

## RESEARCH ARTICLE

# Determinants of unit nonresponse in multi-mode data collection: A multilevel analysis

Finaba Berete<sup>1,2\*</sup>, Johan Van der Heyden<sup>1</sup>, Stefaan Demarest<sup>1</sup>, Rana Charafeddine<sup>1</sup>, Lydia Gisle<sup>1</sup>, Elise Braekman<sup>1</sup>, Jean Tafforeau<sup>1</sup>, Geert Molenberghs<sup>3,4</sup>

**1** Department of Epidemiology and public health, Sciensano, Brussels, Belgium, **2** Department of Public Health, Epidemiology and Health Economics, University of Liège, Liège, Belgium, **3** Interuniversity Institute for Biostatistics and statistical Bioinformatics, Hasselt University, Diepenbeek, Belgium, **4** Interuniversity Institute for Biostatistics and statistical Bioinformatics, KU Leuven, Leuven, Belgium

\* [Finaba.Berete@sciensano.be](mailto:Finaba.Berete@sciensano.be)



## Abstract

### Background

Multi-mode data collection is widely used in surveys. Since several modes of data collection are successively applied in such design (e.g. self-administered questionnaire after face-to-face interview), partial nonresponse occurs if participants fail to complete all stages of the data collection. Although such nonresponse might seriously impact estimates, it remains currently unexplored. This study investigates the determinants of nonresponse to a self-administered questionnaire after having participated in a face-to-face interview.

### Methods

Data from the Belgian Health Interview Survey 2013 were used to identify determinants of nonresponse to self-administered questionnaire ( $n = 1,464$ ) among those who had completed the face-to-face interview ( $n = 8,133$ ). The association between partial nonresponse and potential determinants was explored through multilevel logistic regression models, encompassing a random interviewer effect.

### Results

Significant interviewer effects were found. Almost half (46.6%) of the variability in nonresponse was attributable to the interviewers, even in the analyses controlling for the area as potential confounder. Partial nonresponse was higher among youngsters, non-Belgian participants, people with a lower educational levels and those belonging to a lower income household, residents of Brussels and Wallonia, and people with poor perceived health. Higher odds of nonresponse were found for interviews done in the last quarters of the survey-year. Regarding interviewer characteristics, only the total number of interviews carried out throughout the survey was significantly associated with nonresponse to the self-administered questionnaire.

## OPEN ACCESS

**Citation:** Berete F, Van der Heyden J, Demarest S, Charafeddine R, Gisle L, Braekman E, et al. (2019) Determinants of unit nonresponse in multi-mode data collection: A multilevel analysis. *PLoS ONE* 14(4): e0215652. <https://doi.org/10.1371/journal.pone.0215652>

**Editor:** Florian Fischer, Bielefeld University, GERMANY

**Received:** October 26, 2018

**Accepted:** April 5, 2019

**Published:** April 26, 2019

**Copyright:** © 2019 Berete et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** BHIS data contains sensitive and identifying information and therefore must only be made available upon request. Requests for data access may be made to the Social Security and Health Chamber of the Information Security Committee (hereinafter referred to as the "Social Security and Health Chamber"). Further information regarding the survey and the data access procedure can be found here: [https://his.wiv-isp.be/nl/SitePages/Procedure\\_gegevens2013.aspx](https://his.wiv-isp.be/nl/SitePages/Procedure_gegevens2013.aspx).

**Funding:** The author(s) received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist.

## Conclusions

The results indicate that interviewers play a crucial role in nonresponse to the self-administered questionnaire. Participant characteristics, interview circumstances and interviewer characteristics only partly explain the interviewer variability. Future research should examine further interviewer characteristics that impact nonresponse. The current study emphasises the importance of training and motivating interviewers to reduce nonresponse in multi-mode data collection.

## Background

Combining various modes of data collection in a single survey has become a common practice in survey research [1, 2]. This approach, referred to as multi-mode data collection (MMDC), can apply to different phases of a survey (pre-contact, main data collection, follow-up) [1, 3–5] and can take different forms. There are four types of MMDC in the main data collection phase [1]:

1. One sample, one time period, one questionnaire  
In this type of MMDC, some respondents of the sample use one mode of data collection while the other respondents use another mode in order to collect the same information. An example of this concurrent mixed-mode design is a paper-and-pencil postal survey offering a web-based option.
2. One sample, one time point, but different modes for different parts of the questionnaire  
This MMDC form refers to the situation when different modes are used for a subset of questions in the questionnaire during a single data collection period. A mix of interview (face-to-face or telephone) and self-administered modes enables to combine the advantages of both methods. For instance, a self-administered questionnaire is used for the more sensitive questions to reduce social desirability and enhance privacy, whereas all other questions are administered by an interviewer.
3. One sample, multiple time points  
The third type of MMDC is used in longitudinal and panel studies where the same respondents are surveyed at multiple time points, using a different mode of data collection from one time point to the other. For instance, the Labour Force Surveys conducted in several countries collect data through face-to-face interviews in the first wave of the survey and through telephone interviews in the second wave, with the same participants.
4. Different samples, different modes  
The last type of MMDC employs different modes of data collection for different populations or subgroups. This type of MMDC is often observed in international or regional comparative studies. Indeed, different countries may have different survey traditions and/or practical constraints, calling for different survey modes for collecting the same information. For instance, in a densely populated country, face-to-face surveys are feasible, but this may not be the case in sparsely populated areas where the data may preferably be collected through distal modes.

The “*One sample, one time point, but different modes for different parts of the questionnaire*” approach is widely applied in European health interview and/or health examination surveys, for example in Finland, in the United Kingdom (UK) and in Belgium [6, 7]. In this current

study the term MMDC refers to the form “*One sample, one time point, but different modes for different parts of the questionnaire*”, an approach inherent to the Belgian Health Survey.

Although this MMDC design has several advantages (e.g. reduction of measurement error and social desirability), an important and rarely explored limitation is the potential increase in nonresponse due to the time lags that may occur between the administration of the questionnaires [3]. Even if the second questionnaire is administered immediately after the first, a relatively large number of cases may drop out when switching from one mode of administration to the other. For instance, Sakshaug et al. [8] in their study among alumni from the University of Maryland reported that 26% of those who completed a screening interview never started to complete the subsequent questionnaire. Partial nonresponse in MMDC might also occur because participants consider the second part of the questionnaire as non-relevant, non-essential or non-compulsory to the first, or even as an independent survey.

Nonresponse is a major concern in population health surveys, since it is a threat to the validity of the results [9, 10]. In MMDC, when a participant fails to complete all the stages of data collection, for example by not completing the self-administered questionnaire (SAQ) after having participated in a face-to-face interview, this nonresponse leads to missing data for all the items included in the SAQ. The main implications of this partial nonresponse are a decrease in statistical power, a larger standard error and a nonresponse bias [11].

