Assessment of participation bias in cohort studies: systematic review and meta-regression analysis

Avaliação do viés de participação em estudos de coorte: uma revisão sistemática e metarregressão

Evaluación del sesgo de participación en estudios de cohortes: una revisión sistemática y metarregresión

Sérgio Henrique Almeida da Silva Junior [1,2]
Simone M. Santos [1]
Cláudia Medina Coeli [3]
Marilia Sá Carvalho [4]

[1] *Escola Nacional de Saúde Pública Sergio Arouca, Fundação Oswaldo Cruz, Rio de Janeiro, Brasil.*
[2] *Instituto Benjamin Constant, Rio de Janeiro, Brasil.*
[3] *Instituto de Estudos em Saúde Coletiva, Universidade Federal do Rio de Janeiro, Brasil.*
[4] *Programa de Computação Científica, Fundação Oswaldo Cruz, Rio de Janeiro, Brasil.*

Correspondence
S. H. A. Silva Junior
Rua Jaguari 79, Nova Iguaçu, RJ 26010-425, Brasil.
sergio.edfisica@gmail.com

Abstract

The proportion of non-participation in cohort studies, if associated with both the exposure and the probability of occurrence of the event, can introduce bias in the estimates of interest. The aim of this study is to evaluate the impact of participation and its characteristics in longitudinal studies. A systematic review (MEDLINE, Scopus and Web of Science) for articles describing the proportion of participation in the baseline of cohort studies was performed. Among the 2,964 initially identified, 50 were selected. The average proportion of participation was 64.7%. Using a meta-regression model with mixed effects, only age, year of baseline contact and study region (borderline) were associated with participation. Considering the decrease in participation in recent years, and the cost of cohort studies, it is essential to gather information to assess the potential for non-participation, before committing resources. Finally, journals should require the presentation of this information in the papers.

Selection Bias; Cohort Studies; Epidemiologic Methods

Resumo

A proporção de não-participação em estudos de coorte está associada também à exposição e à probabilidade de ocorrência do evento poder gerar viés nas estimativas de interesse. O objetivo do presente trabalho é realizar uma revisão sistemática e metanálise de artigos que descrevem a participação em estudos de coorte e avaliar as características associadas à participação. Foi realizada uma revisão sistemática (MEDLINE, Scopus e Web of Science), buscando-se artigos que descrevessem a proporção de participação na linha de base de estudos de coorte. De 2.964 artigos inicialmente identificados, foram selecionados 50. Entre esses, a proporção média de participação foi de 64,7%. Utilizando-se o modelo de metarregressão com efeitos mistos, somente a idade, ano da linha de base e a região do estudo (limítrofe) estiveram associados à participação. Considerando a diminuição na participação em anos mais recentes e o custo dos estudos de coorte, é essencial buscar informações que permitam avaliar o potencial de não-participação antes de comprometer os recursos.

Viés de Seleção; Estudos de Coortes; Métodos Epidemiológicos

Background

Among observational studies, the advantages of prospective cohort studies are that they are able to estimate incidence measures directly and are less vulnerable to information bias. However, participation refusal at baseline or follow-up can introduce selection bias when simultaneously associated with both exposure and the outcome [1,2]. As a result, the association between exposure and outcome may differ between participants and non-participants.

Morton et al. [3] observed a tendency for participation in cohort studies to decrease between 1970 and 2003. As the non-participation proportion rises, vulnerability to selection bias tends to increase. Therefore, it is recommended reporting participation proportion in observational studies [4], designing methodological studies to evaluate the impacts of non-participation and evaluating study characteristics that may influence participation [5].

To the best of our knowledge, and in spite of its importance, no systematic evaluation of participation in observational cohort studies is available to guide choices and scientific assessment of validity of conclusions. This present study aims to perform a systematic review and meta-regression of papers describing non-participation bias in cohort studies, and evaluate the studies' characteristics associated with participation proportion.

Methods

We performed a systematic review and meta-regression following the methodology proposed by Higgins & Green [6] and PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) criteria [7].

Search strategy

We searched MEDLINE, Scopus and Web of Science data bases for papers published between January 1978 and November 2014. The query used for the MEDLINE search strategy was: (cooperation[Title/Abstract/MESH] or noncooperation[Title/Abstract/MESH] or non-cooperation[Title/Abstract/MESH] or participant*[Title/Abstract/MESH] or nonparticipant*[Title/Abstract/MESH] or non-participant*[Title/Abstract/MESH] or compliance[Title/Abstract/MESH] or noncompliance[Title/Abstract/MESH] or non-compliance[Title/Abstract/MESH]) AND bias*[Title/Abstract/MESH] AND (cohort*[Title/Abstract/MESH] OR prospective [Title/Abstract/MESH] OR longitudinal [Title/Abstract/MESH]). For the other data bases, the specific syntaxes corresponding to each base were used.

Article titles and abstracts were evaluated by two reviewers working independently in order to ascertain whether they met the criteria for inclusion in the study. Disagreements were assessed by a third reviewer.

Eligibility criteria and data extraction

As specific populations and health problems may induce large differences in participation proportions related to theses specificities, we only included population-based cohort studies on adult (18 to 75 years old) healthy people. We excluded studies that addressed specific populations (eg. pregnant women, patients with specific ailments), review studies and others (eg. genetic studies, surgery, drug therapies). Figure 1 depicts the review flow chart.

The references identified were stored and processed using the JabRef 2.10 software (http://jabref.sourceforge.net/). We collected the participation proportion, the general characteristics of the study (year of baseline contact, place, selection strategy and study outcome). We also evaluated the characteristics of the study population including type (general population vs. working population), participation of women and the mean age. The relevant data was extracted reading the full paper.

Data analysis

A meta-analysis of participation proportion was conducted using mixed-effects models, often called binominal-normal models [8]. Given the heterogeneity of the studies ($I^2 = 99.97\%$; $\tau^2 = 0.54$; $p < 0.001$), we investigated the variables associated with the participation proportion, initially by simple meta-regression models. When the value of *variance accounted for (VAF)* by the model was greater than 5%, the variable was included in the multiple model. VAF indicates the percentage of total heterogeneity that is explained by each moderator. The goodness of fit of the multiple model was evaluated by the likelihood ratio test (LRT).

We analyzed the following variables: year of the baseline contact, participant mean age, proportion of women, selection strategy, population type (general population vs. employees population), study outcome – cardiovascular (baseline category), general health or others (cancer, accident, substance use, incapacity and smoking) – and study region, as divided by United Nations Statistics Division [9] into Continental Europe (baseline category), Northern Europe, USA, and

Others (Asia or Oceania). Spearman correlation coefficient was used to evaluate the relation between the year of the baseline contact and the participation proportion.

The analyses were performed using the metafor [10] library of R software (The R Foundation for Statistical Computing, Vienna, Austria; http://www.r-project.org).

Results

Of the 2,964 original papers initially identified, 50 were selected. Figure 1 summarizes the study selection process.

Table 1 describes the objectives, database, analysis and main results of the selected papers. To evaluate participation, 29 (58%) papers compared participants and non-participants using secondary databases, 15 (30%) used the information available at baseline, and six (12%) used some way of contacting the non-respondents with small questionnaires. Logistic regression models were the most used technique to evaluate participation, used in 18 (40%) of the papers. Passive follow-up studies applied survival (7) and Poisson regression models (4), and a few some combination of different techniques. In eight papers the evaluation was based on frequencies comparison, using baseline characteristics and/or questionnaires. Imputation, weighted regression and simulations were applied in four papers to evaluate and propose analytical methods for correcting potential bias.

