in health surveys compared with the rest [3 - 11] However, not all studies have consistently supported these observations [10, 12, 13]. Furthermore, recent studies have extended analysis of non-response beyond demographic variables showing a lower rate of hospital admissions [2] and mortality among the respondents [14, 15].

Studies on causes for non-participation to health surveys have revealed incorrect address or incorrect delivery by post to contribute to some of the drop out [5, 8, 16]. The non-respondents themselves have reported various kinds of reasons for their behaviour, such as gaining no benefit for participation [5, 16], no interest in the topic [7, 8, 16, 17], feeling of intrusion of privacy [8], lack of time [7, 8, 17], forgetfulness [8], and present illness [17]. Surveys including questions on issues perceived as intimate have often decreased participation rates [18]. It has also been speculated that late respondents might resemble more the non-participants [5, 17, 19] but to date decisive evidence for this pattern is lacking [20]. In follow-up studies, participation has been explored according to the number of rounds

the individual has taken part in [14] and more occasions of non-response found to be positively associated with subsequent mortality rates. On the whole, health selection among participants might decrease the generalizability of prevalence estimates but this effect is until now not satisfactorily described [21]. However, even in the case of health selection the results related to the strength of association between the variables studied need not necessarily be biased [21].

A non-response analysis of the postal survey of the population-based Health and Social Support (HeSSup) study - which achieved a relatively modest (40 %) response rate [9] - replicated previous findings on the differences in demographic characteristics between respondents and non-respondents. The aim of the present study was to extend these analyses to explore whether survey non-respondents differ from the respondents in terms of mortality (all-cause, disease mortality, mortality for external causes).

MATERIALS AND METHODS

The study is based on the complete sample (N=64, 797) of the Health and Social Support (HeSSup) prospective mail survey initiated in 1998. The concurrent joint Ethics Committee of the University of Turku and the Turku University Central Hospital considered approval not necessary for a normal cohort study, but all participants were requested to sign a consent form containing information about the study and to grant permission to allow subsequent studies with the same

data set and possibility to link with national health registries. The sample represented the concurrent age groups of 20-24, 30-34, 40-44 and 50-54 years of age in Finland [9]. However, by purpose the Swedish speaking Finns (5% of the general population) as well as the Turku region were slightly overrepresented. Of the 64,797 persons, 22 could not be included in the present study since the follow-up had to be set to begin from the first death among respondents which was September 22nd 1998. Certain cases of deaths among non respondents had occurred already earlier and potentially before sending out the initial questionnaire. Totally 25,898 satisfactory responses were returned. In 2007, the survey material was by means of an unique social security number linked to Statistics Finland data on mortality between the years 1998 and 2005. Totally 1 174 cases of deaths among 25,290 observations that could be linked with registry data were identified. Moreover, mortality data of non-respondents from the same time period was likewise as for respondents obtained from Statistics Finland and further analyzed by age group and gender.

The outcome variable was primary causes of death encoded with the International Classification of Diseases (ICD-10). Mortality for external causes (S00 - Y98) and disease mortality (A00 – R99) were examined separately. The differences in mortality between non-respondents and respondents were analysed separately for women and men using Cox proportional hazards regression. The analyses were carried out by the research group. Hazard ratios (HR) with 95% confidence intervals (95% CI) for mortality (total, external causes,

disease) of non-respondents vs. respondents according to age (1998) were reported. Cox proportional hazard assumptions were tested by visual examination of log-minus-log plots showing parallelism among the selected strata variable (responding status). The statistical analysis was performed by the SAS ® software 9.2. for Windows.

RESULTS

Between the years 1998 to 2005, 1,174 individuals belonging to the complete sample died (Table 1). Of the deaths, 70% occurred in men.

Table 1. Demographics of respondents and non-respondents and number of deaths during follow-up from 1998 to 2005 (N and %).

	Respondents		Non-respondents		Total	
	N	%	N	%	N	%
Gender						
Women	14,922	59.0	17,137	43.4	32,059	49.5
Men	10,368	41.0	22,348	56.6	32,716	50.5
Total	25,290	100.0	39,485	100.0	64,775	100.0
Age in 1998						
20-24 y	6,783	26.8	9,405	23.8	16,188	25.0
30-34 y	5,981	23.7	10,267	26.0	16,248	25.1
40-44 y	6,073	24.0	10,198	25.8	16,271	25.1
50-54 y	6,453	25.5	9,615	24.4	16,068	24.8
Total	25,290	100.0	39,485	100.0	64,775	100.0
Number of deaths by gender	N	%	N	%	N	%
In women	110	0.7	240	1.4	350	1.1
In men	218	2.1	606	2.7	824	2.5
Total	328	1.3	846	2.1	1,174	1.8
Number of deaths by age in 1998						
20-24 y	21	0.3	47	0.5	68	0.4
30-34 y	38	0.6	105	1.0	143	0.8
40-44 y	85	1.4	252	2.5	337	2.1
50-54 y	184	2.9	442	4.6	626	3.9
Total	328	1.3	846	2.1	1,174	1.8
Number of deaths from external causes						
In women	26	0.2	48	0.3	74	0.2
In men	55	0.5	217	1.0	272	0.8
Total	81	0.3	265	0.7	346	0.5
Number of deaths from external causes by age in 1998						
20-24 y	10	<0.1	32	<0.1	42	<0.1
30-34 y	16	<0.1	66	0.2	82	0.1
40-44 y	22	<0.1	90	0.2	112	0.2
50-54 y	33	0.1	77	0.2	110	0.2
Total	81	0.3	265	0.7	346	0.5

Table 1 continued. Demographics of respondents and non-respondents and number of deaths during follow-up from 1998 to 2005 (N and %).

	Respondents		Non-respondents		Total	
	N	%	N	%	N	%
Number of deaths due to diseases						
In women	84	0.6	192	1.1	276	0.9
In men	163	1.6	389	1.7	552	1.7
Total	247	1.0	581	1.5	828	1.3
Number of deaths due to diseases by age in 1998						
20-24 y	11	<0.1	15	<0.1	26	<0.1
30-34 y	22	<0.1	39	<0.1	61	<0.1
40-44 y	63	0.2	162	0.4	225	0.3
50-54 y	151	0.6	365	0.9	516	0.8
Total	247	1.0	581	1.5	828	1.3

Total mortality was higher for non-respondent women in age group 50-54 years and for non-respondent men in age group 40-44 and 50-54 years and for each gender with all age groups combined. In analyses combining women and men, excess total mortality associated with non-response was observed in age groups 40-44 and 50-54 when age groups were examined separately as well as when all observations were combined (Table 2). Non-respondent men had a higher mortality for external causes in age groups 30-34 and 40-44 years and with all age groups combined. In analyses with genders combined, this was seen in age groups 30-34 and 40-44 years and when all observations were combined. In women, no statistically significant differences in mortality for external causes were detected (Table 3). Non-respondent women showed significantly higher disease-related mortality in age group 50-54 years as well as when all age groups were combined. The same held true for both genders when age groups 40-44 and 50-54 years were examined separately as well as when all

observations were combined. On the other hand, in separate analyses for men non-respondents showed a slightly increased disease mortality compared to respondents only when all age groups were combined (Table 4).

Table 2. Hazard ratios for total mortality of non-respondents vs. respondents and the 95% CIs according to gender and age at the beginning of the follow-up in 1998.

Non-respondents vs. respondents (=1.00)	Age in1998 20-24	Age in 1998 30-34	Age in 1998 40-44	Age in 1998 50-54	Total age or age & gender adjusted
Women	1.00 (0.34-2.97)	1.45 (0.77-2.73)	1.40 (0.93-2.10)	**2.17 (1.58-2.98)**	**1.75 (1.40-2.19)**
Men	1.28 (0.70-2.34)	1.41 (0.89-2.24)	**1.71 (1.25-2.35)**	**1.31 (1.07-1.61)**	**1.41 (1.21-1.65)**
Total gender or age & gender adjusted	1.21 (0.71-2.04)	1.42 (0.98-2.07)	**1.59 (1.24-2.04)**	**1.54 (1.29-1.83)**	**1.52 (1.34-1.73)**

Statistically significant associations are in bold.

Table 3. Hazard ratios for deaths due to external causes in non-respondents vs. respondents and the 95% CIs according to gender and age at the beginning of the follow-up in 1998.

Non-respondents vs. respondents (=1.00)	Age in 1998 20-24	Age in 1998 30-34	Age in 1998 40-44	Age in 1998 50-54	Total age or age & gender adjusted
Women	4.65 (0.52-41.62)	1.94 (0.68-5.50)	1.38 (0.58-3.28)	1.18 (0.57-2.45)	1.50 (0.93-2.42)
Men	1.36 (0.64-2.88)	**2.04 (1.07-3.90)**	**2.42 (1.37-4.28)**	1.63 (0.99-2.67)	**1.87 (1.39-2.52)**
Total gender or age & gender adjusted	1.61 (0.78-3.30)	**2.01 (1.16-3.49)**	**2.07 (1.30-3.31)**	1.47 (0.98-2.22)	**1.78 (1.39-2.29)**

Statistically significant associations are in bold.

Table 4. Hazard ratios for disease mortality of non-respondents vs. respondents and the 95% CIs according to gender and age at the beginning of the follow-up in 1998.