The selection mechanism of partial nonresponse in MMDC shows some similarities with what can be observed in second stage nonresponse. In a second stage recruitment, participants of one survey are asked to participate in yet another survey [12]. In both situations, data collection takes place in a population that has already been willing to participate in the survey.

To date, numerous studies have explored determinants of nonresponse in a Health Interview Survey (HIS). A relationship has been shown with societal factors and characteristics of the survey design [13], characteristics of the sampled persons [11, 13–16], characteristics of the interviewer [13, 17–20], area characteristics [14, 20, 21] and participant-interviewer interaction [13]. However, most of these studies have explored initial nonresponse. For instance, findings regarding nonresponse in the Belgian HIS (BHIS) concern initial nonresponse, are restricted to the household level [22, 23] and only one of them assessed interviewer effects [23].

Much fewer studies have investigated the selection mechanism in second-stage nonresponse in health surveys. Among these, second stage nonresponse was shown to be related to the characteristics of the sampled persons (e.g. socioeconomic status and nationality) [11, 12]. No study has yet assessed the determinants of partial nonresponse in MMDC population health surveys.

This study aims to investigate nonresponse in MMDC using the 2013 Belgian Health Interview Survey (BHIS 2013). The BHIS questions are administered through a face-to-face interview followed by a SAQ. This is of particular interest due to the increasing nonresponse to the SAQ in the BHIS over the years, especially in 2013. Knowing the determinants of SAQ nonresponse is useful for implementing measures to reduce this nonresponse in future surveys.

The purpose of this study is therefore to examine whether there are systematic differences between interviews regarding nonresponse to the SAQ in the BHIS, and whether these differences can be explained by participant characteristics, interview circumstances and interviewer characteristics.

## Methods

### Ethics statement

The BHIS 2013 was carried out in line with the Belgian privacy legislation. Ethical approval for the data collection was given by the Ethical Committee of Ghent University Hospital on

October 1, 2012. There was no formal written and signed consent foreseen as participation was considered as consent. The selected households were notified about the survey, its practical organization, the institution in charge, the commissioners of the survey and its content via a letter and an information leaflet personally addressed to them. It was also clearly stipulated in the letter and the leaflet that participation is voluntary. All data were fully anonymized before their access. The BHIS data are personal data that contain sensitive information about the respondents. They are considered as coded data where the law of public statistics applies. In order to comply with the current regulations in this domain, access to these data is only possible through a request to the Health Committee of the Data Protection Authority. Further information regarding the survey and the data access procedure can be found here: <http://www.healthsurvey.be>.

### Study population and data

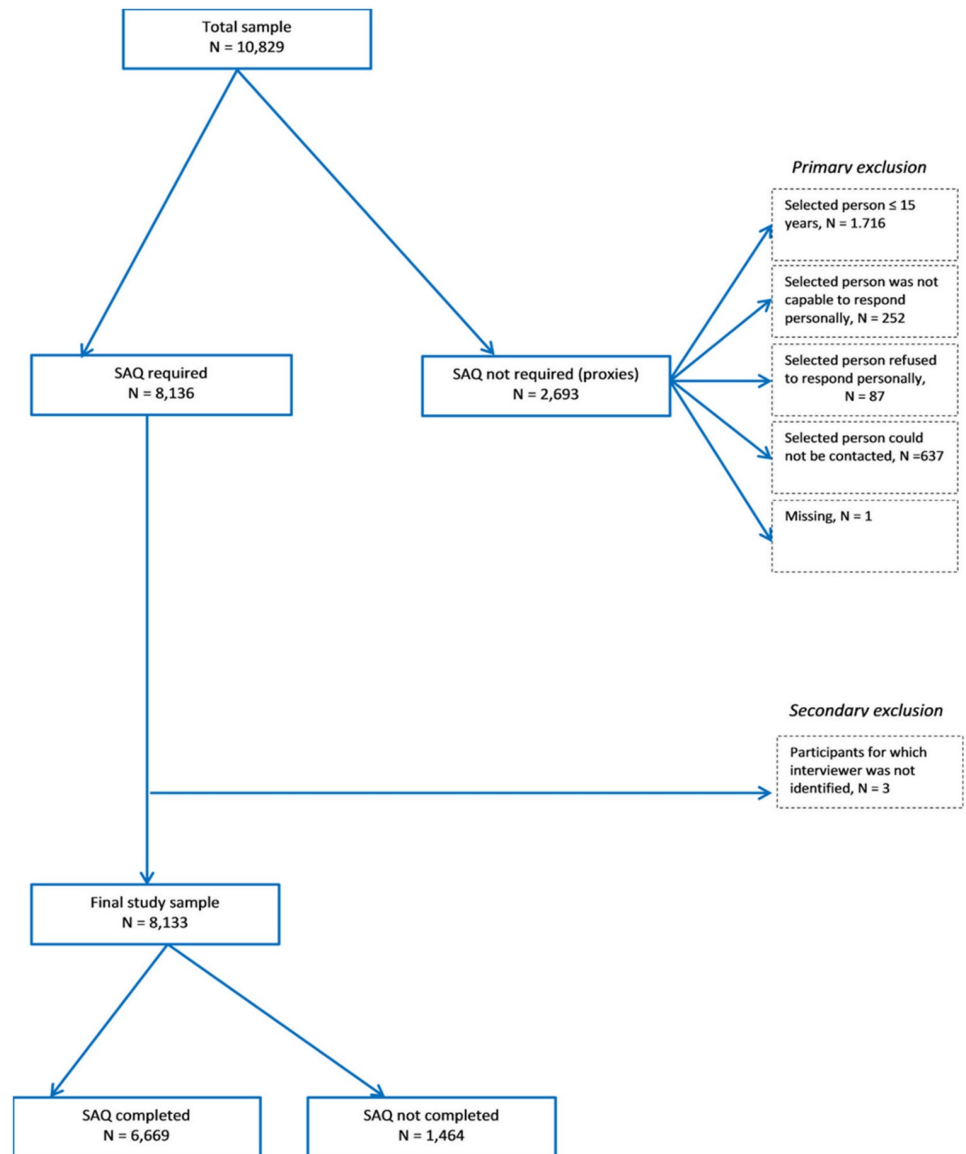
The BHIS is a cross-sectional nationwide household survey on health status, medical consumption and lifestyle habits of the Belgian population. It has been organised every 4 to 5 years since 1997. The participants are selected from the national population register through a multi-stage stratified sample of the population. For this study we used data from the BHIS 2013. In 2013, 9,561 households were invited to participate in the survey. Among them, 5,049 participated, 3,801 refused to participate, 497 were not contactable and 304 were not eligible. So, the response rate was 54% (i.e.,  $5,049/(3,801+5,049+497)$ ) and the participation rate was 57% (i.e.,  $5,049/(3,801+5,049)$ ), at the household level. The detailed methodology of the BHIS is described elsewhere [24]. Data collection is done by means of computer-aided personal interviews (CAPI) and a SAQ at the end of the CAPI interviews. Proxy interviews were conducted for participants younger than 15 years, for those not capable of responding personally, for those who refused to respond personally or for those who were not contactable for at least 3 months. Proxy participants were not eligible to complete the SAQ.