Table 2 describes of the overall study characteristics and sample characteristics potentially associated with participation proportion. Most of the publications are concentrated in the years from 2005 to 2014, the oldest having been published in 1978. The studies comprised 40 (80%) geographically population-based, while the remainder were of workers (8), students (1) and recruits (1).

Most of the studies were conducted in Northern Europe (40%). Regarding participant selection, 60% were random sample, the remainder census-based. The most frequent outcomes were overall health condition in twenty-three (46%), and cardiovascular health in forteen. Other outcomes included cancer, accident, substance use, incapacity and smoking. Participant mean age was 49.5 years (SD = 8.2 years). Mean participation proportion was 64.7%, and ranged from 32.2% to 87.3%. Women participation was slightly larger (52.6%) (Table 2).

A negative correlation was found between study year and participation proportion (ρ = -0.38). Figure 2 shows the downward trend in participation proportion. The dotted line indicates the linear regression, an annual rate of decrease of 0.66% (R^2 = 0.1; p = 0.01). The continuous line (a smooth spline) indicates a downward trend in participation, since 1985. The diameters of the circles of each study, identified by the number of the study (id) in Table 1, is proportional to the inverse of the corresponding standard errors in the meta-regression. The larger circles are more influential in the meta-regression.

The simple meta-regression showed association only between participation proportion and year of the baseline contact (OR = 0.97; 95%CI: 0.95-0.99). The multiple meta-regression showed an association between participation proportion, year of the baseline contact (OR = 0.97; 95%CI: 0.95-0.99) and age (OR = 0.97; 95%CI: 0.95-1.00) (Table 3). In other words, for one-year increase in the year of the baseline contact of the study we expect a 3% decrease in the odds of study participation. Likewise, for one-year increase in the mean age of the study participants we expect a 3% reduction in the odds of study participation.

The analysis shows residual heterogeneity τ^2 = 0.41 (p < 0.001) for the participation proportion, suggesting that 18.1% of total heterogeneity can be accounted for by including year of the baseline contact and age. The test for residual heterogeneity is significant (LRT = 42,252.5, df = 33, p = 0.00), indicating that other covariates not considered in the model are influencing the participation proportion.

Discussion

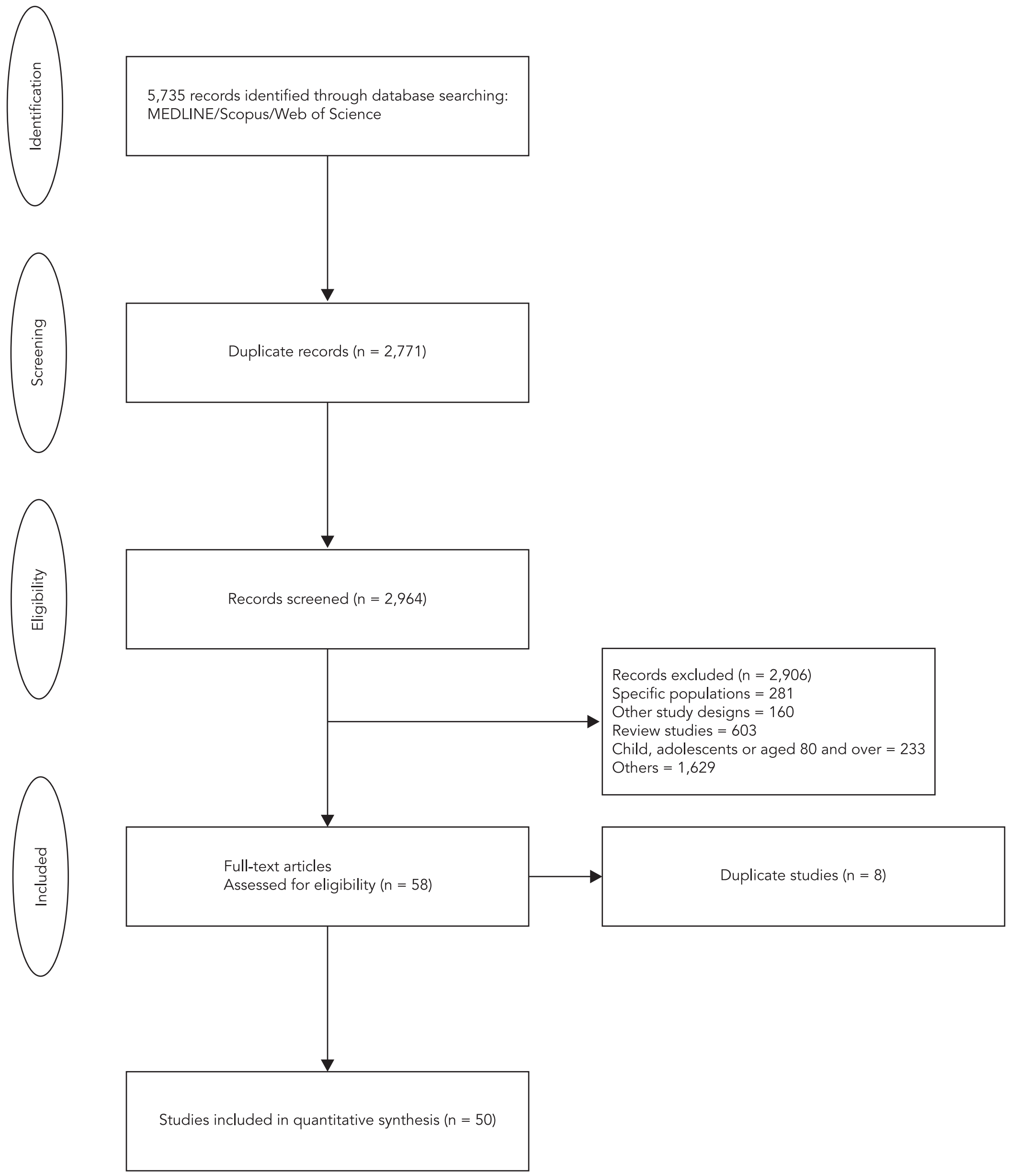
We found a high heterogeneity in participation proportions among the papers evaluating non-participation bias. The most referred characteristics described in the systematic reviewed papers were sociodemographic profile, hospitalization and cancer incidence. Mortality was larger among non-participants. However, in the meta-regression performed only year of the baseline contact and age was associated with participation.

Several strategies involving comparison between participants and non-participants have been proposed to evaluate the potential selection bias in cohort studies: questionnaires to non-participants, comparison of participants according to recruitment moment [4] and passive monitoring of the eligible population using secondary database to assess the outcome [11], the majority of papers in our study.

The results show a decrease in participation in studies over time. The reasons for this decline are not clear, but social changes, and changes in selection and recruitment and in study designs may influence participation [3]. The decrease in participa-

Figure 1

Flowchart of the search and selection of studies included in the meta-analysis.



Source: Moher et al. [7].

tion may be related particularly to the increasing number of studies in recent decades, as well as the proliferation of political and marketing surveys [5]. In addition, increased requests for biological material in epidemiological studies may influence adherence negatively [3].

Previous studies have reported the association between young age and participation cohort studies. Contrary to other articles [12,13,14,15] the proportion of women in the studies showed no association with participation, not even in the simple model. The outcome of the studies was not associ-

Table 1

Characteristics of studies potentially associated with participation.