Non-respondents vs. respondents (=1.00)	Age in 1998 20-24	Age in 1998 30-34	Age in 1998 40-44	Age in 1998 50-54	Total age or age & gender adjusted
Women	0.39 (0.08-1.92)	1.21 (0.54-2.69)	1.41 (0.88-2.23)	**2.46 (1.73-3.52)**	**1.83 (1.41-2.36)**
Men	1.13 (0.40-3.18)	0.83 (0.42-1.66)	1.42 (0.97-2.07)	1.25 (0.99-1.56)	**1.25 (1.04-1.50)**
Total gender or age & gender adjusted	0.81 (0.37-1.80)	0.98 (0.58-1.67)	**1.41 (1.05-1.89)**	**1.55 (1.28-1.87)**	**1.42 (1.23-1.65)**

Statistically significant associations are in bold.

DISCUSSION

In this large population-based epidemiological study, comparison of mortality of non-respondents with respondents showed, as expected, a higher mortality among the former group. The differences between respondents and non-respondents, however, were not very large and of quite similar magnitude for both genders although with varying causes. Among women the greatest differences were seen in disease mortality in the oldest age group of initially 50–54 years of age. The greatest differences for men were caused by external causes of death and involved age groups 30- 34 and 40-44 years.

Postal surveys are frequently used in population based health research. The impact of health selection on the results has been studied but not very

extensively . An affirmative view on the potential demographic difference between respondents and non-respondents is still lacking. According to previous findings [3, 8, 9], respondents tend to be somewhat healthier and report a more favourable health behaviour compared to non-respondents. However, the bias caused by this was limited and applied mainly to health problems or risk behaviours that generally are not eagerly communicated to others, such as mental problems or binge drinking. From previous studies it is also known that women as well as those well off in the society tend to participate more actively in health related survey research on the whole. Hence, this might result in under-estimation of the prevalence of health problems common among men as well as individuals from lower social strata. Results from a previous non-response analysis of the initial phase of this study supports this view [9]. Women as well as indivduals with high level of education were somewhat over represented among respondents but no clear health-related selection could be shown. According to our present results 7 years later the potential health-related selection can be further clarified. In Finland, mortality for external causes is intimately associated with alcohol consumption and alcoholism [22]. As could be expected, we could see a significantly higher mortality for external causes among male non-respondents as compared with the respondents.

Moreover, in previous studies [23] it has also been pointed out that non-respondents are not necessarily a homogenic group but can differ internally, e.g. depending on to which wave of the survey if any they have taken part in. Also the

correspondence between late respondents and total non-respondents has been questioned [20].

Given the health selection related to postal surveys it has to be kept in mind that population studies usually do not focus solely on prevalence estimates anymore but more or less on potential risk or protective factors of certain health problems. From previous research [21] there are indications that even if prevalence estimates are not accurate the associations between the variables studies are not necessarily biased. This gives further justificaton to continue postal surveys in health research since the results can help us understand the etiology and prevention process of a certain kind of disease.

Strengths and limitations of the study

A major strength of this study is that the linkage to mortality data was successful for virtually all individuals of the original sample. Furthermore, the registry data on mortality in Finland could be considered as quite reliable. All deaths with suspicion of an external cause and in the age groups studied are investigated by autopsy. Naturally, inaccuracy concerning the final diagnosis might be present but nevertheless the data can be considered as reliable as registry data based on clinical medical work generally can become. Moreover, there is no reason to assume that problems in reliability of the diagnoses would systematically be associated with the response status of the study members. The legislation in

Finland enables use of mortality data for research purposes and thus we were able by ourselves to perform a mortality analysis for all respondents.

The large sample size secures that the conclusions drawn from the statistical analyses are reliable and can not have been caused by random effects. To the best of our knowledge a corresponding study based on as large a sample as in this study has previously not been carried out.

Conclusions

Total mortality was moderately higher for the non-respondents of a nationwide mail survey compared to the respondents in both genders. For women this was mostly due to disease mortality in age group 50-54 years but for men due to mortality for external causes in age groups 30-34 and 40-44 years. The most prominent excess mortality was seen for total mortality for both genders and for mortality due to external causes among men. Selection by health, especially mental health in men can cause bias in health related population surveys. However, this applies to prevalence estimates and does not necessarily jeopardize results from studies on risk and protective factors. Hence, it seems justified to continue postal surveys in health research in order to improve understanding of etiology and prevention of diseases.

REFERENCES

1. Helakorpi S., Prättälä R., Uutela A. Suomalaisen aikuisväestön terveyskäyttäytyminen ja terveys, kevät 2007. [Health Behaviour and Health among the Finnish Adult Population. Spring 2007] Publications of the National Public Health Institute. 2008;B 6.

2. Kjøller M, Thoning H. Characteristics of non-response in the Danish Health Interview Surveys, 1987-1994. *Eur J Public Health* 2005;**15**(5):528-35 doi: 10.1093/eurpub/cki023 [Advance Access published online July 28, 2005].

3. Cottler LB, Zipp JF, Robins LN, et al. Difficult-to-recruit respondents and their effect on prevalence estimates in an epidemiologic survey. *Am J Epidemiol* 1987;**125**(2):329-39.

4. Jacobsen BK, Thelle DS. The Tromsø Heart Study: responders and non-responders to a health questionnaire, do they differ? *Scand J Soc Med* 1988;**16**(2):101-4.

5. Bakke P, Gulsvik A, Lilleng P, et al. Postal survey on airborne occupational exposure and respiratory disorders in Norway: causes and consequences of non-response. *J Epidemiol Community Health* 1990;**44**(4):316-20.

6. Jooste PL, Yach D, Steenkamp HJ, et al. Drop-out and newcomer bias in a community cardiovascular follow-up study. *Int J Epidemiol* 1990;19(2):284-9.

7. Brussaard JH, Brants HA, Bouman M, et al. The study population: general characteristics and potential confounding factors. *Eur J Clin Nutr* 1997;**51** Suppl 3:S19-24.

8. Rönmark E, Lundqvist A, Lundbäck B, et al. Non-responders to a postal questionnaire on respiratory symptoms and diseases. *Eur J Epidemiol* 1999;**15**(3):293-9.

9. Korkeila K, Suominen S, Ahvenainen J, et al. Non-response and related factors in a nation-wide health survey. *Eur J Epidemiol*. 2001;**17**(11):991-9.

10. van den Akker M, Buntinx F, Metsemakers JF, et al. Morbidity in responders and non-responders in a register-based population survey. *Fam Pract* 1998;**15**(3):261-3.

11. Sonne-Holm S, Sørensen TI, Jensen G, et al. Influence of fatness, intelligence, education and sociodemographic factors on response rate in a health survey. *J Epidemiol Community Health* 1989;**43**(4):369-74.

12. Asch DA, Jedrziewski MK, Christakis NA. Response rates to mail surveys published in medical journals. *J Clin Epidemiol* 1997;**50**(10):1129-36.

13. Heilbrun LK, Ross PD, Wasnich RD, et al. Characteristics of respondents and nonrespondents in a prospective study of osteoporosis. *J Clin Epidemiol* 1991;**44**(3):233-9.

14. Ferrie JE, Kivimäki M, Singh-Manoux A, et al. Non-response to baseline, non-response to follow-up and mortality in the Whitehall II cohort. *Int J Epidemiol* 2009;**38**(3):831-7 doi:10.1093/ije/dyp153 [Advance Access publication 5 March 2009].

15. Mattila VM, Parkkari J, Rimpelä A. Adolescent survey non-response and later risk of death. A prospective cohort study of 78 609 persons with 11-year follow-up. *BMC Public Health* 2007;**7**(1):87 doi:10.1186/1471-2458-7-87 [Published: 22 May 2007].

16. Foster K, Campbell D, Crum J, et al. Non-response in a population study after an environmental disaster. *Public Health* 1995;**109**(4):267-73.

17. Janzon L, Hanson BS, Isacsson SO, et al. Factors influencing participation in health surveys. Results from prospective population study 'Men born in 1914' in Malmö, Sweden. *J Epidemiol Community Health* 1986;**40**(2):174-7.

18. Dunne MP, Martin NG, Bailey JM, et al. Participation bias in a sexuality survey: psychological and behavioural characteristics of responders and non-responders. *Int J Epidemiol* 1997;**26**(4):844-54.

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**(1):81-5.

20. Lahaut, V. M. H. C. J., Jansen HAM, van de Mheen D, et al. Estimating non-response bias in a survey on alcohol consumption: comparison of response waves. *Alcohol Alcohol* 2003;**38**(2):128-34 [doi:10.1093/alcalc/agg044].

21. Martikainen P, Laaksonen M, Piha K, et al. Does survey non-response bias the association between occupational social class and health? *Scand J Public Health* 2007;**35**(2):212-5 [doi:10.1080/14034940600996563].

22. Koskenvuo M, Kaprio J, Kesaniemi A, et al. Alcohol-related diseases associated with ischaemic heart disease: a three-year follow-up of middle-aged male hospital patients. *Alcohol Alcohol* 1986;**21**(3):251-6.

23. Boersma F, Eefsting JA, van den Brink W, et al. Characteristics of non-responders and the impact of non-response on prevalence estimates of dementia. *Int J Epidemiol* 1997;**26**(5):1055-62.