The total sample of the BHIS 2013 was 10,829 individuals interviewed by 183 trained interviewers. As in most face-to-face interview methodology, for logistical reasons, interviewers were not randomly allocated to municipalities (primary sampling unit). In the BHIS fieldwork, one interviewer may work in more than one municipality and sampled cases in one municipality may be assigned to more than one interviewer as municipalities are divided into groups of 50 interviews. However, such situations were rather rare as 86% of the interviewers worked in only one single municipality and as 71% of the municipalities had only one operating interviewer. Completing the SAQ was required for 8,136 individuals and has been completed by 6,669 respondents (82%). After excluding individuals for whom complete information on the interviewer was not available, the final sample size for this study is 8,133 individuals, Fig 1.

### Potential determinants of nonresponse

Determinants were chosen based on their predictive capacity for nonresponse or because they were related to nonresponse in other studies [10, 11, 25]. We considered determinants at three levels: the participant, the interview and the interviewer.

Variables selected at the participant level were gender (male/female), age group (15–34, 35–54 and  $\geq 55$  years), nationality (Belgian/European/Non-European), educational level (primary or no degree/secondary inferior/secondary superior/superior), household size (1 to 4+ members), household income level (quintiles), region of residence (Flanders/Brussels/Wallonia) and having a chronic health condition (yes/no). The presence of chronic conditions (suffering from one or more long-standing illnesses, chronic conditions or handicaps) was used as proxy for health status.



**Fig 1. Flow diagram of the number of participants included in the study.**

<https://doi.org/10.1371/journal.pone.0215652.g001>

Information on the interview circumstances were the duration of the interview (short/average/long) and the quarter of the year in which the interview took place. The duration of the interview is assumed to negatively influence the probability of completing the SAQ. The longer the duration, the higher the probability of not completing the SAQ at the end of the CAPI interview. Also the quarter of the year in which the interview took place may influence nonresponse to the SAQ.

The included interviewer-related variables were gender (male/female), age group (20-44/45-64/>64 years), educational level (primary/secondary inferior/ secondary superior/superior/unknown), professional status (active/unemployed or retired/others or unknown), prior experience with interview surveys (yes/no), number of interviews achieved per interviewer (low/average/high /very high). The number of interviews achieved per interviewer is based on

the total interviews each interviewer completed during the HIS project. It was included in order to evaluate the impact of the interviewer's workload on nonresponse rate.

An additional and important factor which can also influence the response rate to SAQ is the way the interviewer presented the SAQ to respondents, either as a required questionnaire to be completed or as a supplementary ('extra') questionnaire which respondents could, or could not complete. While we do not have the information to verify this hypothesis, it remains an important aspect relating to nonresponse in MMDC.

## Statistical analyses

Weighted proportions and means were calculated to present the characteristics of the population.

A multilevel logistic regression model was used to assess the determinants associated with nonresponse. This approach is advocated and implemented in various cross-sectional studies to disentangle effects associated with each level [25–27]. The survey design and the related within-cluster dependence make ordinary regression modelling inappropriate, so the use of multilevel models is recommended to accommodate such dependence [28]. The interviewer level is therefore included as a random effect to account for possible correlation within clusters [29]. The reasoning behind adding the characteristics of the interviewer as fixed effects is to understand to which extent they explain interviewer variability in nonresponse rate.

To obtain the best fitting and most parsimonious model for the data set and research questions, the 2-levels Generalised Linear Mixed Models with Random intercepts only is used. Individuals (level 1 units) are modelled as nested within interviewers (level 2 units). In addition to the main effects, some interaction effects (between participant and interviewer age group; participant and interviewer gender) were also studied. The model building process was done in 4 steps, as follows:

### Step 1: Model 0 (empty model)

In the first step, the model was fitted without covariates at any level (i.e. only random effects for the intercept) to assess whether there was a significant variation at the interviewer level [26, 27].

### Step 2: Model 1

Participant characteristics were included as fixed effects. The results of model 1 indicate the relationship between the level 1 predictors and the nonresponse to the SAQ.

### Step 3: Model 2

The interview circumstances were added as fixed effects in order to assess to what extent their addition altered the relationships observed in the previous model and to identify the relationship between these predictors and the outcome.

### Step 4: Model 3

The interviewer characteristics were added to model 2. The result of this final model allows to identify the relationship between the level 2 variables and the outcome.

No survey settings were used in the multilevel analyses. However, all variables used in the calculation of the survey weights were included as covariates in the final model.

Across these models, we then compared the estimated interviewer variance components. To allow for this comparison, a rescaling procedure was performed to take into account the implicit scale changes in logit model [30]. We estimate the multilevel logit model with the Stata command `xtmelogit` and rescale parameters by means of the command `meresc` developed by Enzmann and Kohler [31].



Variance components were tested against zero using the likelihood ratio test, asymptotically distributed as a mixture of a Chi-Squared with zero and a Chi-Squared with one degree of freedom [32]. The intraclass correlation coefficient (ICC) was used to measure how much of the total variation in the probability of being a nonrespondent is accounted for by the interviewers [26]. The proportional change in variance (PCV) between the initial model and the model with more terms was also calculated [33, 34]. Associations were expressed as odds ratios (OR) together with their 95% confidence interval (CI). Likelihood ratio (deviance) tests were conducted to compare the relative fit of the different models [26]. The difference in deviance of two nested models follows a chi-square distribution with degrees-of-freedom equal to the number of additional parameters in the larger model [35]. All statistical analyses were carried out using SAS v.9.3 and Stata 14.

## Results

### Background characteristics of BHIS participants and interviewers

The eligible sample consisted of 8,133 participants aged 15 years and older, of which 1,464 (18%) did not complete the SAQ.

Table 1 presents the characteristics of participants and interviewers. BHIS participants were more likely female, highly educated and without chronic conditions. Additional participant characteristics can be found in Table 1.

About half of the interviews were conducted in a time period running between 13 and 37 minutes. The average duration of an interview was 56 minutes (SD approx. 2.5 minutes). The number of interviews carried out was slightly higher in the last two quarters of the year. One hundred and eighty-three interviewers were involved in this study. Interviewers were more often male (62.1%), aged between 45–64 years (53.9%) and highly educated (54.7%). More than 50% of the interviewers held another job and 87.9% of them were experienced interviewers. The average number of interviews carried out per interviewer (in a 12-month period) was 94.5 (SD  $\pm$  1.1), the minimum was 1 and the maximum was 565 interviews. More than half of the interviewers performed at least 47 interviews (Table 1).