Id	Reference	Baseline year	Source population	Study region	Outcome	Selection	Mean age	N	Participation rate (%)
1	Studer et al. [37]	2010	Recruits	Continental Europe	Other	Sampling	20	5,457	Overall = 46.2
2	Kaerlev et al. [38]	2007	Workers in general	Northern Europe	General health	Sampling	45	4,489	Overall = 44.7; Men = 79.2; Women = 20.8
3	Langley et al. [39]	2007	General population	Other	Other	Sampling	41.4	2,856	Overall = 79.9; Men = 61.0; Women = 39.0
4	Alkerwi et al. [40]	2006	General population	Continental Europe	Cardiovascular diseases	Sampling	44.3	1,432	Overall = 32.2; Men = 48.7; Women = 51.3
5	Langhammer et al. [41]	2006	General population	Northern Europe	General health	Census	53.1	50,807	Overall = 54.1; Men = 45.4; Women = 54.6
6	Eriksson et al. [42]	2005	General population	Northern Europe	Cardiovascular diseases	Census	47	25,173	Overall = 82.6; Men = 39.4; Women = 60.6
7	Osler et al. [43]	2004	General population	Northern Europe	General health	Census	51	6,292	Overall = 66.2
8	Buckley et al. [44]	2003	General population	Northern Europe	Cardiovascular diseases	Sampling	63.9	493	Overall = 45.6; Men = 80.9; Women = 19.1
9	Schmidt et al. [45]	2003	General population	Continental Europe	General health	Sampling	46.7	7,189	Overall = 64.5; Men = 45.7; Women = 54.3
10	Martikainen et al. [46]	2002	Workers in general	Northern Europe	General health	Sampling	49.6	8,960	Overall = 67.1
11	Holden et al. [47]	2001	General population	Other	General health	Census	65	1,115	Overall = 42.6; Men = 49.9; Women = 50.1
12	Lissner et al. [17]	2001	General population	Northern Europe	General health	Sampling	46.8	850	Overall = 71
13	Stang et al. [16]	2001	General population	Continental Europe	Cardiovascular diseases	Sampling	58.8	8,413	Overall = 53.3; Men = 54.3; Women = 45.7
14	Goldberg et al. [20]	2000	Electric and gas utility workers	Continental Europe	General health	Census	45.1	20,328	Overall = 44.1; Men = 72.9; Women = 27.1
15	Taylor et al. [48]	2000	General population	Other	General health	Sampling	46	6,073	Overall = 49.6; Men = 48.9; Women = 51.1
16	Alonso et al. [49]	1999	Students	Continental Europe	Cardiovascular diseases	Census	35.4	9,907	Overall = 87.3; Men = 40.7; Women = 59.3
17	Knudsen et al. [19]	1999	General population	Northern Europe	Cardiovascular diseases	Census	48.8	18,565	Overall = 63.2; Men = 51.2; Women = 48.8
18	Manjer et al. [29]	1999	General population	Northern Europe	Other	Census	52.9	28,098	Overall = 60.5; Men = 39.4; Women = 60.6
19	Barchielli & Balzi [24]	1998	General population	Continental Europe	Other	Sampling	61.8	1,776	Overall = 85.8; Men = 44.3; Women = 55.7
20	Bergman et al. [50]	1998	General population	Northern Europe	Other	Census	42.7	19,742	Overall = 52.9; Men = 44.5; Women = 55.5
21	Petersen et al. [51]	1998	General population	Northern Europe	Other	Census	63.9	791	Overall = 38.4; Men = 41.8; Women = 58.2
22	Rao et al. [52]	1998	Radiologists	U.S.A.	Other	Census	50.1	90,305	Overall = 68.4
23	Haring et al. [53]	1997	General population	Continental Europe	General health	Sampling	54.1	7,008	Overall = 47.1; Men = 48.2; Women = 51.8
24	Van Loon et al. [30]	1997	General population	Continental Europe	General health	Sampling	42.2	12,097	Overall = 56.5; Men = 44.5; Women = 55.5
25	Drivsholm et al. [54]	1996	General population	Northern Europe	Other	Census	60	1,077	Overall = 64.5; Men = 46.8; Women = 53.2

(continues)

Table 1 (continued)

Id	Reference	Baseline year	Source population	Study region	Outcome	Selection	Mean age	N	Participation rate (%)
26	Jackson et al. [15]	1996	General population	U.S.A.	Cardiovascular diseases	Sampling	54	15,800	Overall = 59.8; Men = 45.1; Women = 54.9
27	Veenstra et al. [28]	1996	General population	Continental Europe	Cardiovascular diseases	Sampling	55.8	15,896	Overall = 51.9; Men = 50.5; Women = 49.5
28	Young et al. [55]	1996	General population	Other	General health	Sampling	47.5	40,395	Overall = 80.4
29	Caetano et al. [56]	1995	General population	U.S.A.	Other	Sampling	42.2	3,106	Overall = 81.8; Men = 48.2; Women = 51.8
30	Garcia et al. [27]	1994	General population	Continental Europe	General health	Sampling	39.1	1,438	Overall = 57.5; Men = 46.5; Women = 53.5
31	Hara et al. [57]	1994	General population	Other	Other	Sampling	55.5	61,447	Overall = 50.5; Men = 46.7; Women = 53.8
32	Kjøller & Thoning [32]	1994	General population	Northern Europe	General health	Sampling	45.6	18,292	Overall = 79.2; Men = 48.5; Women = 51.5
33	Jacobsen et al. [31]	1993	General population	U.S.A.	Cardiovascular diseases	Sampling	60.8	963	Overall = 50.6; Men = 47.3; Women = 52.7
34	Montgomery et al. [58]	1993	Pesticide applicators	U.S.A.	General health	Census	47.3	50,764	Overall = 65.9; Men = 97.0; Women = 3.0
35	Jousilahti et al. [59]	1992	General population	Northern Europe	General health	Sampling	48.1	6,051	Overall = 84.4; Men = 47.1; Women = 52,9
36	May et al. [60]	1992	General population	Other	General health	Sampling	52.2	375,815	Overall = 81.6; Men = 27.6; Women = 72.4
37	Batty & Gale [61]	1991	General population	Northern Europe	Cardiovascular diseases	Sampling	51	6,484	Overall = 70.8; Men = 44.7; Women = 55.3
38	Dugué et al. [62]	1990	General population	Northern Europe	General health	Census	33.2	1,156,671	Overall = 78.1
39	Hara et al. [22]	1990	General population	Other	Cardiovascular diseases	Census	49.6	43,140	Overall = 79.3; Men = 48.0; Women = 52.0
40	Benfante et al. [63]	1989	General population	U.S.A.	Cardiovascular diseases	Census	54.3	8,006	Overall = 71.9
41	Ferrie et al. [23]	1988	Office workers	Northern Europe	General health	Census	46.3	10,297	Overall = 87.1; Men = 67.0; Women = 33.0
42	François et al. [64]	1987	General population	Continental Europe	Other	Sampling	43,3	1,910	Overall = 83.1; Men = 48.9; Women = 51.1
43	Walker et al. [21]	1987	General population	Northern Europe	Cardiovascular diseases	Sampling	46.4	15,364	Overall = 74.3
44	David et al. [65]	1986	General population	Other	Other	Sampling	48.7	2,095	Overall = 78.0; Men = 43.8; Women = 56.2
45	Froom et al. [66]	1985	Industrial employees	Other	General health	Census	45	5,302	Overall = 71.6
46	Bopp et al. [67]	1984	General population	Continental Europe	Cardiovascular diseases	Census	47.6	10,160	Overall = 33.9; Men = 49.1; Women 50.9
47	Criqui et al. [68]	1978	General population	U.S.A.	Cardiovascular diseases	Sampling	52.5	5,052	Overall = 82.1; Men = 46.0; Women = 54.0
48	Lindsted et al. [69]	1976	General population	U.S.A.	General health	Census	53	39,886	Overall = 78.0; Men = 40.8; Women = 59.2
49	Thygesen et al. [70]	1976	General population	Northern Europe	General health	Sampling	53.1	24,464	Overall = 72.0; Men = 45.8; Women = 54.2
50	Vestbo & Rasmussen [71]	1974	Workers in general	Northern Europe	Other	Sampling	55.1	1,404	Overall = 66.1

Table 2

Objectives, database, analysis and results of the selected papers.