Acknowledgements

The study was supported by the Academy of Finland (three different decisions on funding)

The authors do not declare any competing interests

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *cohort studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	3
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	3
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	5
Objectives	3	State specific objectives, including any pre-specified hypotheses	6
Methods			
Study design	4	Present key elements of study design early in the paper	7-8
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	7-8
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	7-8
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	8
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	8
Bias	9	Describe any efforts to address potential sources of bias	7-8
Study size	10	Explain how the study size was arrived at	7-8
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	8
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	8-9
		(*b*) Describe any methods used to examine subgroups and interactions	8
		(*c*) Explain how missing data were addressed	7-8
		(*d*) If applicable, explain how loss to follow-up was addressed	7-8
		(*e*) Describe any sensitivity analyses	
Results			

Participants	13*	(a) Report numbers of individuals at each stage of study—e.g. numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	7-8
		(b) Give reasons for non-participation at each stage	7-8
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (e.g. demographic, clinical, social) and information on exposures and potential confounders	Table 1, 10-11
		(b) Indicate number of participants with missing data for each variable of interest	Table 1, 10-11
		(c) Summarise follow-up time (e.g., average and total amount)	7-8, Table 1, 10-11
Outcome data	15*	Report numbers of outcome events or summary measures over time	Table 1, 10-11
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Table 2-4, 12-13
		(*b*) Report category boundaries when continuous variables were categorized	7-8
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—e.g. analyses of subgroups and interactions, and sensitivity analyses	
Discussion			
Key results	18	Summarise key results with reference to study objectives	13
Limitations			
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	13-16
Generalisability	21	Discuss the generalisability (external validity) of the study results	14-16
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	21

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

Non-response in a nationwide follow-up postal survey in Finland: A register-based mortality analysis of respondents and non-respondents of the Health and Social Support (HeSSup) Study

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2011-000657.R1
Article Type:	Research
Date Submitted by the Author:	09-Jan-2012
Complete List of Authors:	Suominen, Sakari; University of Turku, Department of Public Health Koskenvuo, Karoliina; The Social Insurance Institution of Finland, Sillanmaki, Lauri; University of Helsinki, Hjelt Institute, Department of Public Health Vahtera, Jussi; University of Turku, Department of Public Health Korkeila, Katariina; City of Raisio, Kivimaki, Mika; London Medical School, Department of Epidemiology and Public Health, Mattila, Kari; University of Tampere, Center of General Practice Virtanen, Pekka; Tampere School of Public Health, University of Tampere Sumanen, Markku; University of Tampere, Rautava, Paivi; University of Turku, Department of Public Health Koskenvuo, Markku; University of Helsinki, Hjelt Institute, Department of Public Health
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Research methods
Keywords:	EPIDEMIOLOGY, PUBLIC HEALTH, SOCIAL MEDICINE

Non-response in a nationwide follow-up postal survey in Finland: A register-based mortality analysis of respondents and non-respondents of the Health and Social Support (HeSSup) Study

All authors critically revised the paper for important intellectual content and approved the final version.

Deleted: Non-respondence to postal survey is related to excess mortality: A register-based analysis of the Health and Social Support (HeSSup) Cohort Study, Finland ¶

Deleted: Sakari Suominen contributed substantially to acquisition of data, the design of the study and to the interpretation of the data, wrote the first draft of the ms and revised it several times critcally for important intellectual content. He has approved the final version of the ms.¶
Karoliina Koskenvuo contributed substantially to acquisition of data, the design of the study, carried out the first data analyses, contributed substantially to the interpretation of the data and revised the ms critically several times for important intellectual content. She has approved the final version of the ms.¶
Lauri Sillanmäki contributed substantially to acquisition of data, the design of the study, carried out data analyses, contributed substantially to the interpretation of the data and revised the ms critically several times for important intellectual content. He has approved the final version of the ms.¶
Jussi Vahtera contributed substantially to the design of the study and to the interpretation of the data and revised the ms critically several times for important intellectual content. He has approved the final version of the ms.¶
Katariina Korkeila contributed substantially to acquisition of data, the design of the study and to the interpretation of the data and revised the ms critically several times for important intellectual content. She has approved the final version of the ms.¶
Mika Kivimäki contributed substantially to design of the study and to the interpretation of the data and revised the ms critically several times for important intellectual content. He has approved the final version of the ms.¶
Kari J. Mattila contributed substantially to acquisition of data, the design of the study and to the interpretation of the data and revised the ms critically several times for important intellectual content. He has approved the final version of the ms.¶
Pekka J. Virtanen contributed substantially to the interpretation of the data and revised the ms critically several times for important intellectual content. He has approved the final version of the ms.¶
Markku Sumanen contributed substantially to the interpretation of the data and revised the ms critically several times for important intellectual content. He has approved the final version of the ms.¶
Päivi Rautava contributed substantially to acquisition of data, the interpretation of the data and revised the ms critically several times for important intellectual content. She has approved the final version of the ms.¶
Markku Koskenvuo contributed substantially to acquisition of data, to the design of the study and to the interpretation of the data and revised the ms critically several times for important intellectual content. He has approved the final version of the ms.¶

ABSTRACT

Objective

To examine difference in mortality between postal survey non-respondents and respondents.

Design

A prospective cohort study with baseline survey in 1998 and comprehensive linkage to national mortality registers until 2005, the Health and Social Support (HeSSup) study.

Setting

A population based postal survey of the working aged population in Finland in 1998

Participants

The original random sample comprised 64,797 working-aged individuals in Finland (20-24, 30-34, 40-44, 50-54 years of age; 32,059 women and 32,716 men), yielding 25,898 (40.0%) responses in the baseline postal survey in 1998.

Primary outcome measure

Registry based primary causes of death encoded with the International Classification of Diseases (ICD-10)

Results

In women, hazard ratio for total mortality was 1.75 (95% confidence interval 1.40 - 2.19) times higher among the non-respondents compared to the respondents. In men, non-response was associated with a 1.41 fold (1.21 - 1.65) excess risk of total mortality. Non-response associated in certain age groups with deaths due to diseases in women and with deaths due to external causes in men. The most prominent excess mortality was seen for total mortality for both genders and for mortality due to external causes among men.

Conclusions

Postal surveys result in slight underestimation of illness prevalence, especially illness involving mental health in men.

Deleted: non-

Deleted: can still be considered a valid method for population based health research, but they

Word Count of Abstract 207

Keywords

Follow-up Studies, Non-respondents, Mortality, Health Surveys, Registries

Article summary

Article focus

Women and individuals from upper social strata tend to participate more actively in postal health surveys

What this exactly means in terms of health selection among respondents is unclear

Postal health surveys are believed to produce underestimates of illness prevalence

Key message

Total mortality was consistently and for women in the age group >=50 years and for men in the age groups >=40 years significantly higher among non-respondents compared to respondents during a 7 year follow-up among a total Finnish nationwide sample in working age comprising almost 65 000 individuals

The excess mortality observed was 1,5 – 2 fold. Among men it was explained by external causes, whereas among women it was due to diseases and was statistically significant only in the age group 50 – 54 years

Postal surveys result in slight underestimation of illness prevalence, especially illness involving mental health in men

Strengths and limitations of the study

The linkage to mortality data was successful for virtually all individuals of the original sample comprising nearly 65 000 individuals. The sample size secures the reliability of the conclusions drawn. Furthermore, the registry data on mortality in Finland can be considered as reliable. To the best of the authors' knowledge a corresponding study based on a as large a sample as in this study has not previously been carried out.

Deleted: 20

Deleted: moderate

Deleted: only seen

Deleted: can still be considered a valid method for population based health research, but they

Some inaccuracy concerning the final diagnosis of death is possible. A further study limitation is that data of socioeconomic status or educational level of non-respondents was not available and hence adjustments of the statistical analyses for these variables was not possible.

Deleted: Inaccuracy of the final diagnosis of death is possible but nevertheless the data can be considered as reliable as registry data based on clinical work generally can become. Moreover, there is no support for the view that inaccuracy of diagnosis of death would systematically be associated with the response status of the study members.¶

INTRODUCTION

Word count without Abstract, Tables and References 1703

Response rates in health related postal surveys are declining in the Western world. In the Nordic countries the situation has been somewhat better compared to the rest of Europe and the U.S.A., but recently even there declining trends have been observed [1, 2]. Previous studies have suggested that women, older persons and persons from upper social strata tend to participate more actively in health surveys compared with the rest [3 - 11] However, not all studies have consistently supported these observations [10, 12, 13]. Furthermore, recent studies have extended analysis of non-response beyond demographic variables showing lower rate of hospital admissions [2], mortality and maternal smoking during pregnancy among the participants as compared to non-participants [14, 15].

Studies on causes for non-participation to health surveys have revealed incorrect address or incorrect delivery by post to contribute to some of the drop out [5, 8, 16]. The non-respondents themselves have reported various kinds of reasons for their behaviour, such as gaining no benefit for participation [5, 16], no interest in the topic [7, 8, 16, 17], feeling of intrusion of privacy [8], lack of time [7, 8, 17], forgetfulness [8], and present illness [17]. Surveys including questions on issues perceived as intimate have often decreased participation rates [18]. It has also been speculated that late respondents might resemble more the non-participants [5, 17, 19] but to date decisive evidence for this pattern is lacking [20]. In follow-

Deleted: 91

Deleted: a

Deleted: and

Deleted: responden

up studies, participation has been explored according to the number of rounds the individual has taken part in [14] and more occasions of non-response found to be positively associated with subsequent mortality rates. On the whole, health selection among participants might decrease the generalizability of prevalence estimates but this effect is until now not satisfactorily described [21]. However, even in the case of health selection the results related to the strength of association between the variables studied need not necessarily be biased [15, 21].