### Determinants of nonresponse to SAQ

Table 2 reports the estimated random effect parameters for the different model specifications, as well as model fit statistics. Model 0 (empty model) captures how much of the total variation in the probability of not completing the SAQ is attributable to interviewers. According to the ICC, interviewer effects were very high. Up to 47% of the variability in the nonresponse was accounted for by the interviewers, leaving 53% of the variability to be accounted for by participant and others unexplored factors. Based on the likelihood ratio test (against the logistic model), one concludes that the interviewer level is required and that there was a significant variability in nonresponse to SAQ between interviewers.

As shown in Table 2, adding participant characteristics to the model (model 1) decreased the interviewer variance and explained 13.3% of this variance in the empty model. The likelihood ratio test (against previous model) reveals that adding participant characteristics has significantly improved the model ( $p < 0.0001$ ).

The additional inclusion of interview circumstances variables increased the interviewer variance (model 2) even after rescaling. Together with participant characteristics, interview circumstances and interviewer characteristics explained 20% of the interviewer variance (model 3). Even after controlling for interviewer characteristics, the interviewer effects remained significant.

**Table 1. Characteristics of the study population, interview circumstances and interviewers' characteristics, Belgian Health Interview Survey, 2013 (weighted percentages).**

	%	N
<b>Participant characteristics (N = 8,133)</b>		
<b>Gender</b>		
Male	48.1	3867
Female	51.9	4266
<b>Age group (years)</b>		
15–34	27.1	2106
35–54	35.7	2810
≥ 55	37.2	3217
<b>Nationality</b>		
Belgian	90.3	7042
European	5.8	690
Non-European	3.8	396
Missing	0.04	5
<b>Education level</b>		
Primary/No degree	10.0	860
Secondary inferior	12.9	1151
Secondary superior	34.0	2602
Superior education	42.2	3406
Missing	0.9	114
<b>Chronic conditions</b>		
Yes	29.0	2562
No	70.9	5570
Missing	0.00	1
<b>Household size</b>		
1	19.1	1701
2	33.6	2751
3	17.7	1398
4+	29.6	2283
<b>Household income</b>		
Quintile 1	18.3	1685
Quintile 2	15.9	1243
Quintile 3	17.9	1381
Quintile 4	19.6	1422
Quintile 5	19.2	1455
Missing	9.1	947
<b>Region</b>		
Flanders	58.4	2831
Brussels	9.7	2062
Wallonia	31.9	3240
<b>Interview circumstances</b>		
<b>Length of interview</b>		
Short ( <i>facetime</i> ≤ 13')	25.1	1829
Average (13' < <i>facetime</i> ≤ 37')	49.5	4091
Long ( <i>facetime</i> > 37')	25.4	2213
<b>Quarter of interview</b>		
Quarter 1	24.0	1367

(Continued)



**Table 1.** (Continued)

	%	N
Quarter 2	24.6	1591
Quarter 3	25.5	2123
Quarter 4	25.8	3052
<b>Interviewer characteristics (N = 183)</b>		
<b>Gender</b>		
Male	62.1	110
Female	37.9	73
<b>Age group (years)</b>		
20–44	19.9	45
45–64	53.9	96
> 64	26.2	42
<b>Education level</b>		
Secondary inferior/Secondary superior	34.7	65
Superior education	54.7	95
Unknown	10.6	23
<b>Professional status</b>		
Active	54.3	90
Unemployed/Retired	30.9	60
Other/Unknown	14.7	33
<b>Prior experience with the HIS</b>		
Yes	87.9	141
No	12.1	42
<b>Number of interviews performed by interviewer</b>		
Low ( <i>count</i> < 47)	46.8	88
Average ( $47 \leq \textit{count} < 60$ )	24.5	41
High ( $60 \leq \textit{count} < 98$ )	23.2	36
Very high ( $\geq 98$ )	5.5	18

<https://doi.org/10.1371/journal.pone.0215652.t001>

Based on the likelihood ratio tests, model 3 was the best fitting model and therefore parameter estimates from this model were used to explore the determinants of SAQ nonresponse.

Table 3 presents the rescaled estimated odds ratios and their 95% CI for the model specifications. As stated above, we comment only on the final model.

After adjusting for all other variables, participant characteristics were found to be important determinants of nonresponse. Participants aged 35 to 54 years were more likely not to complete the SAQ compared to those aged 55 years and older. Furthermore, compared to their counterparts, non-Belgian participants, who are lower educated, who suffer from chronic conditions and who belong to a lower income household had higher odds of not responding to the SAQ. As well, significant differences in nonresponse were observed between participants living in Brussels (OR = 2.10, 95% CI 1.29–3.42) and in Wallonia (OR = 1.73, 95% CI 1.06–2.83) compared to those living in Flanders.

Concerning the circumstances of the interview interviews performed during the second, third and fourth quarter of the year were significantly associated with higher odds of nonresponse, with respectively (OR = 1.65, 95% CI 1.26–2.17); (OR = 2.06, 95% CI 1.55–2.66); (OR = 2.39, 95% CI 1.85–3.08) compared to those performed in the first quarter of the year. Furthermore, participants whose interviews took a short time (maximum 13 minutes) had significantly higher odds of being nonrespondent for the SAQ than those with an interview

**Table 2. Estimated variance components, intraclass correlations and model fit statistics for different specifications of the multilevel models for nonresponse to self-administered questionnaire, Belgian Health Interview Survey, 2013.**

Model	0	1	2	3
	Intercept only	With respondent characteristics	With interview circumstance	With interviewer characteristics
<b>Variance components</b>				
<i>Not scaled:</i>				
$\sigma_e^2$ (individual level)	3.290			
$\sigma_u^2$ (interviewer level)	2.869	2.663	2.879	2.593
<i>Scaled</i>				
$\sigma_e^2$ (individual level)		3.072	3.025	2.898
$\sigma_u^2$ (interviewer level)		2.486	2.654	2.284
ICC <sub>interviewers</sub>	0.466	0.447	0.467	0.441
PCV		13.3%	7.5%	20.4%
<b>Model fit</b>				
Log likelihood	-3099.214	-2442.017	-2416.269	-2410.129
LR-Test vs. logistic model (p-value of LR test)	1469.44 (0.000)	802.09 (0.000)	807.92 (0.000)	760.09 (0.000)
LR-Test against previous model (df; p-value of LR test)		1314.394 (7; 0.000)	51.496 (2; 0.000)	12.280 (1; <0.005)

<https://doi.org/10.1371/journal.pone.0215652.t002>

length of average duration (i.e. between 13 and 37 minutes). However, this difference is not statistically significant, but is close to reaching the level of significance.