Id	Reference	Objectives *	Data source	Analysis	Results
1	Studer et al. [38]	To evaluate differences in substance use between late and early respondents, non-consenters or silent refusers, and whether converting former non-respondents can reduce non-response bias	Baseline information	Logistic model	Late respondents presented a midway pattern of substance use higher than early respondents, but lower than non-consenters
2	Kaerlev et al. [39]	To examine bias on the association between occupational stressors and mental health due to non-participation in a prospective cohort	Secondary data	Survival model	Proportions of gender, age, employment status, sick leave and hospitalization for affective disorders were different in respondents and non-respondents, but low participation at baseline was not associated with mental health outcome
3	Langley et al. [40]	To evaluate factors associated with non-participation in two follow-up contacts of a prospective cohort study of injury outcomes	Baseline information	Poisson model	Non-participation in the closest follow-up contact did not mean non-participation in the next contact; sociodemographic factors were the most important for non-participation
4	Alkerwi et al. [41]	To evaluate the representativeness of the sample with respect to the population and compare characteristics of participants and non participants	Baseline information	Logistic model	Non-participants were similar to participants in gender and place of residence; younger people were under-represented while adults and elderly were over-represented; no discriminating health profiles were detected
5	Langhammer et al. [42]	To study potential participation bias for common symptoms, diseases and socioeconomic status and mortality by participation status	Secondary data, mailed questionnaire.	Negative binomial and survival models	Questionnaire answers indicated higher prevalences of cardiovascular diseases, diabetes mellitus and psychiatric disorders among non-participants; registry data showed higher mortality and lower socioeconomic status among non-participants
6	Eriksson et al. [43]	To assess selective non-response in population-based cohort study on type 2 diabetes, using the population-based drug register for the Stockholm Diabetes Prevention Program	Secondary data	Logistic model	At baseline, non-participants and participants were similar. At follow-up, risks were higher among non-participants
7	Osler et al. [44]	To evaluate changes in association measures in early-life aspects and later health outcomes due to non-response in a follow-up survey	Secondary data	Logistic model and comparison of odds ratios between respondents and complete cohort	A low response rate at age 50 years was related to having a single mother at birth, low educational attainment at age 18, and low cognitive function at ages 12 and 18. The risk of alcohol overuse and tobacco-related diseases was also highest among non-respondents
8	Buckley et al. [45]	To assess baseline differences in participation in a secondary prevention of ischemic heart disease program	Secondary data	Logistic model	Enrollment was lower for women in general and for men with uncontrolled total cholesterol level
9	Schmidt et al. [46]	To identify back-pain-related indicators that could predict attrition in longitudinal studies	Baseline information	Logistic model	The best predictors of attrition were age and baseline response behavior. No bias was found in relation to back pain indicators
10	Martikainen et al. [47]	To estimate impact on social class inequalities in health due to non-response	Secondary data	Linear regression model	Higher social class employees and women were more likely to participate, and sickness absence was higher in non-respondents. Social classes differences did not impact sickness absence in participants or non-participants

(continues)

Table 2 (continued)

Id	Reference	Objectives *	Data source	Analysis	Results
11	Holden et al. [48]	To explore reasons for non-participation in a chronic disease management program	Secondary data	Logistic and multinomial model	Reasons for loss-to-follow-up were: refusals – related to older age, female gender and heart failure; untraceable people – younger, single, indigenous; and death – older individuals, male, who had cancer or heart failure
12	Lissner et al. [18]	To describe 32 years of follow-up of a cohort of women receiving several health examinations	Baseline information, home visits to non-respondents	Linear regression model	Among the 64% of survivors, non-participants and home visited subjects were similar in regard to anthropometry and blood pressure, and both groups were similar to participants in social indicators
13	Stang et al. [17]	To compare recruitment strategies and baseline characteristics of participants and non-participants	Sample of the population	Frequencies comparison	Nonparticipants were more often smokers and of lower social class. A regular relationship with a partner was more frequent among participants
14	Goldberg et al. [21]	To evaluate several variables associated with participation in the French GAZEL cohort	Baseline information	Mixed effects logistic model	Male and older employees in managerial position or retired presented higher response rates. Smoking and alcohol drinking predicted lower participation. Health problems were strong predictors of attrition
15	Taylor et al. [49]	To analyze the association between health-related and socio-demographic indicators and participation in a biomedical cohort study	Sample of the population	Frequencies comparison	Cohort participants were similar to the source population, except for alcohol consumption, which, at an intermediate to high risk level was more frequent among participants
16	Alonso et al. [50]	To evaluate potential predictors of retention in a cohort study and selection bias effect in rate ratio estimates due to loss-to-follow-up	Baseline information	Inverse probability weight logistic model	Several variables (age, smoking, marital status, obesity, past vehicle injury and self-reported history of cardiovascular disease) were associated with the probability of attrition. Obesity, when adjusted for confounding, was similarly associated with hypertension in models with and without inverse probability weight
17	Knudsen et al. [20]	To evaluate characteristics such as health status and specific health problems of non-participants in population-based study, and the potential resulting bias in association measures	Secondary data	Survival model, simulation	Nonparticipants were twice as likely to receive disability pensions (outcome) than participants, and even more if the pension was received for mental disorders. Simulation excluding participants with a similar profile to non-participants reduced the association between common mental disorders and the outcome
18	Manjer et al. [30]	To investigate the effect of non-participation on cancer incidence and mortality	Secondary data, mailed health survey	Survival model	Non-participants presented lower cancer incidence prior to recruitment and higher cancer incidence during recruitment. The proportion of participants in the cohort reporting better health was higher than in the mailed survey
19	Barchielli & Balzi [25]	To analyze the effect on mortality of non-response in a smoking prevalence survey	Secondary data	Poisson model, life table method	All causes mortality was significantly higher among non-respondents, with higher risks for smoking related causes
20	Bergman et al. [51]	To analyze the consequences of attrition in three years after baseline in the PART study	Baseline information, sample of non-respondents	Logistic model	Variables associated with non-participation – low income and education, non-Nordic origin and marital status – were related with depressive mood as well in the first wave
21	Petersen et al. [52]	To investigate wether terminally ill patients' reported quality-of-life scores should be adjusted for non-participation bias	Baseline information	Imputation methods for missing data	Significant underestimation of symptoms in 4 out of 30 comparisons suggest that imputation of quality-of-life scores of non-participants in palliative care is biased based on the available predictors

(continues)

Table 2 (continued)