A non-response analysis of the postal survey of the population-based Health and Social Support (HeSSup) study - which achieved a relatively modest (40 %) response rate [9] - replicated previous findings on the differences in demographic characteristics between respondents and non-respondents. The aim of the present study was to extend these analyses to explore whether survey non-respondents differ from the respondents in terms of mortality (all-cause, disease mortality, mortality for external causes).

MATERIALS AND METHODS

The study is based on the complete sample (N=64, 797) of the Health and Social Support (HeSSup) prospective mail survey initiated in 1998. The concurrent joint Ethics Committee of the University of Turku and the Turku University Central Hospital considered approval not necessary for a normal cohort study, but all

participants were requested to sign a consent form containing information about the study and to grant permission to allow subsequent studies with the same data set and possibility to link with national health registries. The sample represented the concurrent age groups of 20-24, 30-34, 40-44 and 50-54 years of age in Finland [9]. However, by purpose the Swedish speaking Finns (5% of the general population) as well as the Turku region were slightly overrepresented. Of the 64,797 persons, 22 could not be included in the present study since the follow-up had to be set to begin from the first death among respondents which was September 22nd 1998. Certain cases of deaths among non-respondents had occurred already earlier and potentially before sending out the initial questionnaire. Totally 25,898 satisfactory responses were returned. In 2007, the survey material was by means of an unique social security number linked to Statistics Finland data on mortality between the years 1998 and 2005. Totally 1,174 cases of deaths among 25,290 observations that could be linked with registry data were identified. Moreover, mortality data of non-respondents from the same time period was likewise as for respondents obtained from Statistics Finland and further analyzed by age group and gender.

Deleted:

Deleted:

The outcome variable was primary causes of death encoded with the International Classification of Diseases (ICD-10). Mortality for external causes (S00 - Y98) and disease mortality (A00 – R99) were examined separately. The differences in mortality between non-respondents and respondents were analysed separately for women and men using Cox proportional hazards

regression. The analyses were carried out by the research group. Hazard ratios (HR) with 95% confidence intervals (95% CI) for mortality (total, external causes, disease) of non-respondents vs. respondents according to age (1998) were reported. Cox proportional hazard assumptions were tested by visual examination of log-minus-log plots showing parallelism among the selected strata variable (responding status). The statistical analysis was performed by the SAS ® software 9.2. for Windows.

RESULTS

Between the years 1998 to 2005, 1,174 individuals belonging to the complete sample died (Table 1). Of the deaths, 70% occurred in men.

Table 1. Demographics of respondents and non-respondents and number of deaths during follow-up from 1998 to 2005 (N and %).

	Respondents		Non-respondents		Total	
	N	%	N	%	N	%
Gender						
Women	14,922	59.0	17,137	43.4	32,059	49.5
Men	10,368	41.0	22,348	56.6	32,716	50.5
Total	25,290	100.0	39,485	100.0	64,775	100.0
Age in 1998						
20-24 y	6,783	26.8	9,405	23.8	16,188	25.0
30-34 y	5,981	23.7	10,267	26.0	16,248	25.1
40-44 y	6,073	24.0	10,198	25.8	16,271	25.1
50-54 y	6,453	25.5	9,615	24.4	16,068	24.8
Total	25,290	100.0	39,485	100.0	64,775	100.0
Number of deaths by gender	N	%	N	%	N	%
In women	110	0.7	240	1.4	350	1.1
In men	218	2.1	606	2.7	824	2.5
Total	328	1.3	846	2.1	1,174	1.8
Number of deaths by age in 1998						
20-24 y	21	0.3	47	0.5	68	0.4
30-34 y	38	0.6	105	1.0	143	0.8
40-44 y	85	1.4	252	2.5	337	2.1
50-54 y	184	2.9	442	4.6	626	3.9
Total	328	1.3	846	2.1	1,174	1.8
Number of deaths from external causes						
In women	26	0.2	48	0.3	74	0.2
In men	55	0.5	217	1.0	272	0.8
Total	81	0.3	265	0.7	346	0.5
Number of deaths from external causes by age in 1998						
20-24 y	10	<0.1	32	<0.1	42	<0.1
30-34 y	16	<0.1	66	0.2	82	0.1
40-44 y	22	<0.1	90	0.2	112	0.2
50-54 y	33	0.1	77	0.2	110	0.2
Total	81	0.3	265	0.7	346	0.5

Table 1 continued. Demographics of respondents and non-respondents and number of deaths during follow-up from 1998 to 2005 (N and %).

	Respondents		Non-respondents		Total	
	N	%	N	%	N	%
Number of deaths due to diseases						
In women	84	0.6	192	1.1	276	0.9
In men	163	1.6	389	1.7	552	1.7
Total	247	1.0	581	1.5	828	1.3
Number of deaths due to diseases by age in 1998						
20-24 y	11	<0.1	15	<0.1	26	<0.1
30-34 y	22	<0.1	39	<0.1	61	<0.1
40-44 y	63	0.2	162	0.4	225	0.3
50-54 y	151	0.6	365	0.9	516	0.8
Total	247	1.0	581	1.5	828	1.3

Total mortality was higher for non-respondent women in age group 50-54 years and for non-respondent men in age group 40-44 and 50-54 years and for each gender with all age groups combined. In analyses combining women and men, excess total mortality associated with non-response was observed in age groups 40-44 and 50-54 when age groups were examined separately as well as when all observations were combined (Table 2). Non-respondent men had a higher mortality for external causes in age groups 30-34 and 40-44 years and with all age groups combined. In analyses with genders combined, this was seen in age groups 30-34 and 40-44 years and when all observations were combined. In women, no statistically significant differences in mortality for external causes were detected (Table 3). Non-respondent women showed significantly higher disease-related mortality in age group 50-54 years as well as when all age groups were combined. The same held true for both genders when age groups 40-44 and 50-54 years were examined separately as well as when all

observations were combined. On the other hand, in separate analyses for men non-respondents showed a slightly increased disease mortality compared to respondents only when all age groups were combined (Table 4).

Table 2. Hazard ratios for total mortality of non-respondents vs. respondents and the 95% CIs according to gender and age at the beginning of the follow-up in 1998.

Non-respondents vs. respondents (=1.00)	Age in1998 20-24	Age in 1998 30-34	Age in 1998 40-44	Age in 1998 50-54	Total age or age & gender adjusted
Women	1.00 (0.34-2.97)	1.45 (0.77-2.73)	1.40 (0.93-2.10)	**2.17 (1.58-2.98)**	**1.75 (1.40-2.19)**
Men	1.28 (0.70-2.34)	1.41 (0.89-2.24)	**1.71 (1.25-2.35)**	**1.31 (1.07-1.61)**	**1.41 (1.21-1.65)**
Total gender or age & gender adjusted	1.21 (0.71-2.04)	1.42 (0.98-2.07)	**1.59 (1.24-2.04)**	**1.54 (1.29-1.83)**	**1.52 (1.34-1.73)**

Statistically significant associations are in bold.

Table 3. Hazard ratios for deaths due to external causes in non-respondents vs. respondents and the 95% CIs according to gender and age at the beginning of the follow-up in 1998.

Non-respondents vs. respondents (=1.00)	Age in 1998 20-24	Age in 1998 30-34	Age in 1998 40-44	Age in 1998 50-54	Total age or age & gender adjusted
Women	4.65 (0.52-41.62)	1.94 (0.68-5.49)	1.38 (0.58-3.28)	1.18 (0.57-2.45)	1.50 (0.93-2.42)
Men	1.36 (0.64-2.88)	**2.04 (1.07-3.90)**	**2.42 (1.37-4.28)**	1.62 (0.99-2.67)	**1.87 (1.39-2.52)**
Total gender or age & gender adjusted	1.61 (0.78-3.30)	**2.01 (1.16-3.49)**	**2.07 (1.30-3.31)**	1.47 (0.98-2.22)	**1.78 (1.39-2.29)**

Statistically significant associations are in bold.

Deleted: 50

Deleted: 3

Table 4. Hazard ratios for disease mortality of non-respondents vs. respondents and the 95% CIs according to gender and age at the beginning of the follow-up in 1998.

Non-respondents vs. respondents (=1.00)	Age in 1998 20-24	Age in 1998 30-34	Age in 1998 40-44	Age in 1998 50-54	Total age or age & gender adjusted
Women	0.39 (0.08-1.92)	1.21 (0.54-2.69)	1.41 (0.88-2.23)	**2.46 (1.73-3.52)**	**1.83 (1.41-2.36)**
Men	1.13 (0.40-3.18)	0.83 (0.42-1.66)	1.43 (0.98-2.09)	1.25 (1.00-1.57)	**1.25 (1.04-1.51)**
Total gender or age & gender adjusted	0.81 (0.37-1.80)	0.98 (0.58-1.67)	**1.42 (1.06-1.90)**	**1.55 (1.28-1.88)**	**1.43 (1.23-1.66)**

Statistically significant associations are in bold.

DISCUSSION

In this large population-based epidemiological study, comparison of mortality of non-respondents with respondents showed, as expected, consistently higher mortality rates among the former group. The differences between non-respondents and respondents were 1.5 – 2 fold and of quite similar magnitude for both genders although with varying causes. Among women the greatest statistically significant differences were seen in disease mortality in the oldest age group of initially 50–54 years of age. The greatest significant differences for men were caused by external causes of death and involved age groups 30- 34 and 40-44 years.