In the full model, interviewer characteristics were no longer significantly related to nonresponse. Only the number of interviews performed per interviewer remained significant. Compared to those who carried out an average number of interviews, interviewers who performed a low, a high and a very high number of interviews reported higher nonresponse rates to SAQ.

## Discussion

This study investigated whether there were systematic differences between interviewers regarding nonresponse to the SAQ in the BHIS, and whether these differences could be explained by participant characteristics, interview circumstances and interviewer characteristics.

### Variability between interviewers

Substantial variability between interviewers was found. Almost half of the total variance in nonresponse was found at the interviewer level, which confirms that there are systematic differences between interviewers for participants' nonresponse to the SAQ. However, in most interview surveys, interviewers work in a limited geographical area and, as people from certain areas may be more or less likely to cooperate, a significant interviewer effect may simply indicate an existing neighbourhood effect [16, 36]. To test whether this variability did not mainly reflect area variability, municipalities were included as random factors in the empty model. These municipalities differed significantly in nonresponse rate. However, the three-level model with municipalities as level 3 units showed a non-significant area effect, implying that, for this study, a simpler two-level model was indeed sufficient, even after the inclusion of explanatory variables. These results are consistent with previous studies [16, 25].

Table 3. Odds ratios (and 95%CI) of SAQ nonresponse based on the multilevel analysis model, Belgian Health Interview Survey, 2013 (rescaled estimates).

*	Model 0 (Empty model)	Model 1	Model 2	Model 3 <sup>a</sup>
		OR (95% IC)	OR (95% IC)	OR (95% IC)
<b>Fixed effects at individual-level</b>				
<b>Gender</b>				
Male		1.11 (0.96–1.28)	1.11 (0.96–1.28)	1.10 (0.96–1.27)
Female		1.00	1.00	1.00
<b>Age group (years)</b>				
15–34		1.08 (0.90–1.30)	1.06 (0.88–1.28)	1.06(0.88–1.28)
35–54		1.00	1.00	1.00
≥ 55		0.77 (0.65–0.93)*	0.77 (0.64–0.92)*	0.78 (0.65–0.93)*
<b>Nationality</b>				
Belgian		1.00	1.00	1.00
European		1.39 (1.10–1.77)*	1.40 (1.11–1.78)*	1.40 (1.11–1.77)*
Non-European		1.58 (1.18–2.13)*	1.60 (1.19–2.15)*	1.59 (1.19–2.12)*
<b>Education level</b>				
Primary/No degree		1.95 (1.50–2.54)*	1.91 (1.47–2.48)*	1.89 (1.46–2.44)*
Secondary inferior		1.10 (0.86–1.40)	1.06 (0.83–1.35)	1.05 (0.83–1.34)
Secondary superior		1.17 (0.97–1.41)	1.15 (0.95–1.37)	1.14 (0.95–1.37)
Superior education		1.00	1.00	1.00
<b>Region</b>				
Flanders		1.00	1.00	1.00
Brussels		2.38 (1.45–3.93)*	2.27 (1.37–3.78)*	2.10 (1.29–3.42)*
Wallonia		1.75 (1.05–2.91)*	1.72 (1.02–2.89)*	1.73 (1.06–2.83)*
<b>Chronic conditions</b>				
Yes		1.17 (0.99–1.38)	1.20 (1.02–1.42)*	1.20 (1.02–1.41)*
No		1.00	1.00	1.00
<b>Household income</b>				
Quintile 1		2.06 (1.61–2.64)*	2.05 (1.60–2.63)*	2.02 (1.59–2.58)*
Quintile 2		1.76 (1.36–2.27)*	1.78 (1.38–2.30)*	1.76 (1.36–2.26)*
Quintile 3		1.09 (0.84–1.41)	1.10 (0.85–1.43)	1.10 (0.85–1.42)
Quintile 4		1.00	1.00	1.00
Quintile 5		0.92 (0.70–1.21)	0.91 (0.69–1.19)	0.91 (0.70–1.19)
<b>Fixed effects at interview-level</b>				
<b>Quarter of interview</b>				
Quarter 1			1.00	1.00
Quarter 2			1.66 (1.26–2.20)*	1.65 (1.26–2.17)*
Quarter 3			2.06 (1.56–2.71)*	2.06 (1.55–2.66)*
Quarter 4			2.43 (1.87–3.15)*	2.39 (1.85–3.08)*
<b>Length of face-to-face interview</b>				
Short (facetime ≤ 13')			1.19 (0.97–1.46)	1.19 (0.97–1.45)
Average (13' < facetime ≤ 37')			1.00	1.00
Long (facetime > 37')			1.00 (0.83–1.20)	1.00 (0.83–1.19)
<b>Fixed effects at interviewer level<sup>b</sup></b>				
<b>Number of interviews performed by interviewer</b>				
Low (count < 47)				3.30 (1.60–6.79)*
Average (47 ≤ count < 60)				1.00
High (60 ≤ count < 98)				3.08 (1.37–6.92)*

(Continued)

Table 3. (Continued)

*	Model 0 (Empty model)	Model 1	Model 2	Model 3 <sup>a</sup>
		OR (95% IC)	OR (95% IC)	OR (95% IC)
Very high ( $\geq 98$ )				(1.43–10.22)*

Observations: 7,089 respondents, 183 interviewers, ICC = intraclass coefficient of correlation; PCV = proportional change in variance. The PCV expresses the change in the interviewer’s level variance between the initial model and the model with more terms.

<sup>a</sup> Best fitting model;

<sup>b</sup> Only variables that were significantly related to nonresponse after adjusting for all other variables.

\* p value <0.05

<https://doi.org/10.1371/journal.pone.0215652.t003>

It is striking that the interviewer effect in this study is much more pronounced compared to other studies that investigated interviewer effects for nonresponse or non-contact in surveys [16, 25, 36–38]. This difference might simply indicate that the interviewer role in MMDC is more relevant than it would be in initial nonresponse. Indeed, in MMDC the interviewer has to reinforce the participant’s motivation to continue with the SAQ. Maintaining participant motivation depends on the way in which the interviewer presents the whole survey and emphasises that both questionnaires, the CAPI and the SAQ, are necessary parts of the same survey. Furthermore, the willingness to continue with the SAQ may also rest on the interaction with the interviewer during the face-to-face interview. Finally, the interviewer’s lack of rigour in the follow-up of the SAQ can influence its missingness. Therefore, further efforts to reduce nonresponse to SAQ should first focus on reducing interviewer variability.

### Influence of participant characteristics

Several characteristics of the participants were identified as independently linked to nonresponse to SAQ, suggesting that nonresponse to SAQ was not random.