Id	Reference	Objectives *	Data source	Analysis	Results
22	Rao et al. [53]	To propose a method based on propensity scores to analytically reduce bias due to non-response	Secondary data	Propensity score based on baseline information and data imputation	Among the respondents, there was a higher frequency of women, Caucasian, married and younger people. Differences due to the proposed weighting scheme were small
23	Haring et al. [54]	To determine attrition predictors and evaluate the effect of extensive recruitment procedures on attrition and bias	Baseline information	Logistic model	The main predictors for attrition were late recruitment at baseline, unemployment, low educational level, female gender, and smoking. However attrition bias was not associated with health-related indicators
24	Van Loon et al. [31]	To investigate possible response bias in prevalence estimation and association measures	Baseline information	Logistic model	Respondents, as compared to non-respondents, presented higher socioeconomic status, better subjective health and healthier behaviors. The association measures were similar in respondents and the entire population source
25	Drivsholm et al. [55]	To compare participants at the 20-year follow-up study with non-participants, and to investigate the representativeness of both groups in relation to the population source	Secondary data	Logistic model	Participation decreased to 65% in the 20th follow-up year, when non-participants had lower socioeconomic status, worse health profile and higher mortality rate than participants
26	Jackson et al. [16]	To compare participants with complete clinical examinations to those with just home interview in the the ARIC study	Baseline information	Frequencies comparison	Response rates was similar for white participants, both male and female, and in all study centers. In general, respondents presented higher socioeconomic status and health, but differences were smaller for women
27	Veenstra et al. [29]	To assess association between health status at baseline and nonresponse; to analyze survival in a 5-year follow-up	Secondary data	Logistic model	Among respondents, prevalence of coronary heart disease was higher. However, their mortality was lower than noncontacts
28	Young et al. [56]	To describe factors associated with attrition in a longitudinal study with three age cohorts of women	Baseline information	Logistic model	Variables associated with loss-to-follow-up were: education (lower), non-English-speaking origin, current smoker, poorer health and difficulty managing their income, varying according to cohort age
29	Caetano et al. [57]	To identify characteristics of non-respondents in a survey among couples on violence and drinking	Secondary data	Logistic model	Male non-respondents were younger, less educated, more often unemployed and drinkers. Among women, having been an abuse victim during childhood increased response
30	Garcia et al. [28]	To evaluate attrition in a Spanish population-based cohort	Baseline information	Logistic model	Death and moving to another town were the main reasons of nonresponse. Refusals were associated with working status (disabled and retired) and place of birth (other regions of Spain or in foreign countries); emigration with civil status, age and education as well
31	Hara et al. [58]	To examine factors influencing the recruitment in a study collecting genetic data	Baseline information	Logistic model	Sex (male) and age (younger) presented lower participation rates. The survey location (easy access to participants' residence) and reminders sent to subjects significantly improved the participation rate
32	Kjoller & Thoning [33]	To analyze trends in nonresponse and assess bias on morbidity prevalence	Secondary data	Logistic model	Refusals increased 4.3% in seven years (from 1987 to 1994). Nonrespondents were defined by a combination of sociodemographic characteristics. Nonrespondents hospital admission rates were higher than respondents six months before data collection, and similar afterwards

(continues)

Table 2 (continued)

Id	Reference	Objectives *	Data source	Analysis	Results
33	Jacobsen et al. [32]	To evaluate associations between socioeconomic factors and participation in the Danish National Birth Cohort	Secondary data	Frequencies comparison	Groups with low socioeconomic status were underrepresented as compared to the background population
34	Montgomery et al. [59]	To investigated potential bias due to non-participation in the follow-up of a large cohort study on pesticide applicators	Secondary data	Logistic model	Non-respondents at follow-up were younger, less educated, with lower body mass index and poorer health behaviors but better health conditions, and lower pesticide use. Estimates of exposure-disease associations did not present strong bias
35	Jousilahti et al. [60]	To evaluate total and cause specific mortality comparing participants cohort study	Secondary data	Survival model	At eight year follow up, mortality of non-participating men and women was higher than participating, except for smoking related causes
36	May et al. [61]	To evaluate potential predictors of non-response that are available at baseline (socio-economic-demographic, health,)follow-up duration and contact strategies	Baseline information	Logistic model	Age (younger), sex (male), marital status (single), poorer health conditions, and undernourishment or obesity were associated with non-response
37	Batty & Gale [62]	To investigated variables associated with non-response and its impact on the association measures of several known risk factors and cardiovascular mortality	Secondary data	Survival model	The non-participants had higher CVD mortality than participants. However, the association measures between the risk factors evaluated and the mortality was not affected by non-response
38	Dugue et al. [63]	To estimate excess mortality comparing participants and non-participants in cervical screening	Secondary data	Survival model	All cause mortality and HPV-related mortality was higher for non-participants in cervical screening, and the hazard ratio increased over time
39	Hara et al. [23]	To evaluate the healthy volunteer effect comparing mortality rates among respondents and nonrespondents	Secondary data	Poisson model	Mortality was higher among nonrespondents for all causes studied, although with different effects according do sex. The relative risk varied as well according to the length of follow-up
40	Benfante et al. [64]	To investigate differences between participants and nonparticipants and the potential introduction of bias in the association measures	Secondary data	Frequencies comparison	Total mortality, cancer mortality, and coronary heart disease incidence rates were higher in non-examined men, but the differences decreased over time. No bias was found
41	Ferrie et al. [24]	To evaluate association between nonresponse at baseline and missing follow-up contacts and general mortality, and mortality by socioeconomic position	Secondary data	Survival model	Non-response at baseline and at any follow-up contact was associated with doubling the mortality hazard
42	François et al. [65]	To demonstrate how it is possible to obtain a satisfactory rate of participation in a cohort study, and to compare participants and nonparticipants	Baseline information	Frequencies comparison	The main factors associated with the response rate were: linguistic region, age, income, civil status, educational and alcohol/drugs consumption
43	Walker et al. [22]	To compare the mortality rates and the demographic characteristics between participants and nonparticipants	Secondary data	Frequencies comparison	Non-participants were younger, more likely to be unmarried and work in less skilled jobs. Their mortality rates were higher in the first three years of follow-up, decreasing afterwards. CVD mortality was similar in both groups
44	David et al. [66]	To assess the performance of two different models with two end points each, in analyzing loss-to-follow-up	Secondary data	Logistic and survival model	Survival models performed better than logistic models

(continues)

Table 2 (continued)

Id	Reference	Objectives *	Data source	Analysis	Results
45	Froom et al. [67]	To investigate the healthy volunteer effect in an occupationally cohort of male industrial employees	Secondary data	Survival model	All cause mortality hazard ratio was higher in nonparticipants, and the difference persisted up to 8 years of follow-up
46	Bopp et al. [68]	To evaluate feasibility and quality of linkage procedure in providing follow-up information	Secondary data	Survival model	Linkage success was independent of any variables. Losses in 10 years were 4.7%. Participants of the study had lower mortality than the general population
47	Criqui et al. [69]	To evaluate differences in cardiovascular health status according to participation in a population based study	Baseline information, non-respondents telephone interview	Frequencies comparison	Non-respondents presented more CVD but did not differ on known hypertension. Impact on prevalence estimates was small due to low proportion of non-response
48	Lindsted et al. [70]	To assess the healthy volunteer effect comparing mortality rates between the respondents to a small questionnaire with respondents to a full detailed questionnaire	Secondary data	Survival model	Hazard ratio for different mortality causes was larger for non-respondents, but the difference decreased over time
49	Thygesen et al. [71]	To estimate the effect of drop-out on the association between alcohol intake and mortality	Secondary data	Poisson model	Loss to-follow-up was associated with increased mortality and incidence rates of heart disease, some cancers, and liver diseases related to alcohol intake
50	Vestbo & Rasmussen [72]	To evaluate if baseline characteristics could provide sufficient information about non-response bias	Secondary data	Logistic model	At baseline, respondents and non-respondents presented similar profiles (smoking, lung function and respiratory symptoms). However, non-respondents had larger rates of hospital admission due to respiratory diseases, indicating that equal baseline profile does not protect against non-response bias

CVD: cardiovascular diseases.

* Objectives presented here were the most related to the objective of this review.

ated with participation, in spite of its importance in some of them [11,18,19,20,21,22,23,24,25].