Deleted: 2
Formatted: English (U.S.)
Formatted: English (U.S.)
Deleted: 7
Deleted: 7
Deleted: 0.99
Deleted: 6
Formatted: English (U.S.)
Deleted: 0
Formatted: English (U.S.)
Formatted: English (U.S.)
Formatted: English (U.S.)
Formatted: English (U.S.)
Formatted: English (U.S.)
Deleted: 1
Formatted: English (U.S.)
Deleted: 2
Formatted: English (U.S.)
Deleted: 5
Formatted: English (U.S.)
Deleted: 89
Formatted: English (U.S.)
Deleted: 7
Formatted: English (U.S.)
Deleted: 5
Deleted: a
Deleted: non-
Deleted: , however,
Deleted: not very large

Postal surveys are frequently used in population based health research. The impact of health selection on the results has been studied but not very extensively. An affirmative view on the potential demographic difference between respondents and non-respondents is still lacking. According to previous findings [3, 8, 9], respondents tend to be somewhat healthier and report a more favourable health behaviour compared to non-respondents. However, the bias caused by this was limited and applied mainly to health problems or risk behaviours that generally are not eagerly communicated to others, such as mental problems or binge drinking. From previous studies it is also known that women as well as those well off in the society tend to participate more actively in health related survey research on the whole. Hence, this might result in under-estimation of the prevalence of health problems common among men as well as individuals from lower social strata. Results from a previous non-response analysis of the initial phase of this study supports this view [9]. Women as well as indivduals with high level of education were somewhat over represented among respondents but no clear health-related selection could be shown. According to our present results 7 years later the potential health-related selection can be further clarified. In Finland, mortality for external causes is intimately associated with alcohol consumption and alcoholism [22]. As could be expected, we could see a significantly higher mortality for external causes among male non-respondents as compared with the respondents.

Deleted:

Moreover, in previous studies [23] it has also been pointed out that non-respondents are not necessarily a homogenic group but can differ internally, e.g. depending on to which wave of the survey if any they have taken part in. Also the correspondence between late respondents and total non-respondents has been questioned [20].

Given the health selection related to postal surveys it has to be kept in mind that population studies usually do not focus solely on prevalence estimates anymore but more or less on potential risk or protective factors of certain health problems. From previous research [21] there are indications that even if prevalence estimates are not accurate the associations between the variables studies are not necessarily biased.

Deleted: This gives further justificaton to continue postal surveys in health research since the results can help us understand the etiology and prevention process of a certain kind of disease. ¶

Strengths and limitations of the study

A major strength of this study is that the linkage to mortality data was successful for virtually all individuals of the original sample. Furthermore, the registry data on mortality in Finland could be considered as quite reliable. All deaths with suspicion of an external cause and in the age groups studied are investigated by autopsy.

The large sample size secures that the conclusions drawn from the statistical analyses are reliable and can not have been caused by random effects. To the best of our knowledge a corresponding study based on as large a sample as in this study has previously not been carried out.

Deleted: Naturally, inaccuracy concerning the final diagnosis might be present but nevertheless the data can be considered as reliable as registry data based on clinical medical work generally can become. Moreover, there is no reason to assume that problems in reliability of the diagnoses would systematically be associated with the response status of the study members. The legislation in Finland enables use of mortality data for research purposes and thus we were able by ourselves to perform a mortality analysis for all respondents.

Deleted: ¶

Some inaccuracy concerning the final diagnosis of death is possible. Another study limitation is that data of socioeconomic status or educational level of non-respondents was not available and hence adjustments of the statistical analyses for these variables was not possible.

Conclusions

Total mortality was consistently 1.5 – 2 fold and for women in the age group >= 50 years and for men in the age groups >=40 years significantly higher for the non-respondents of a nationwide mail survey compared to the respondents. For women this was mostly due to disease mortality in age group 50-54 years but for men due to mortality for external causes in age groups 30-34 and 40-44 years. The most prominent excess mortality was seen for total mortality for both genders and for mortality due to external causes among men. Selection by health, especially mental health in men can cause bias in health related population surveys. However, this applies to prevalence estimates and does not necessarily jeopardize results from studies on risk and protective factors.

Deleted: moderately higher

Deleted:

Deleted: in both genders

Deleted: Hence, it seems justified to continue postal surveys in health research in order to improve understanding of etiology and prevention of diseases.

REFERENCES

1. Helakorpi S., Prättälä R., Uutela A. Suomalaisen aikuisväestön terveyskäyttäytyminen ja terveys, kevät 2007. [Health Behaviour and Health

among the Finnish Adult Population. Spring 2007] Publications of the National Public Health Institute. 2008;B 6.

2. Kjøller M, Thoning H. Characteristics of non-response in the Danish Health Interview Surveys, 1987-1994. *Eur J Public Health* 2005;**15**(5):528-35 doi: 10.1093/eurpub/cki023 [Advance Access published online July 28, 2005].

3. Cottler LB, Zipp JF, Robins LN, et al. Difficult-to-recruit respondents and their effect on prevalence estimates in an epidemiologic survey. *Am J Epidemiol* 1987;**125**(2):329-39.

4. Jacobsen BK, Thelle DS. The Tromsø Heart Study: responders and non-responders to a health questionnaire, do they differ? *Scand J Soc Med* 1988;**16**(2):101-4.

5. Bakke P, Gulsvik A, Lilleng P, et al. Postal survey on airborne occupational exposure and respiratory disorders in Norway: causes and consequences of non-response. *J Epidemiol Community Health* 1990;**44**(4):316-20.

6. Jooste PL, Yach D, Steenkamp HJ, et al. Drop-out and newcomer bias in a community cardiovascular follow-up study. *Int J Epidemiol* 1990;19(2):284-9.

7. Brussaard JH, Brants HA, Bouman M, et al. The study population: general characteristics and potential confounding factors. *Eur J Clin Nutr* 1997;**51** Suppl 3:S19-24.

8. Rönmark E, Lundqvist A, Lundbäck B, et al. Non-responders to a postal questionnaire on respiratory symptoms and diseases. *Eur J Epidemiol* 1999;**15**(3):293-9.

9. Korkeila K, Suominen S, Ahvenainen J, et al. Non-response and related factors in a nation-wide health survey. *Eur J Epidemiol*. 2001;**17**(11):991-9.

10. van den Akker M, Buntinx F, Metsemakers JF, et al. Morbidity in responders and non-responders in a register-based population survey. *Fam Pract* 1998;**15**(3):261-3.

11. Sonne-Holm S, Sørensen TI, Jensen G, et al. Influence of fatness, intelligence, education and sociodemographic factors on response rate in a health survey. *J Epidemiol Community Health* 1989;**43**(4):369-74.

12. Asch DA, Jedrziewski MK, Christakis NA. Response rates to mail surveys published in medical journals. *J Clin Epidemiol* 1997;**50**(10):1129-36.

13. Heilbrun LK, Ross PD, Wasnich RD, et al. Characteristics of respondents and nonrespondents in a prospective study of osteoporosis. *J Clin Epidemiol* 1991;**44**(3):233-9.

14. Ferrie JE, Kivimäki M, Singh-Manoux A, et al. Non-response to baseline, non-response to follow-up and mortality in the Whitehall II cohort. *Int J Epidemiol* 2009;**38**(3):831-7 doi:10.1093/ije/dyp153 [Advance Access publication 5 March 2009].

15. Greene N, Greenland S, Olsen J, Aagard Nohr E. Estimating Bias From Loss to Follow-up in the Danish National Birth Cohort. *Epidemiology* 2011;**22**:815–822.

16. Foster K, Campbell D, Crum J, et al. Non-response in a population study after an environmental disaster. *Public Health* 1995;**109**(4):267-73.

17. Janzon L, Hanson BS, Isacsson SO, et al. Factors influencing participation in health surveys. Results from prospective population study 'Men born in 1914' in Malmö, Sweden. *J Epidemiol Community Health* 1986;**40**(2):174-7.

18. Dunne MP, Martin NG, Bailey JM, et al. Participation bias in a sexuality survey: psychological and behavioural characteristics of responders and non-responders. *Int J Epidemiol* 1997;**26**(4):844-54.

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**(1):81-5.

20. Lahaut, V. M. H. C. J., Jansen HAM, van de Mheen D, et al. Estimating non-response bias in a survey on alcohol consumption: comparison of response waves. *Alcohol Alcohol* 2003;**38**(2):128-34 [doi:10.1093/alcalc/agg044].

21. Martikainen P, Laaksonen M, Piha K, et al. Does survey non-response bias the association between occupational social class and health? *Scand J Public Health* 2007;**35**(2):212-5 [doi:10.1080/14034940600996563].

Formatted: Font: 12 pt

Deleted: Mattila VM, Parkkari J, Rimpelä A. Adolescent survey non-response and later risk of death. A prospective cohort study of 78 609 persons with 11-year follow-up. *BMC Public Health* 2007;**7**(1):87 doi:10.1186/1471-2458-7-87 [Published: 22 May 2007].