A higher nonresponse was found among participants younger than 55 years old. This group may have less time to take part in the face-to-face interview and to complete the SAQ the same day due to work obligations and family commitments. This result is in agreement with previous studies [12, 39]. Furthermore, in accordance with other studies on initial nonresponse [9, 15, 40] or second stage nonresponse [11, 12, 39, 41], nonrespondents were more likely to be less educated, non-Belgian and belonging to a lower income household. Higher nonresponse among non-Belgians participants, especially among non-Europeans might be related to a lower socio-economic status. The SAQ is a self-administered written questionnaire in French, Dutch, German or English, which probably favours non-Belgians who are well-integrated or well educated. Earlier studies on initial nonresponse in BHIS at the household level reported similar results [22, 42]. Unfortunately, a rigorous comparison with these studies is not feasible due to methodological differences. Nevertheless, it is possible that the mechanisms of initial nonresponse in BHIS were not very different from those found in this study with regard to the background characteristics of the participants, since the characteristics of the reference person are closely related to those of other household members.

Participants living in Brussels were more likely not to respond to the SAQ compared to those living in Flanders. This might be due to the fact that Brussels has a higher proportion of people with lower socio economic status. This population has more difficulties with the language/comprehension and is less interested in participating in SAQs where they have to read the questions themselves. This finding was in line with another study [40]. Furthermore, Brussels is predominantly an urban area and the finding of a higher nonresponse in Brussels is

consistent with other studies that reported lower survey response rates in urbanized settings, relative to rural and less urbanized areas [43].

The results support the hypothesis that the mechanisms of nonresponse to SAQ are not very different from those found in initial nonresponse, as confirmed by earlier studies [11, 12]. Further efforts to reduce nonresponse to the SAQ in BHIS should take into account participants who were more likely to be nonrespondents, i.e. those of working age, those with lower socio-economic status and those living in Brussels.

### **Influence of interview circumstances**

Interview circumstances were found to be important determinants of nonresponse. However, it must be recognised that there were some outliers in the face-to-face interview duration. The minimum length was 0 minutes and the maximum was 24 hours, which are unrealistic values. Obviously, this might be due to technical issues in the CAPI application. Sensitivity analyses were conducted with and without outliers, but the difference was negligible. Therefore, these outliers were kept in the analyses.

Participants interviewed during the last quarters of the year have higher odds of being nonrespondents to the SAQ than those interviewed during the first quarter of the year. This might be a result of the BHIS fieldwork. Indeed, to compensate for the BHIS non-participating rate, the number of interviews to be carried out by the interviewer is increased from one quarter to the next, which leads to an increase in interviewer workload for the last quarters. The aim of the working method is to ensure reaching the number of planned interviews by the end of the 12-month fieldwork period. Japac (2008, cited by [44]) mentioned that interviewer workload is generally seen as a negative influence on their performance. Other authors argue that increased workload give interviewers less time to attempt contact during the most productive times [45]. Similarly, it is possible that in the BHIS, interviewers with heavier workload have less time to wait for the interviewees to complete the SAQ and are therefore tempted to leave the questionnaire to the participant and do not return to get it.

It is striking that a short interview duration yields higher odds of being a SAQ nonrespondent, which was the opposite of what was expected. An explanation of this result could be that short duration interviews reflect the poor motivation of respondents and their lack of interest to continue with a SAQ.

### **Influence of interviewer characteristics**

We observed significant and important interviewer effects. After adjusting for participant characteristics and interview circumstances, none of the interviewers' sex, age, education, professional status and experience explained this variation. These characteristics are thus not a major source of interviewer effects. Only the number of interviews performed per interviewer was significantly related to SAQ nonresponse. Compared to those who carried out an average number of interviews, interviewers who performed a high or very high number of interviews recorded a high proportion of nonresponse for the SAQ. Another surprising result was that interviewers who carried out a few interviews also have higher odds of nonresponse compared to those who carried out an average number of interviews. This might partially be attributed to their characteristics. In fact, although they are all experienced interviewers, almost two thirds of the interviewers who performed a low number of interviews held another job compared to barely a third of those who carried out an average number of interviews. Combining their regular employment with the job of interviewer leads to tighter time constraints and organisation.

These findings are in line with previous studies that found that interviewer background characteristics (sex, age, education, socioeconomic status, work experience) were not good predictors of interviewer level variance [10, 16, 35], whilst respondent-interviewer interaction has found to be a good predictor. For instance, Pickery and colleagues [16], emphasised the importance of the first positive interview experience for the response in the subsequent waves of a panel survey. Another author showed that a positive survey experience enhances the chance that people will participate in subsequent surveys, whereas those without such an experience are less likely to participate [46]. Although this study is not a panel survey, one can hypothesise that the experience of participants during the face-to-face interview and their interaction with the interviewer may affect their propensity to complete the subsequent SAQ. Moreover, another study [10] reported that interviewer level response rate can be predicted by interviewer attitude towards the interviewer's role. Interviewers who were more inclined to favour persuading the participants had a higher response rate, while those who were more inclined to favour acceptance of refusals and not persuading the participants had a lower response rate. Furthermore, other explanations of the higher nonresponse for the SAQ could be of course fatigue or time constraints for respondents and interviewers. However, we suspect that interviewers presented the SAQ to participants as a "second stage recruitment". Although not scheduled as such, it looks that a substantial portion of interviewers thought and introduced the SAQ as a supplementary ('extra') questionnaire which respondents could or not complete. So, after completing the CAPI, respondents "were asked" if they were willing to complete the SAQ. This was of course not the idea, nor the correct procedure. Interviewers should have presented the survey as consisting of two necessary parts and not as a main core interview (CAPI) with complementary questions (SAQ).

Even after taking into account interviewer characteristics, a significant part of interviewer variability remained unexplained. This might be because the relevant interviewer characteristics were not available in our data. Future investigation should include, for example, interviewer attitude and motivation [19, 47], participant-interviewer interaction and interviewer expectation [13], interviewer attitude regarding the persuasion of reluctant respondents [10, 44], and interviewer own behaviour regarding data collection requests [18].

### Strengths and limitations

The major strengths of this study include its relatively large sample of participants and interviewers and the use of multilevel analysis. The large sample size for both participants and interviewers increased the ability to detect specific interviewer effects. A multilevel analysis allowed to separate the potential sources of variability and to explore clustering effects.