Study region showed no association with participation, in spite of the diversity of places evaluated. Participating in studies voluntarily, giving time, information and biological material is all related to ideas of social capital and volunteering [16], and we expected variation according to local cultural components.

Participation in studies has also been associated with behavioral variables and with general state of health. Non-participants report greater consumption of alcohol, smoking and poor general state of health [12,15,2019,21,22,23,26,27,28,29,30,31,32,33,34]. This information, however, are not available in most publications, limiting the scope of our study.

Strategies to increase participation proportion have been proposed in terms of persuading individuals who are reluctant or hesitant; however, willingness to participate is not always accompanied by commitment to adhere to the study in the long term [35]. Lastly, we agree with the argument of Morton et al. [3] that more information should be requested on the profile of participation and its potential bias.

There is a major need to pursue methodological studies to evaluate the impacts of non-participation on measures of effect in cohort studies. Strategies for that kind of evaluation include comparing participants with non-participants through administrative data bases (sex, age, place of residence), application of summary questionnaires and passive follow-up of eligible population to evaluate mortality [4]. Recent publications from journals with high impact factors show that nonparticipation is mostly ignored or dismissed by many authors, although some are attempting to reduce it or mention it as a limitation in their study [36].

In conclusion, our findings suggest that the drive for participation and compliance should be assessed previously to funding the cohort study, and specific local knowledge should be included in addressing the potential participants.

Figure 2

Correlation of year the baseline year and participation rate.

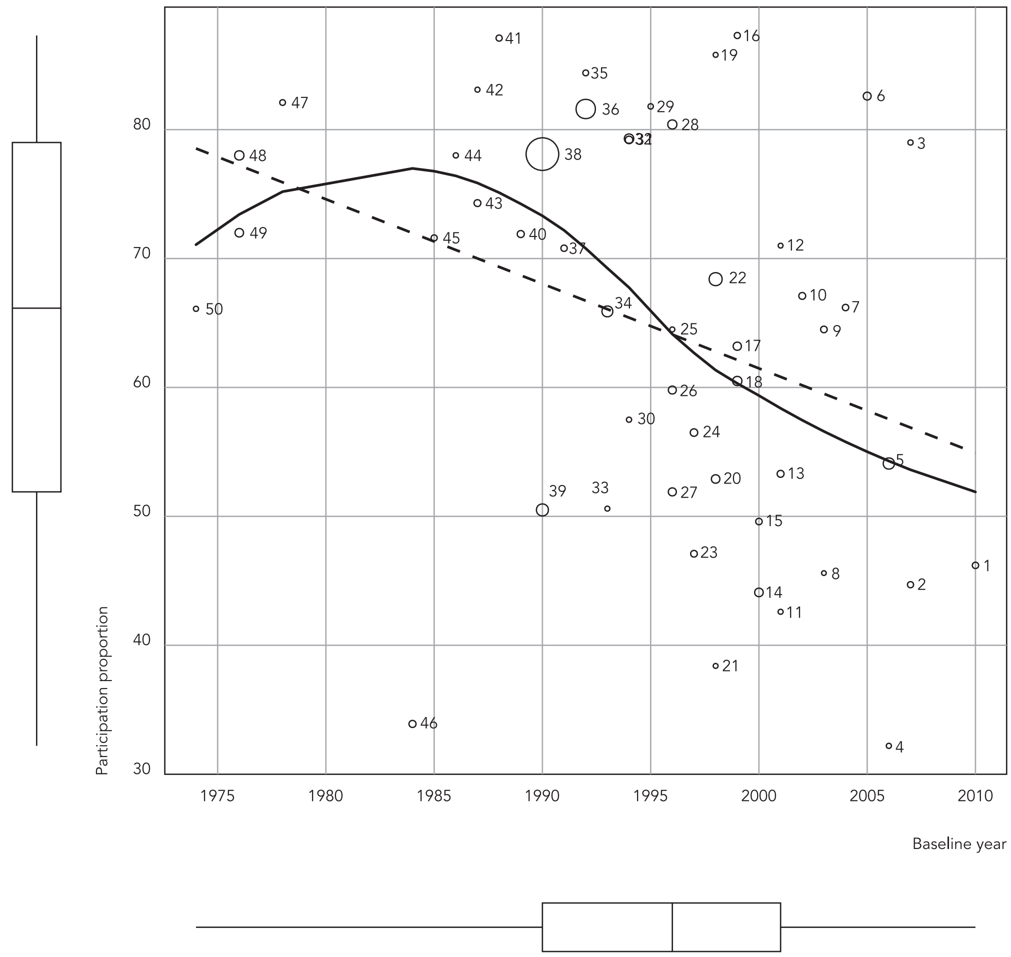


Table 3

Univariate and multiple meta-regression models.

Variables	VAF	Univariate meta-regression models			Multiple meta-regression models VAF = 18.1%		
		OR	95%CI	p-value	OR	95%CI	p-value
Age (mean)	2.6%	0.99	0.96-1.01	0.24	0.97 *	0.95-1.00	0.04
Proportion of women	6.7%	1.01	0.99-1.03	0.21	1,01	0.99-1.03	0.18
Baseline year	11.9%	0.97	0.95-0.99	0.01	0.97 *	0.94-0.99	0.02
Selection (baseline: Sampling)							
Census	0.0%	0.99	0.66-1.48	0.95	-	-	-
Population (baseline: General population)							
Other	0.0%	1.02	0.63-1.67	0.93	-	-	-
Outcomes (baseline: General health)							
Cardiovascular diseases	3.6%	0.78	0.49-1.24	0.30	-	-	-
Other		1.12	0.70-1.80	0.65	-	-	-
Study region (baseline: Northern Europe)							
U.S.A.	7.5%	1.17	0.67-2.04	0.59	0.94	0.51-1.73	0.85
Continental Europe		0.69	0.43-1.12	0.13	0.64	0.38-1.07	0.09
Other		1.10	0.65-1.89	0.72	0.94	0.52-1.69	0.83

95%CI: 95% confidence interval; OR: odds ratio; VAF: variance accounted for.

* For the change of one unit in the variable causes decline the odds of participation.

Resumen

La proporción de no participación en estudios de cohorte se asocia también con la exposición y probabilidad de ocurrencia de hechos que pueden generar sesgos en las estimaciones de interés. El objetivo de este estudio es realizar una revisión sistemática y un metaanálisis de artículos que describen la participación en estudios de cohortes y evaluar las características asociadas con la participación. Una revisión sistemática fue realizada (MEDLINE, Scopus y Web of Science), en busca de artículos que describen la relación de participación basada en estudios de cohortes. Se seleccionaron 2964 artículos, de los cuales se identificaron preliminarmente 50. Entre estos, la proporción promedio de participación fue de un 64,7%. Utilizando la metarregresión, sólo la edad, años de referencia y la región de estudio (borderline) se asociaron con la participación. Teniendo en cuenta la disminución de la participación en los últimos años, y el coste de los estudios de cohortes, es esencial buscar información para evaluar el potencial de la no participación antes de comprometer recursos.

Sesgo de Selección; Estudios de Cohortes; Métodos Epidemiológicos

Contributors

S. H. A. Silva Junior was responsible for the first draft of the manuscript and data analyzes and contributed to the conception and design of the study. S. M. Santos, C. M. Coeli and M. S. Carvalho contributed to the conception and design of the study. All authors contributed significantly to interpreting the results, commented extensively on subsequent revisions, and read and approved the final manuscript.