Formatted: Font: 12 pt

Formatted: Normal, Line spacing: 1.5 lines, Don't adjust space between Latin and Asian text, Don't adjust space between Asian text and numbers

Formatted: Font: (Default) Arial, 12 pt, English (U.S.), Check spelling and grammar

Formatted: Font: 12 pt, Check spelling and grammar

Formatted: Font: (Default) Arial, 12 pt, English (U.S.), Check spelling and grammar

Formatted: Font: (Default) Arial, 12 pt, Check spelling and grammar

Formatted: Font: (Default) Arial, 12 pt, Bold, Check spelling and grammar

Formatted: Font: (Default) Arial, 12 pt, Check spelling and grammar

Formatted: Font: 12 pt, English (U.S.), Check spelling and grammar

22. Koskenvuo M, Kaprio J, Kesaniemi A, et al. Alcohol-related diseases associated with ischaemic heart disease: a three-year follow-up of middle-aged male hospital patients. *Alcohol Alcohol* 1986;**21**(3):251-6.

23. Boersma F, Eefsting JA, van den Brink W, et al. Characteristics of non-responders and the impact of non-response on prevalence estimates of dementia. *Int J Epidemiol* 1997;**26**(5):1055-62.

Acknowledgements

The study was supported by the Academy of Finland (three different decisions on funding)

The authors do not declare any competing interests

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *cohort studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	4
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	6-7
Objectives	3	State specific objectives, including any pre-specified hypotheses	7
Methods			
Study design	4	Present key elements of study design early in the paper	8-9
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	8-9
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	8-9
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	8-9
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	8-9
Bias	9	Describe any efforts to address potential sources of bias	8-9
Study size	10	Explain how the study size was arrived at	9
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	9-10
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	9-10
		(*b*) Describe any methods used to examine subgroups and interactions	9-10
		(*c*) Explain how missing data were addressed	9
		(*d*) If applicable, explain how loss to follow-up was addressed	8-9
		(*e*) Describe any sensitivity analyses	
Results			

Participants	13*	(a) Report numbers of individuals at each stage of study—e.g. numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	8-9
		(b) Give reasons for non-participation at each stage	8-9
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (e.g. demographic, clinical, social) and information on exposures and potential confounders	Table 1, 11-12
		(b) Indicate number of participants with missing data for each variable of interest	Table 1, 11-12
		(c) Summarise follow-up time (e.g., average and total amount)	8-9, Table 1, 11-12
Outcome data	15*	Report numbers of outcome events or summary measures over time	Table 1, 11-12
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Table 2-4, 13-14
		(*b*) Report category boundaries when continuous variables were categorized	8-9
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—e.g. analyses of subgroups and interactions, and sensitivity analyses	
Discussion			
Key results	18	Summarise key results with reference to study objectives	14
Limitations			
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	15-18
Generalisability	21	Discuss the generalisability (external validity) of the study results	15-18
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	22

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ open

Non-response in a nationwide follow-up postal survey in Finland: A register-based mortality analysis of respondents and non-respondents of the Health and Social Support (HeSSup) Study

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2011-000657.R2
Article Type:	Research
Date Submitted by the Author:	27-Jan-2012
Complete List of Authors:	Suominen, Sakari; University of Turku, Department of Public Health Koskenvuo, Karoliina; The Social Insurance Institution of Finland, Sillanmaki, Lauri; University of Helsinki, Hjelt Institute, Department of Public Health Vahtera, Jussi; University of Turku, Department of Public Health Korkeila, Katariina; City of Raisio, Kivimaki, Mika; University College London, Department of Epidemiology and Public Health Mattila, Kari; Hospital District of Pirkanmaa, Centre of General Practice Virtanen, Pekka; Tampere School of Public Health, University of Tampere Sumanen, Markku; University of Tampere, Rautava, Paivi; University of Turku, Department of Public Health Koskenvuo, Markku; University of Helsinki, Hjelt Institute, Department of Public Health
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Research methods, Occupational and environmental medicine
Keywords:	EPIDEMIOLOGY, PUBLIC HEALTH, SOCIAL MEDICINE

Non-response in a nationwide follow-up postal survey in Finland: A register-based mortality analysis of respondents and non-respondents of the Health and Social Support (HeSSup) Study

Sakari Suominen contributed substantially to acquisition of data, the design of the study and to the interpretation of the data, wrote the first draft of the ms and revised it several times critcally for important intellectual content.

Karoliina Koskenvuo contributed substantially to acquisition of data, the design of the study, carried out the first data analyses, contributed substantially to the interpretation of the data and revised the ms critically several times for important intellectual content.

Lauri Sillanmäki contributed substantially to acquisition of data, the design of the study, carried out data analyses, contributed substantially to the interpretation of the data and revised the ms critically several times for important intellectual content.

Jussi Vahtera and Mika Kivimäki contributed substantially to the design of the study and to the interpretation of the data and revised the ms critically several times for important intellectual content.

Katariina Korkeila,Kari J. Mattila and Markku Koskenvuo contributed substantially to acquisition of data, the design of the study and to the interpretation of the data and revised the ms critically several times for important intellectual content.

Pekka J. Virtanen and Markku Sumanen contributed substantially to the interpretation of the data and revised the ms critically several times for important intellectual content.

Päivi Rautava contributed substantially to acquisition of data, the interpretation of the data and revised the ms critically several times for important intellectual content.

All the authors have read and approved the final version of the ms.

ABSTRACT

Objective

To examine difference in mortality between postal survey non-respondents and respondents.

Design

A prospective cohort study with baseline survey in 1998 and comprehensive linkage to national mortality registers until 2005, the Health and Social Support (HeSSup) study.

Setting

A population based postal survey of the working aged population in Finland in 1998

Participants

The original random sample comprised 64,797 working-aged individuals in Finland (20-24, 30-34, 40-44, 50-54 years of age; 32,059 women and 32,716 men), yielding 25,898 (40.0%) responses in the baseline postal survey in 1998.

Primary outcome measure

Registry based primary causes of death encoded with the International Classification of Diseases (ICD-10)

Results

In women, hazard ratio for total mortality was 1.75 (95% confidence interval 1.40 - 2.19) times higher among the non-respondents compared to the respondents. In men, non-response was associated with a 1.41 fold (1.21 - 1.65) excess risk of total mortality. Non-response associated in certain age groups with deaths due to diseases in women and with deaths due to external causes in men. The most prominent excess mortality was seen for total mortality for both genders and for mortality due to external causes among men.

Conclusions

Postal surveys result in slight underestimation of illness prevalence.

Word Count of Abstract 201

Keywords

Follow-up Studies, Non-respondents, Mortality, Health Surveys, Registries

Article summary

Article focus

Women and individuals from upper social strata tend to participate more actively in postal health surveys

What this exactly means in terms of health selection among respondents is unclear

Postal health surveys are believed to produce underestimates of illness prevalence

Key message

Total mortality was consistently and for women in the age group >=50 years and for men in the age groups >=40 years significantly higher among non-respondents compared to respondents during a 7 year follow-up among a total Finnish nationwide sample in working age comprising almost 65 000 individuals

The excess mortality observed was 1,5 – 2 fold. Among men it was explained by external causes, whereas among women it was due to diseases and was statistically significant only in the age group 50 – 54 years

Postal surveys result in slight underestimation of illness prevalence

Strengths and limitations of the study

The linkage to mortality data was successful for virtually all individuals of the original sample comprising nearly 65 000 individuals. The sample size secures the reliability of the conclusions drawn. Furthermore, the registry data on mortality in Finland can be considered as reliable. To the best of the authors' knowledge a corresponding study based on a as large a sample as in this study has not previously been carried out.

Some inaccuracy concerning the final diagnosis of death is possible. A further study limitation is that data of socioeconomic status or educational level of non-

respondents was not available and hence adjustments of the statistical analyses for these variables was not possible.

INTRODUCTION

Word count without Abstract, Tables and References 1735

Response rates in health related postal surveys are declining in the Western world. In the Nordic countries the situation has been somewhat better compared to the rest of Europe and the U.S.A., but recently even there declining trends have been observed [1, 2]. Previous studies have suggested that women, older persons and persons from upper social strata tend to participate more actively in health surveys compared with the rest [3 - 11] However, not all studies have consistently supported these observations [10, 12, 13]. Furthermore, recent studies have extended analysis of non-response beyond demographic variables showing lower rate of hospital admissions [2], mortality and maternal smoking during pregnancy among the participants as compared to non-participants [14, 15].

Studies on causes for non-participation to health surveys have revealed incorrect address or incorrect delivery by post to contribute to some of the drop out [5, 8, 16]. The non-respondents themselves have reported various kinds of reasons for their behaviour, such as gaining no benefit for participation [5, 16], no interest in the topic [7, 8, 16, 17], feeling of intrusion of privacy [8], lack of time [7, 8, 17], forgetfulness [8], and present illness [17]. Surveys including questions on issues perceived as intimate have often decreased participation rates [18]. It has also been speculated that late respondents might resemble more the non-participants [5, 17, 19] but to date decisive evidence for this pattern is lacking [20]. In follow-

up studies, participation has been explored according to the number of rounds the individual has taken part in [14] and more occasions of non-response found to be positively associated with subsequent mortality rates. On the whole, health selection among participants might decrease the generalizability of prevalence estimates but this effect is until now not satisfactorily described [21]. However, even in the case of health selection the results related to the strength of association between the variables studied need not necessarily be biased [15, 21].