Conveniently, an interpenetrated design is used when interviewers are allocated at random to participants [48, 49]. However, in practice, randomly allocating interviewers to participants is expensive and difficult to organise [44, 50]. Consequently, much of the literature on interviewer effects consists of either telephone surveys with small numbers of interviewers or face-to-face surveys without interpenetrated designs [51]. In this study, for practical reasons, interviewers are not randomly assigned to participants. Without these interpenetrating designs, the interviewer and participants explanatory variables are correlates and interviewer intraclass correlation no longer estimates interviewer effects only [50]. This is a limitation of the current study. However, this difficulty has been overcome statistically by adding participant characteristics and interviewer characteristics in two separate steps as recommended by several authors [44, 50]. This therefore allowed distinguishing participants and interviewer effects. Interviewer effects remained significant even after controlling for participant characteristics.

A number of additional study limitations need to be acknowledged. Firstly, the clustering effect at the household level was not explored. However, earlier studies have shown that the impact of intra-household effect was negligible [22, 42]. Secondly, even if the findings are relevant for Belgium, they are restricted to specific situations considered in this study. Therefore, the results cannot automatically be extrapolated to different settings or other surveys. Further research investigating different situations and data structures should be carried out to confirm our results.

The impact of SAQ nonresponse on BHIS results might potentially be important as it took place among participants who might already be a selective sample from the general population. While the analysis of this impact is beyond the present study, an interesting idea for future research would be to assess the consequences of this nonresponse on the estimates.

## Conclusions

To the best of our knowledge, this is the first study investigating partial nonresponse in MMDC. The results contribute to a better understanding of the nonresponse in such a design. The findings highlight that interviewers do play a crucial role in nonresponse to the SAQ in the context of MMDC. Unfortunately, the interviewers' characteristics included in our analyses do not contribute to explain this variability.

The inter-interviewer variability is partially explained by the participant characteristics, interview circumstances and interviewer characteristics involved in the analyses. However, further interviewer characteristics and interviewer-participant interaction should be considered for future research. Nevertheless, this study has a major implication for survey researchers since it provides some support to reduce SAQ nonresponse. Although some of the determinants of nonresponse are beyond the control of the researchers (e.g. participant characteristics), they might be highly useful regarding survey implementation in order to reduce nonresponse.

## Acknowledgments

The authors wish to thank Sandrine de Waleffe from Statistics Belgium for providing interviewer supplementary data.

## Author Contributions

**Conceptualization:** Finaba Berete, Johan Van der Heyden.

**Formal analysis:** Finaba Berete.

**Methodology:** Finaba Berete, Johan Van der Heyden, Geert Molenberghs.

**Supervision:** Johan Van der Heyden.

**Validation:** Finaba Berete, Johan Van der Heyden, Geert Molenberghs.

**Writing – original draft:** Finaba Berete.

**Writing – review & editing:** Finaba Berete, Stefaan Demarest, Rana Charafeddine, Lydia Gisle, Elise Braekman, Jean Tafforeau, Geert Molenberghs.

## References

1. De Leeuw D. To mix or not to mix data collection modes in surveys. *Journal of Official Statistics*. 2005; 21(2): 233.



2. Hox JJ, De Leeuw ED, Zijlmans EA. Measurement equivalence in mixed mode surveys. *Frontiers in psychology*. 2015; 6.
3. Betts P, Lound C. The application of alternative modes of data collection on UK Government social surveys. ONS, London. 2010.
4. Dillman DA, Smyth JD, Christian LM. *Internet, phone, mail, and mixed-mode surveys: the tailored design method*. John Wiley & Sons, 2014.
5. Roberts C. Mixing modes of data collection in surveys: A methodological review. 2007.
6. Aromaa A, Koponen P, Tafforeau J, Vermeire C. Evaluation of health interview surveys and health examination surveys in the European Union. *European Journal of Public Health*. 2003; 13(suppl\_3): 67–72.
7. Koponen P, Aromaa A. Survey Design and methodology in national health interview and health examination surveys. Review of literature, European survey experiences and recommendations. [https://www.researchgate.net/publication/242188042\\_Survey\\_Design\\_and\\_Methodology\\_in\\_National\\_Health\\_Interview\\_and\\_Health\\_Examination\\_Surveys\\_Review\\_of\\_literature\\_European\\_survey\\_experiences\\_and\\_recommendations](https://www.researchgate.net/publication/242188042_Survey_Design_and_Methodology_in_National_Health_Interview_and_Health_Examination_Surveys_Review_of_literature_European_survey_experiences_and_recommendations). 2000. 7-27-2017.
8. Sakshaug JW, Yan T, Tourangeau R. Nonresponse Error, Measurement Error, And Mode Of Data Collection: Tradeoffs in a Multi-mode Survey of Sensitive and Non-sensitive Items. *Public Opinion Quarterly*. 2010; 74(5): 907–933.
9. 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. *Medical Research Methodology*. 2010.
10. De Leeuw E, Hox J, Snijders G, De Heer W. Interviewer Opinions, Attitudes and Strategies Regarding Survey Participation and their Effect on Response. *ZUMA Nachrichten Sepezial*. 1998.
11. Volken T. Second-stage non-response in the Swiss health survey: determinants and bias in outcomes. *BMC Public Health*. 2013; 13(167).
12. Boshuizen HC, Viet AL, Picavet HS, 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(4): 297–308. <https://doi.org/10.1016/j.puhe.2005.09.008> PMID: 16364384
13. Groves RM, Cialdini RB, Couper MP. Understanding the decision to participate in a survey. *Public Opinion Quarterly*. 1992; 56(4): 475–495.
14. Brunton-Smith I, Sturgis P, Williams J. Is success in obtaining contact and cooperation correlated with the magnitude of interviewer variance? *Public Opinion Quarterly*. 2012; 76(2): 265–286.
15. Picavet HS. National health surveys by mail or home interview: effects on response. *Journal of Epidemiology and Community Health*. 2001; 55(6): 408–413. <https://doi.org/10.1136/jech.55.6.408> PMID: 11350999
16. Pickery J, Loosveldt G, Carton A. The effects of interviewer and respondent characteristics on response behavior in panel surveys a multilevel approach. *Sociological Methods & Research*. 2001; 29(4): 509–523.
17. Blom AG, Leeuw ED, Hox J. Interviewer effects on nonresponse in the European Social Survey. 2010.
18. Blom AG, Korbmacher JM. Measuring interviewer characteristics pertinent to social surveys: A conceptual framework. *Survey Methods: Insights from the Field (SMIF)*. 2013.
19. Durrant GB, Groves RM, Staetsky L, Steele F. Effects of interviewer attitudes and behaviors on refusal in household surveys. *Public Opinion Quarterly*. 2010; nfp098.
20. Singer E, Frankel MR, Glassman MB. The effect of interviewer characteristics and expectations on response. *Public Opinion Quarterly*. 1983; 47(1): 68–83.
21. Haunberger S. The effects of interviewer, respondent and area characteristics on cooperation in panel surveys: a multilevel approach. *Quality & Quantity*. 2010; 44(5): 957–969.
22. Burzykowski T, Molenberghs G, Tafforeau J, Van Oyen H, Demarest S, Bellammer L. Missing data in the Health Interview Survey 1997 in Belgium. *Arch Public Health*. 1999; 57: 107–129.
23. Van der Heyden J, Demarest S, Tafforeau J. Interviewer variability in response rates in the Belgian Health Interview Survey.: 2013.
24. Demarest S, Van der Heyden J, Charafeddine R, Drieskens S, Gisle L, Tafforeau J. Methodological basics and evolution of the Belgian Health Interview Survey 1997–2008. *Archives of Public Health*. 2013; 71(1): 24. <https://doi.org/10.1186/0778-7367-71-24>
25. Vassallo R, Durrant GB, Smith PW, Goldstein H. Interviewer effects on non-response propensity in longitudinal surveys: a multilevel modelling approach. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 2015; 178(1): 83–99.