Acknowledgments

We acknowledge Wolfgang Viechtbauer (at Maastricht University, Maastricht, Netherlands) for his help with the R metafor package, and Israel Souza for revising the selected paper. C. M. Coeli was supported by research fellowship grants from CNPq (304101/2011-7) and Faperj (E26/102.771/2012).

References

1. Kelsey JL. Methods in observational epidemiology. New York: Oxford University Press; 1996.
2. Greenland S. Response and follow-up bias in cohort studies. Am J Epidemiol 1977; 106:184-7.
3. Morton LM, Cahill J, Hartge P. Reporting participation in epidemiologic studies: a survey of practice. Am J Epidemiol 2006; 163:197-203.
4. Stang A. Nonresponse research: an underdeveloped field in epidemiology. Eur J Epidemiol 2003; 18:929-31.
5. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol 2007; 17:643-53.
6. Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions. Volume 5. Chichester: Wiley-Blackwell; 2008.
7. Moher D, Liberati A, Tetzlaff J, Altman DG; The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. PLoS Med 2009; 6:e1000097.
8. De Ridder MAJ, Stijnen T, Hokken-Koelega ACS. A new method to determine mean adult height from incomplete follow-up data. Horm Res 2007; 67:205-10.
9. United Nations Statistics Division. Composition of macro geographical (continental) regions, geographical sub-regions, and selected economic and other groupings. https://unstats.un.org/unsd/methods/m49/m49regin.htm (accessed on 01/Mar/2013).
10. Viechtbauer W. Conducting meta-analyses in R with the metafor package. J Stat Softw 2010; 36(3). http://www.jstatsoft.org/v36/i03/paper.
11. Bopp M, Braun J, Faeh D. Variation in mortality patterns among the general population, study participants, and different types of nonparticipants: evidence from 25 years of follow-up. Am J Epidemiol 2014; 180:1028-35.
12. Harald K, Salomaa V, Jousilahti P, Koskinen S, Vartiainen E. Non-participation and mortality in different socioeconomic groups: the FINRISK population surveys in 1972-92. J Epidemiol Community Health 2007; 61:449-54.
13. Carlsson F, Merlo J, Lindström M, Ostergren P-O, Lithman T. Representativity of a postal public health questionnaire survey in Sweden, with special reference to ethnic differences in participation. Scand J Public Health 2006; 34:132-9.
14. Lundberg I, Damström Thakker K, Hällström T, Forsell Y. Determinants of non-participation, and the effects of non-participation on potential cause-effect relationships, in the PART study on mental disorders. Soc Psychiatry Psychiatr Epidemiol 2005; 40:475-83.
15. Jackson R, Chambless LE, Yang K, Byrne T, Watson R, Folsom A, et al. Differences between respondents and nonrespondents in a multicenter community-based study vary by gender ethnicity. The Atherosclerosis Risk in Communities (ARIC) Study Investigators. J Clin Epidemiol 1996; 49:1441-6.
16. Stang A, Moebus S, Dragano N, Beck EM, Möhlenkamp S, Schmermund A, et al. Baseline recruitment and analyses of nonresponse of the Heinz Nixdorf Recall Study: identifiability of phone numbers as the major determinant of response. Eur J Epidemiol 2005; 20:489-96.
17. Lissner L, Skoog I, Andersson K, Beckman N, Sundh V, Waern M, et al. Participation bias in longitudinal studies: experience from the Population Study of Women in Gothenburg, Sweden. Scand J Prim Health Care 2003; 21:242-7.
18. Boshuizen HC, Viet AL, Picavet HSJ, Botterweck A, van Loon AJM. Non-response in a survey of cardiovascular risk factors in the Dutch population: determinants and resulting biases. Public Health 2006; 120:297-308.
19. Knudsen AK, Hotopf M, Skogen JC, Overland S, Mykletun A. The health status of nonparticipants in a population-based health study: the Hordaland Health Study. Am J Epidemiol 2010; 172:1306-14.
20. Goldberg M, Chastang JF, Zins M, Niedhammer I, Leclerc A. Health problems were the strongest predictors of attrition during follow-up of the GAZEL cohort. J Clin Epidemiol 2006; 59:1213-21.
21. Walker M, Shaper AG, Cook DG. Non-participation and mortality in a prospective study of cardiovascular disease. J Epidemiol Community Health 1987; 41:295-9.
22. Hara M, Sasaki S, Sobue T, Yamamoto S, Tsugane S. Comparison of cause-specific mortality between respondents and nonrespondents in a population-based prospective study: ten-year follow-up of JPHC Study Cohort I. Japan Public Health Center. J Clin Epidemiol 2002; 55:150-6.
23. Ferrie JE, Kivimäki M, Singh-Manoux A, Shortt A, Martikainen P, Head J, et al. Non-response to baseline, non-response to follow-up and mortality in the Whitehall II cohort. Int J Epidemiol 2009; 38:831-7.
24. Barchielli A, Balzi D. Nine-year follow-up of a survey on smoking habits in Florence (Italy): higher mortality among non-responders. Int J Epidemiol 2002; 31:1038-42.
25. Putnam RD. Tuning in, tuning out: the strange disappearance of social capital in America. PS Polit Sci Polit 1995; 28:664-83.
26. Shahar E, Folsom AR, Jackson R. The effect of nonresponse on prevalence estimates for a referent population: insights from a population-based cohort study. Atherosclerosis Risk in Communities (ARIC) Study Investigators. Ann Epidemiol 1996; 6:498-506.
27. Garcia M, Fernandez E, Schiaffino A, Borrell C, Marti M, Borras JM. Attrition in a population-based cohort eight years after baseline interview: The Cornella Health Interview Survey Follow-up (CHIS.FU) Study. Ann Epidemiol 2005; 15:98-104.