A non-response analysis of the postal survey of the population-based Health and Social Support (HeSSup) study - which achieved a relatively modest (40 %) response rate [9] - replicated previous findings on the differences in demographic characteristics between respondents and non-respondents. The aim of the present study was to extend these analyses to explore whether survey non-respondents differ from the respondents in terms of mortality (all-cause, disease mortality, mortality for external causes).

MATERIALS AND METHODS

The study is based on the complete sample (N=64, 797) of the Health and Social Support (HeSSup) prospective mail survey initiated in 1998. The concurrent joint Ethics Committee of the University of Turku and the Turku University Central Hospital considered approval not necessary for a normal cohort study, but all

participants were requested to sign a consent form containing information about the study and to grant permission to allow subsequent studies with the same data set and possibility to link with national health registries. The sample represented the concurrent age groups of 20-24, 30-34, 40-44 and 50-54 years of age in Finland [9]. However, by purpose the Swedish speaking Finns (5% of the general population) as well as the Turku region were slightly overrepresented. Of the 64,797 persons, 22 could not be included in the present study since the follow-up had to be set to begin from the first death among respondents which was September 22nd 1998. Certain cases of deaths among non-respondents had occurred already earlier and potentially before sending out the initial questionnaire. Totally 25,898 satisfactory responses were returned. In 2007, the survey material was by means of an unique social security number linked to Statistics Finland data on mortality between the years 1998 and 2005. Totally 1,174 cases of deaths among 25,290 observations that could be linked with registry data were identified. Moreover, mortality data of non-respondents from the same time period was likewise as for respondents obtained from Statistics Finland and further analyzed by age group and gender.

The outcome variable was primary causes of death encoded with the International Classification of Diseases (ICD-10). Mortality for external causes (S00 - Y98) and disease mortality (A00 – R99) were examined separately. The differences in mortality between non-respondents and respondents were analysed separately for women and men using Cox proportional hazards

regression. The analyses were carried out by the research group. Hazard ratios (HR) with 95% confidence intervals (95% CI) for mortality (total, external causes, disease) of non-respondents vs. respondents according to age (1998) were reported. Cox proportional hazard assumptions were tested by visual examination of log-minus-log plots showing parallelism among the selected strata variable (responding status). The statistical analysis was performed by the SAS ® software 9.2. for Windows.

RESULTS

Between the years 1998 to 2005, 1,174 individuals belonging to the complete sample died (Table 1). Of the deaths, 70% occurred in men.

Table 1. Demographics of respondents and non-respondents and number of deaths during follow-up from 1998 to 2005 (N and %).

	Respondents		Non-respondents		Total	
	N	%	N	%	N	%
Gender						
Women	14,922	59.0	17,137	43.4	32,059	49.5
Men	10,368	41.0	22,348	56.6	32,716	50.5
Total	25,290	100.0	39,485	100.0	64,775	100.0
Age in 1998						
20-24 y	6,783	26.8	9,405	23.8	16,188	25.0
30-34 y	5,981	23.7	10,267	26.0	16,248	25.1
40-44 y	6,073	24.0	10,198	25.8	16,271	25.1
50-54 y	6,453	25.5	9,615	24.4	16,068	24.8
Total	25,290	100.0	39,485	100.0	64,775	100.0
Number of deaths by gender	N	%	N	%	N	%
In women	110	0.7	240	1.4	350	1.1
In men	218	2.1	606	2.7	824	2.5
Total	328	1.3	846	2.1	1,174	1.8
Number of deaths by age in 1998						
20-24 y	21	0.3	47	0.5	68	0.4
30-34 y	38	0.6	105	1.0	143	0.8
40-44 y	85	1.4	252	2.5	337	2.1
50-54 y	184	2.9	442	4.6	626	3.9
Total	328	1.3	846	2.1	1,174	1.8
Number of deaths from external causes						
In women	26	0.2	48	0.3	74	0.2
In men	55	0.5	217	1.0	272	0.8
Total	81	0.3	265	0.7	346	0.5
Number of deaths from external causes by age in 1998						
20-24 y	10	<0.1	32	<0.1	42	<0.1
30-34 y	16	<0.1	66	0.2	82	0.1
40-44 y	22	<0.1	90	0.2	112	0.2
50-54 y	33	0.1	77	0.2	110	0.2
Total	81	0.3	265	0.7	346	0.5

Table 1 continued. Demographics of respondents and non-respondents and number of deaths during follow-up from 1998 to 2005 (N and %).

	Respondents		Non-respondents		Total	
	N	%	N	%	N	%
Number of deaths due to diseases						
In women	84	0.6	192	1.1	276	0.9
In men	163	1.6	389	1.7	552	1.7
Total	247	1.0	581	1.5	828	1.3
Number of deaths due to diseases by age in 1998						
20-24 y	11	<0.1	15	<0.1	26	<0.1
30-34 y	22	<0.1	39	<0.1	61	<0.1
40-44 y	63	0.2	162	0.4	225	0.3
50-54 y	151	0.6	365	0.9	516	0.8
Total	247	1.0	581	1.5	828	1.3

Total mortality was higher for non-respondent women in age group 50-54 years and for non-respondent men in age group 40-44 and 50-54 years and for each gender with all age groups combined. In analyses combining women and men, excess total mortality associated with non-response was observed in age groups 40-44 and 50-54 when age groups were examined separately as well as when all observations were combined (Table 2). Non-respondent men had a higher mortality for external causes in age groups 30-34 and 40-44 years and with all age groups combined. In analyses with genders combined, this was seen in age groups 30-34 and 40-44 years and when all observations were combined. In women, no statistically significant differences in mortality for external causes were detected (Table 3). Non-respondent women showed significantly higher disease-related mortality in age group 50-54 years as well as when all age groups were combined. The same held true for both genders when age groups 40-44 and 50-54 years were examined separately as well as when all

observations were combined. On the other hand, in separate analyses for men non-respondents showed a slightly increased disease mortality compared to respondents only when all age groups were combined (Table 4).

Table 2. Hazard ratios for total mortality of non-respondents vs. respondents and the 95% CIs according to gender and age at the beginning of the follow-up in 1998.

Non-respondents vs. respondents (=1.00)	Age in1998 20-24	Age in 1998 30-34	Age in 1998 40-44	Age in 1998 50-54	Total age or age & gender adjusted
Women	1.00 (0.34-2.97)	1.45 (0.77-2.73)	1.40 (0.93-2.10)	**2.17 (1.58-2.98)**	**1.75 (1.40-2.19)**
Men	1.28 (0.70-2.34)	1.41 (0.89-2.24)	**1.71 (1.25-2.35)**	**1.31 (1.07-1.61)**	**1.41 (1.21-1.65)**
Total gender or age & gender adjusted	1.21 (0.71-2.04)	1.42 (0.98-2.07)	**1.59 (1.24-2.04)**	**1.54 (1.29-1.83)**	**1.52 (1.34-1.73)**

Statistically significant associations are in bold.

Table 3. Hazard ratios for deaths due to external causes in non-respondents vs. respondents and the 95% CIs according to gender and age at the beginning of the follow-up in 1998.

Non-respondents vs. respondents (=1.00)	Age in 1998 20-24	Age in 1998 30-34	Age in 1998 40-44	Age in 1998 50-54	Total age or age & gender adjusted
Women	4.65 (0.52-41.62)	1.94 (0.68-5.49)	1.38 (0.58-3.28)	1.18 (0.57-2.45)	1.50 (0.93-2.42)
Men	1.36 (0.64-2.88)	**2.04 (1.07-3.90)**	**2.42 (1.37-4.28)**	1.62 (0.99-2.67)	**1.87 (1.39-2.52)**
Total gender or age & gender adjusted	1.61 (0.78-3.30)	**2.01 (1.16-3.49)**	**2.07 (1.30-3.31)**	1.47 (0.98-2.22)	**1.78 (1.39-2.29)**

Statistically significant associations are in bold.

Table 4. Hazard ratios for disease mortality of non-respondents vs. respondents and the 95% CIs according to gender and age at the beginning of the follow-up in 1998.

Non-respondents vs. respondents (=1.00)	Age in 1998 20-24	Age in 1998 30-34	Age in 1998 40-44	Age in 1998 50-54	Total age or age & gender adjusted
Women	0.39 (0.08-1.92)	1.21 (0.54-2.69)	1.41 (0.88-2.23)	**2.46 (1.73-3.52)**	**1.83 (1.41-2.36)**
Men	1.13 (0.40-3.18)	0.83 (0.42-1.66)	1.43 (0.98-2.09)	1.25 (1.00-1.57)	**1.25 (1.04-1.51)**
Total gender or age & gender adjusted	0.81 (0.37-1.80)	0.98 (0.58-1.67)	**1.42 (1.06-1.90)**	**1.55 (1.28-1.88)**	**1.43 (1.23-1.66)**

Statistically significant associations are in bold.

DISCUSSION

In this large population-based epidemiological study, comparison of mortality of non-respondents with respondents showed, as expected, consistently higher mortality rates among the former group. The differences between non-respondents and respondents were 1.5 – 2 fold and of quite similar magnitude for both genders although with varying causes. Among women the greatest statistically significant differences were seen in disease mortality in the oldest age group of initially 50–54 years of age. The greatest significant differences for men were caused by external causes of death and involved age groups 30- 34 and 40-44 years. However, it is worth to notice that an increase of total mortality of the magnitude seen in this study implies approximately 300 extra deaths among non-respondents during the follow-up of 7 years.