26. Ene M, Leighton EA, Blue GL, Bell BA. *Multilevel Models for Categorical Data Using SAS PROC GLIMMIX: The Basics.*: 2015.
27. Smiley W, Leighton E, Guo Z, Ene M, Bell BA. *An Intermediate Guide to Estimating Multilevel Models for Categorical Data using SAS PROC GLIMMIX.* 2015.
28. Min Z. *Analyzing Multilevel Models with the GLIMMIX Procedure.* SAS Institute Inc. 2014.
29. Rosell-Murphy M, Rodriguez-Blanco T, Morán J, Pons-Vigués M, Elorza-Ricart JM, Rodríguez J et al. Variability in screening prevention activities in primary care in Spain: a multilevel analysis. *BMC Public Health.* 2015; 15(1): 1.
30. Hox JJ, Moerbeek M, Van de Schoot R. *Multilevel analysis: Techniques and applications.* Routledge, 2017.
31. Bristle J, Celidoni M, Dal Bianco C, Weber G. The contribution of paradata to panel cooperation in SHARE. 2014. SHARE Working Paper (19–2014), Munich: MEA, Max Planck Institute for Social Law and Social Policy, 2014.
32. Zhang D, Lin X. Variance component testing in generalized linear mixed models for longitudinal/clustered data and other related topics. *Random effect and latent variable model selection.* Springer; 2008. pp. 19–36.
33. Merlo J, Yang M, Chaix B, Lynch J, Rastam L. A brief conceptual tutorial on multilevel analysis in social epidemiology: investigating contextual phenomena in different groups of people. *Journal of Epidemiology and Community Health.* 2005; 59(9): 729–736. <https://doi.org/10.1136/jech.2004.023929> PMID: 16100308
34. Merlo J, Chaix B, Ohlsson H, Beckman A, Johnell K, Hjerpe P et al. A brief conceptual tutorial of multilevel analysis in social epidemiology: using measures of clustering in multilevel logistic regression to investigate contextual phenomena. *Journal of Epidemiology and Community Health.* 2006; 60(4): 290–297. <https://doi.org/10.1136/jech.2004.029454> PMID: 16537344
35. Winters S, Strating MH, Klazinga NS, Kool RB, Huijsman R. Determining the interviewer effect on CQ Index outcomes: a multilevel approach. *BMC medical research methodology.* 2010; 10(1): 75.
36. Vassallo R, Durrant G, Smith P. Separating interviewer and area effects by using a cross-classified multilevel logistic model: simulation findings and implications for survey designs. *Journal of the Royal Statistical Society: Series A (Statistics in Society).* 2017; 180(2): 531–550.
37. Campanelli P, O'muircheartaigh C. Interviewers, interviewer continuity, and panel survey nonresponse. *Quality & Quantity.* 1999; 33(1): 59–76.
38. Durrant GB, Steele F. Multilevel modelling of refusal and non-contact in household surveys: evidence from six UK Government surveys. *Journal of the Royal Statistical Society: Series A (Statistics in Society).* 2009; 172(2): 361–381.
39. May AM, Adema LE, Romaguera D, Vergnaud AC, 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 medical research methodology.* 2012; 12(1): 1.
40. Dey M, Mohler-Kuo M. An analysis of non-response in a Swiss national survey. *Int J Public Health.* 2013; 58: 323–326. <https://doi.org/10.1007/s00038-012-0377-6> PMID: 22814478
41. Van Loon AJ, Tjihuis M, Picavet HS, Surtees PG, Ormel J. Survey non-response in the Netherlands: effects on prevalence estimates and associations. *Annals of epidemiology.* 2003; 13(2): 105–110. PMID: 12559669
42. Van der Heyden J, Demarest S, Van Herck K, De Bacquer D, Tafforeau J, Van Oyen H. Association between variables used in the field substitution and post-stratification adjustment in the Belgian health interview survey and non-response. *International Journal of Public Health.* 2014; 59(1): 197–206. <https://doi.org/10.1007/s00038-013-0460-7> PMID: 23619721
43. Nummela O, Sulander T, Helakorpi S, Haapola I, Uutela A, Heinonen H et al. Register-based data indicated nonparticipation bias in a health study among aging people. *Journal of Clinical Epidemiology.* 2011; 64(12): 1418–1425. <https://doi.org/10.1016/j.jclinepi.2011.04.003> PMID: 21764556
44. Blom AG, Leeuw ED, Hox J. Interviewer effects on nonresponse in the European Social Survey. *Journal of Official Statistics.* 2011; 27(2): 359–377.
45. Botman SL, Thornberry OT. Survey design features correlates of nonresponse. *ASA Proceedings of the section on Survey Research Methods.* 1992; 309–314.
46. Nederhof AJ. When neutrality is negative: Pleasantness of Most Recent Survey Experience and Nonresponse. *Quality & Quantity* 21[4], 425–432. 1987. Springer.
47. Hox J, De Leeuw E, Couper MP, Groves RM, De Heer W, Kuusela V et al. The influence of interviewers attitude and behavior on household survey nonresponse: An international comparison. *Survey nonresponse.* 2002; 103–120.

48. Groves RM, Fultz NH. Gender effects among telephone interviewers in a survey of economic attitudes. *Sociological Methods & Research*. 1985; 14(1): 31–52.
49. Schnell R, Kreuter F. Separating interviewer and sampling-point effects. *Journal of Official Statistics*. 2005; 21(3): 389.
50. Hox JJ, de Leeuw ED, Kreft IG. The effect of interviewer and respondent characteristics on the quality of survey data: a multilevel model. *Measurement errors in surveys*. 1991; 439–461.
51. Davis RE, Couper MP, Janz NK, Caldwell CH, Resnicow K. Interviewer effects in public health surveys. *Health education research*. 2010; 25(1): 14–26. <https://doi.org/10.1093/her/cyp046> PMID: 19762354