28. Veenstra MY, Friesema IHM, Zwietering PJ, Garretsen HFL, Knottnerus JA, Lemmens PHHM. Lower prevalence of heart disease but higher mortality risk during follow-up was found among nonrespondents to a cohort study. J Clin Epidemiol 2006; 59:412-20.
29. Manjer J, Carlsson S, Elmståhl S, Gullberg B, Janzon L, Lindström M, et al. The Malmö Diet and Cancer Study: representativity, cancer incidence and mortality in participants and non-participants. Eur J Cancer Prev 2001; 10:489-99.
30. Van Loon AJM, Tijhuis M, Picavet HSJ, Surtees PG, Ormel J. Survey non-response in the Netherlands: effects on prevalence estimates and associations. Ann Epidemiol 2003; 13:105-10.
31. Jacobsen TN, Nohr EA, Frydenberg M. Selection by socioeconomic factors into the Danish National Birth Cohort. Eur J Epidemiol 2010; 25:349-55.
32. Kjøller M, Thoning H. Characteristics of non-response in the Danish Health Interview Surveys, 1987-1994. Eur J Public Health 2005; 15:528-35.
33. Carter KN, Imlach-Gunasekara F, McKenzie SK, Blakely T. Differential loss of participants does not necessarily cause selection bias. Aust N Z J Public Health 2012; 36:218-22.
34. Forthofer RN. Investigation of nonresponse bias in NHANES II. Am J Epidemiol 1983; 117:507-15.
35. Groves RM, Peytcheva E. The impact of nonresponse rates on nonresponse bias: a meta-analysis. Public Opin Q 2008; 72:167-89.
36. Keeble C, Barber S, Law GR, Baxter PD. Participation bias assessment in three high-impact journals. SAGE Open 2013; 3(4):2158244013511260.
37. Studer J, Baggio S, Mohler-Kuo M, Dermota P, Gaume J, Bertholet N, et al. Examining non-response bias in substance use research: are late respondents proxies for non-respondents? Drug Alcohol Depend 2013; 132:316-23.
38. Kaerlev L, Kolstad HA, Hansen AM, Thomsen JF, Kærgaard A, Rugulies R, et al. Are risk estimates biased in follow-up studies of psychosocial factors with low base-line participation? BMC Public Health 2011; 11:539.
39. Langley JD, Lilley R, Wilson S, Derrett S, Samaranayaka A, Davie G, et al. Factors associated with non-participation in one or two follow-up phases in a cohort study of injured adults. Inj Prev 2013; 19:428-33.
40. Alkerwi A, Sauvageot N, Couffignal S, Albert A, Lair M-L, Guillaume M. Comparison of participants and non-participants to the ORISCAV-LUX population-based study on cardiovascular risk factors in Luxembourg. BMC Med Res Methodol 2010; 10:80.
41. Langhammer A, Krokstad S, Romundstad P, Heggland J, Holmen J. The HUNT study: participation is associated with survival and depends on socioeconomic status, diseases and symptoms. BMC Med Res Methodol 2012; 12:143.
42. Eriksson A-K, Ekbom A, Hilding A, Ostenson C-G. The influence of non-response in a population-based cohort study on type 2 diabetes evaluated by the Swedish Prescribed Drug Register. Eur J Epidemiol 2012; 27:153-62.
43. Osler M, Kriegbaum M, Christensen U, Holstein B, Nybo Andersen A-M. Rapid report on methodology: does loss to follow-up in a cohort study bias associations between early life factors and lifestyle-related health outcomes? Ann Epidemiol 2008; 18:422-4.
44. Buckley B, Murphy AW, Glynn L, Hennigan C. Selection bias in enrollment to a programme aimed at the secondary prevention of ischaemic heart disease in general practice: a cohort study. Int J Clin Pract 2007; 61:1767-72.
45. Schmidt CO, Raspe H, Pfingsten M, Hasenbring M, Basler HD, Eich W, et al. Does attrition bias longitudinal population-based studies on back pain? Eur J Pain 2011; 15:84-91.
46. Martikainen PT, Valkonen T. Excess mortality of unemployed men and women during a period of rapidly increasing unemployment. Lancet 1996; 348:909-12.
47. Holden L, Ware RS, Passey M. Characteristics of nonparticipants differed based on reason for nonparticipation: a study involving the chronically ill. J Clin Epidemiol 2008; 61:728-32.
48. Taylor AW, Dal Grande E, Gill T, Chittleborough CR, Wilson DH, Adams RJ, et al. Do people with risky behaviours participate in biomedical cohort studies? BMC Public Health 2006; 6:11.
49. Alonso A, Seguí-Gómez M, de Irala J, Sánchez-Villegas A, Beunza JJ, Martínez-Gonzalez MA. Predictors of follow-up and assessment of selection bias from dropouts using inverse probability weighting in a cohort of university graduates. Eur J Epidemiol 2006; 21:351-8.
50. Bergman P, Ahlberg G, Forsell Y, Lundberg I. Non-participation in the second wave of the PART study on mental disorder and its effects on risk estimates. Int J Soc Psychiatry 2010; 56:119-32.
51. Petersen MA, Pedersen L, Groenvold M. Does non-participation in studies of advanced cancer lead to biased quality-of-life scores? J Palliat Med 2009; 12:1023-8.
52. Rao RS, Sigurdson AJ, Doody MM, Graubard BI. An application of a weighting method to adjust for nonresponse in standardized incidence ratio analysis of cohort studies. Ann Epidemiol 2005; 15:129-36.
53. Haring R, Alte D, Völzke H, Sauer S, Wallaschofski H, John U, et al. Extended recruitment efforts minimize attrition but not necessarily bias. J Clin Epidemiol 2009; 62:252-60.
54. Drivsholm T, Eplov LF, Davidsen M, Jørgensen T, Ibsen H, Hollnagel H, et al. Representativeness in population-based studies: a detailed description of non-response in a Danish cohort study. Scand J Public Health 2006; 34:623-31.

55. Young AF, Powers JR, Bell SL. Attrition in longitudinal studies: who do you lose? Aust N Z J Public Health 2006; 30:353-61.
56. Caetano R, Ramisetty-Mikler S, McGrath C. Characteristics of non-respondents in a US national longitudinal survey on drinking and intimate partner violence. Addict 2003; 98:791-7.
57. Hara M, Higaki Y, Imaizumi T, Taguchi N, Nakamura K, Nanri H, et al. Factors influencing participation rate in a baseline survey of a genetic cohort in Japan. J Epidemiol 2010; 20:40-5.
58. Montgomery MP, Kamel F, Hoppin JA, Beane Freeman LE, Alavanja MCR, Sandler DP. Effects of self-reported health conditions and pesticide exposures on probability of follow-up in a prospective cohort study. Am J Ind Med 2010; 53:486-96.
59. Jousilahti P, Salomaa V, Kuulasmaa K, Niemelä M, Vartiainen E. Total and cause specific mortality among participants and non-participants of population based health surveys: a comprehensive follow up of 54 372 Finnish men and women. J Epidemiol Community Health 2005; 59:310-5.
60. May AM, Adema LE, Romaguera D, Vergnaud A-C, Agudo A, Ekelund U, et al. Determinants of non-response to a second assessment of lifestyle factors and body weight in the EPIC-PANACEA study. BMC Med Res Methodol 2012; 12:148.
61. Batty GD, Gale CR. Impact of resurvey non-response on the associations between baseline risk factors and cardiovascular disease mortality: prospective cohort study. J Epidemiol Community Health 2009; 63:952-5.
62. Dugué P-A, Lynge E, Rebolj M. Mortality of non-participants in cervical screening: register-based cohort study. Int J Cancer 2014; 134:2674-82.
63. Benfante R, Reed D, MacLean C, Kagan A. Response bias in the Honolulu Heart Program. Am J Epidemiol 1989; 130:1088-100.
64. François Y, Truan P, Gmel G. Response rate and analysis of non-responses in a cohort study. Soz Praventivmed 1997; 42:186-91.
65. David MC, van der Pols JC, Williams GM, Alati R, Green AC, Ware RS. Risk of attrition in a longitudinal study of skin cancer: logistic and survival models can give different results. J Clin Epidemiol 2013; 66:888-95.
66. Froom P, Melamed S, Kristal-Boneh E, Benbassat J, Ribak J. Healthy volunteer effect in industrial workers. J Clin Epidemiol 1999; 52:731-5.
67. Bopp M, Braun J, Faeh D, Gutzwiller F; Swiss National Cohort Study Group. Establishing a follow-up of the Swiss MONICA participants (1984-1993): record linkage with census and mortality data. BMC Public Health 2010; 10:562.
68. Criqui MH, Barrett-Connor E, Austin M. Differences between respondents and non-respondents in a population-based cardiovascular disease study. Am J Epidemiol 1978; 108:367-72.
69. Lindsted KD, Fraser GE, Steinkohl M, Beeson WL. Healthy volunteer effect in a cohort study: temporal resolution in the Adventist Health Study. J Clin Epidemiol 1996; 49:783-90.
70. Thygesen LC, Johansen C, Keiding N, Giovannucci E, Grønbaek M. Effects of sample attrition in a longitudinal study of the association between alcohol intake and all-cause mortality. Addict 2008; 103:1149-59.
71. Vestbo J, Rasmussen FV. Baseline characteristics are not sufficient indicators of non-response bias follow up studies. J Epidemiol Community Health 1992; 46:617-9.

Submitted on 04/Sep/2014
Final version resubmitted on 08/May/2015
Approved on 21/May/2015