Postal surveys are frequently used in population based health research. The impact of health selection on the results has been studied but not very extensively. An affirmative view on the potential demographic difference between respondents and non-respondents is still lacking. According to previous findings [3, 8, 9], respondents tend to be somewhat healthier and report a more favourable health behaviour compared to non-respondents. However, the bias caused by this was limited and applied mainly to health problems or risk behaviours that generally are not eagerly communicated to others, such as mental problems or binge drinking. From previous studies it is also known that women as well as those well off in the society tend to participate more actively in health related survey research on the whole. Hence, this might result in under-estimation of the prevalence of health problems common among men as well as individuals from lower social strata. Results from a previous non-response analysis of the initial phase of this study supports this view [9]. Women as well as indivduals with high level of education were somewhat over represented among respondents but no clear health-related selection could be shown. According to our present results 7 years later the potential health-related selection can be further clarified. In Finland, mortality for external causes is intimately associated with alcohol consumption and alcoholism [22]. As could be expected, we could see a significantly higher mortality for external causes among male non-respondents as compared with the respondents.

Moreover, in previous studies [23] it has also been pointed out that non-respondents are not necessarily a homogenic group but can differ internally, e.g. depending on to which wave of the survey if any they have taken part in. Also the correspondence between late respondents and total non-respondents has been questioned [20].

Given the health selection related to postal surveys it has to be kept in mind that population studies usually do not focus solely on prevalence estimates anymore but more or less on potential risk or protective factors of certain health problems. From previous research [21] there are indications that even if prevalence estimates are not accurate the associations between the variables studies are not necessarily biased.

Strengths and limitations of the study

A major strength of this study is that the linkage to mortality data was successful for virtually all individuals of the original sample. Furthermore, the registry data on mortality in Finland could be considered as quite reliable. All deaths with suspicion of an external cause and in the age groups studied are investigated by autopsy.

The large sample size secures that the conclusions drawn from the statistical analyses are reliable and can not have been caused by random effects. To the best of our knowledge a corresponding study based on as large a sample as in this study has previously not been carried out.

Some inaccuracy concerning the final diagnosis of death is possible. Another study limitation is that data of socioeconomic status or educational level of non-respondents was not available and hence adjustments of the statistical analyses for these variables was not possible.

Conclusions

Total mortality was consistently 1.5 – 2 fold and for women in the age group >= 50 years and for men in the age groups >=40 years significantly higher for the non-respondents of a nationwide mail survey compared to the respondents. For women this was mostly due to disease mortality in age group 50-54 years but for men due to mortality for external causes in age groups 30-34 and 40-44 years. The most prominent excess mortality was seen for total mortality for both genders and for mortality due to external causes among men. Selection by health, especially mental health in men can cause bias in health related population surveys. However, this applies to prevalence estimates and does not necessarily jeopardize results from studies on risk and protective factors.

REFERENCES

1. Helakorpi S., Prättälä R., Uutela A. Suomalaisen aikuisväestön terveyskäyttäytyminen ja terveys, kevät 2007. [Health Behaviour and Health

among the Finnish Adult Population. Spring 2007] Publications of the National Public Health Institute. 2008;B 6.

2. Kjøller M, Thoning H. Characteristics of non-response in the Danish Health Interview Surveys, 1987-1994. *Eur J Public Health* 2005;**15**(5):528-35 doi: 10.1093/eurpub/cki023 [Advance Access published online July 28, 2005].

3. Cottler LB, Zipp JF, Robins LN, et al. Difficult-to-recruit respondents and their effect on prevalence estimates in an epidemiologic survey. *Am J Epidemiol* 1987;**125**(2):329-39.

4. Jacobsen BK, Thelle DS. The Tromsø Heart Study: responders and non-responders to a health questionnaire, do they differ? *Scand J Soc Med* 1988;**16**(2):101-4.

5. Bakke P, Gulsvik A, Lilleng P, et al. Postal survey on airborne occupational exposure and respiratory disorders in Norway: causes and consequences of non-response. *J Epidemiol Community Health* 1990;**44**(4):316-20.

6. Jooste PL, Yach D, Steenkamp HJ, et al. Drop-out and newcomer bias in a community cardiovascular follow-up study. *Int J Epidemiol* 1990;19(2):284-9.

7. Brussaard JH, Brants HA, Bouman M, et al. The study population: general characteristics and potential confounding factors. *Eur J Clin Nutr* 1997;**51** Suppl 3:S19-24.

8. Rönmark E, Lundqvist A, Lundbäck B, et al. Non-responders to a postal questionnaire on respiratory symptoms and diseases. *Eur J Epidemiol* 1999;**15**(3):293-9.

9. Korkeila K, Suominen S, Ahvenainen J, et al. Non-response and related factors in a nation-wide health survey. *Eur J Epidemiol*. 2001;**17**(11):991-9.

10. van den Akker M, Buntinx F, Metsemakers JF, et al. Morbidity in responders and non-responders in a register-based population survey. *Fam Pract* 1998;**15**(3):261-3.

11. Sonne-Holm S, Sørensen TI, Jensen G, et al. Influence of fatness, intelligence, education and sociodemographic factors on response rate in a health survey. *J Epidemiol Community Health* 1989;**43**(4):369-74.

12. Asch DA, Jedrziewski MK, Christakis NA. Response rates to mail surveys published in medical journals. *J Clin Epidemiol* 1997;**50**(10):1129-36.

13. Heilbrun LK, Ross PD, Wasnich RD, et al. Characteristics of respondents and nonrespondents in a prospective study of osteoporosis. *J Clin Epidemiol* 1991;**44**(3):233-9.

14. Ferrie JE, Kivimäki M, Singh-Manoux A, et al. Non-response to baseline, non-response to follow-up and mortality in the Whitehall II cohort. *Int J Epidemiol* 2009;**38**(3):831-7 doi:10.1093/ije/dyp153 [Advance Access publication 5 March 2009].

15. Greene N, Greenland S, Olsen J, Aagard Nohr E. Estimating Bias From Loss to Follow-up in the Danish National Birth Cohort. *Epidemiology* 2011;**22**:815–822.

16. Foster K, Campbell D, Crum J, et al. Non-response in a population study after an environmental disaster. *Public Health* 1995;**109**(4):267-73.

17. Janzon L, Hanson BS, Isacsson SO, et al. Factors influencing participation in health surveys. Results from prospective population study 'Men born in 1914' in Malmö, Sweden. *J Epidemiol Community Health* 1986;**40**(2):174-7.

18. Dunne MP, Martin NG, Bailey JM, et al. Participation bias in a sexuality survey: psychological and behavioural characteristics of responders and non-responders. *Int J Epidemiol* 1997;**26**(4):844-54.

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**(1):81-5.

20. Lahaut, V. M. H. C. J., Jansen HAM, van de Mheen D, et al. Estimating non-response bias in a survey on alcohol consumption: comparison of response waves. *Alcohol Alcohol* 2003;**38**(2):128-34 [doi:10.1093/alcalc/agg044].

21. Martikainen P, Laaksonen M, Piha K, et al. Does survey non-response bias the association between occupational social class and health? *Scand J Public Health* 2007;**35**(2):212-5 [doi:10.1080/14034940600996563].

22. Koskenvuo M, Kaprio J, Kesaniemi A, et al. Alcohol-related diseases associated with ischaemic heart disease: a three-year follow-up of middle-aged male hospital patients. *Alcohol Alcohol* 1986;**21**(3):251-6.

23. Boersma F, Eefsting JA, van den Brink W, et al. Characteristics of non-responders and the impact of non-response on prevalence estimates of dementia. *Int J Epidemiol* 1997;**26**(5):1055-62.

Acknowledgements

The study was supported by the Academy of Finland (three different decisions on funding)

The authors do not declare any competing interests

Data Sharing

At this point we are not willing to share data based on a totally open principle. However, naturally we are willing to collaborate with other researchers based on plans approved by our research group.

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *cohort studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	4
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	6-7
Objectives	3	State specific objectives, including any pre-specified hypotheses	7
Methods			
Study design	4	Present key elements of study design early in the paper	8-9
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	8-9
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	8-9
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	8-9
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	8-9
Bias	9	Describe any efforts to address potential sources of bias	8-9
Study size	10	Explain how the study size was arrived at	9
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	9-10
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	9-10
		(*b*) Describe any methods used to examine subgroups and interactions	9-10
		(*c*) Explain how missing data were addressed	9
		(*d*) If applicable, explain how loss to follow-up was addressed	8-9
		(*e*) Describe any sensitivity analyses	
Results			

Participants	13*	(a) Report numbers of individuals at each stage of study—e.g. numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	8-9
		(b) Give reasons for non-participation at each stage	8-9
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (e.g. demographic, clinical, social) and information on exposures and potential confounders	Table 1, 11-12
		(b) Indicate number of participants with missing data for each variable of interest	Table 1, 11-12
		(c) Summarise follow-up time (e.g., average and total amount)	8-9, Table 1, 11-12
Outcome data	15*	Report numbers of outcome events or summary measures over time	Table 1, 11-12
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Table 2-4, 13-14
		(*b*) Report category boundaries when continuous variables were categorized	8-9
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—e.g. analyses of subgroups and interactions, and sensitivity analyses	
Discussion			
Key results	18	Summarise key results with reference to study objectives	14
Limitations			
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	15-18
Generalisability	21	Discuss the generalisability (external validity) of the study results	15-18
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	22

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.