

Original Paper

# Unit Response and Costs in Web Versus Face-To-Face Data Collection: Comparison of Two Cross-sectional Health Surveys

Elise Braekman<sup>1</sup>, PhD; Stefaan Demarest<sup>1</sup>, MSc; Rana Charafeddine<sup>1</sup>, PhD; Sabine Drieskens<sup>1</sup>, BSc; Finaba Berete<sup>1</sup>, MPH; Lydia Gisle<sup>1</sup>, MSc; Johan Van der Heyden<sup>1</sup>, PhD; Guido Van Hal<sup>2</sup>, PhD

<sup>1</sup>Lifestyle and chronic diseases, Epidemiology and public health, Sciensano, Brussels, Belgium

<sup>2</sup>Social Epidemiology and Health Policy, Antwerp University, Antwerp, Belgium

**Corresponding Author:**

Elise Braekman, PhD

Lifestyle and chronic diseases

Epidemiology and public health

Sciensano

Juliette Wytmanstraat 14

Brussels

Belgium

Phone: 32 2 642 57 06

Email: [elise.braekman@sciensano.be](mailto:elise.braekman@sciensano.be)

## Abstract

**Background:** Potential is seen in web data collection for population health surveys due to its combined cost-effectiveness, implementation ease, and increased internet penetration. Nonetheless, web modes may lead to lower and more selective unit response than traditional modes, and this may increase bias in the measured indicators.

**Objective:** This research assesses the unit response and costs of a web study versus face-to-face (F2F) study.

**Methods:** Alongside the Belgian Health Interview Survey by F2F edition 2018 (BHISF2F; net sample used: 3316), a web survey (Belgian Health Interview Survey by Web [BHISWEB]; net sample used: 1010) was organized. Sociodemographic data on invited individuals was obtained from the national register and census linkages. Unit response rates considering the different sampling probabilities of both surveys were calculated. Logistic regression analyses examined the association between mode system and sociodemographic characteristics for unit nonresponse. The costs per completed web questionnaire were compared with the costs for a completed F2F questionnaire.

**Results:** The unit response rate is lower in BHISWEB (18.0%) versus BHISF2F (43.1%). A lower response rate was observed for the web survey among all sociodemographic groups, but the difference was higher among people aged 65 years and older (15.4% vs 45.1%), lower educated people (10.9% vs 38.0%), people with a non-Belgian European nationality (11.4% vs 40.7%), people with a non-European nationality (7.2% vs 38.0%), people living alone (12.6% vs 40.5%), and people living in the Brussels-Capital (12.2% vs 41.8%) region. The sociodemographic characteristics associated with nonresponse are not the same in the 2 studies. Having another European (OR 1.60, 95% CI 1.20-2.13) or non-European nationality (OR 2.57, 95% CI 1.79-3.70) compared to a Belgian nationality and living in the Brussels-Capital (OR 1.72, 95% CI 1.41-2.10) or Walloon (OR 1.47, 95% CI 1.15-1.87) regions compared to the Flemish region are associated with a higher nonresponse only in the BHISWEB study. In BHISF2F, younger people (OR 1.31, 95% CI 1.11-1.54) are more likely to be nonrespondents than older people, and this was not the case in BHISWEB. In both studies, lower educated people have a higher probability of being nonrespondent, but this effect is more pronounced in BHISWEB (low vs high education level: Web, OR 2.71, 95% CI 2.21-3.39 and F2F OR 1.70, 95% CI 1.48-1.95). The BHISWEB study had a considerable advantage; the cost per completed questionnaire was almost 3 times lower (€1 [US \$48]) compared with F2F data collection (€11 [US \$131]).

**Conclusions:** The F2F unit response rate was generally higher, yet for certain groups the difference between web and F2F was more limited. Web data collection has a considerable cost advantage. It is therefore worth experimenting with adaptive mixed-mode designs to optimize financial resources without increasing selection bias (eg, only inviting sociodemographic groups who are keener to participate online for web surveys while continuing to focus on increasing F2F response rates for other groups).

(*J Med Internet Res* 2022;24(1):e26299) doi: [10.2196/26299](https://doi.org/10.2196/26299)

**KEYWORDS**

health interview surveys; data collection mode; face-to-face; web; unit response; response rate; nonresponse; data collection costs; web data; health surveys; internet penetration; web survey; costs

**Introduction**

General population health surveys are an important data source for monitoring the health of the population and for policy making. In this regard, the Belgian Health Interview Survey (BHIS) provides periodic statistics on the health status, health care use, and health determinants of the country's population [1]. Through the BHIS, the statistics requested by the European Statistical Office (Eurostat) in the framework of the European Health Interview Survey (EHIS) are collected. Since its inception in 1997, BHIS data collection has been undertaken through face-to-face (F2F) interviews at participants' homes. In addition, participants aged 15 years and older are asked to fill out a paper-and-pencil questionnaire covering the most sensitive topics. This mix of the interview and self-administered modes is used to exploit the advantages of each [2]. F2F interviewing is particularly suited for the rather long and complex BHIS questionnaire [3,4], and the paper-and-pencil questionnaire reduces the risk of social desirability bias for sensitive topics and enhances privacy [2].

High internet penetration rates and widespread adaptation of the general public to the internet have encouraged experimentation with online HISs in developed countries to exploit the advantages of the web mode. Sending emails is the most cost-effective recruitment strategy for web surveys [5]. Even when using postal mail instead of email invitations, however, a web survey may have a considerable cost advantage over an F2F survey [6]. Moreover, web data collection shortens the duration of data collection [6,7] and is less demanding from a logistical point of view (eg, no interviewer training, no intensive interviewer follow-up) [4] compared to F2F data collection.

Unlike an F2F survey with a paper-and-pencil self-administered part, a web survey is completely self-administered, and this also provides advantages [5]. For example, the burden on respondents is expected to be lower when using the web mode since respondents can complete the questionnaire at a time convenient for them, possibly even spread over several time points, and no interviewer appointment is required [6]. While interviewers can guide and motivate respondents to complete the questionnaire optimally, their presence in itself (ie, social desirability bias), their characteristics (eg, age, gender, and ethnicity), and their interview strategies (eg, incorrect reading of the questions or inadequate probing) can affect the respondent's answers [3,8].

In spite of these advantages, the web mode is rarely used as a single mode of data collection for population-based HISs due to expected noncoverage and nonresponse issues. Although the internet access rate has increased substantially over time in Europe, noncoverage remains an issue; women, older people, and lower educated people still have lower access rates [9]. Moreover, according to a recent study internet users have a better subjective health status than internet nonusers and weighting for sociodemographic characteristics does not

eliminate this observed health difference [10]. In contrast, F2F data collection is expected to be the most reliable approach to obtain a nationally representative sample [11]. In theory, any household in the country can be accessed by an interviewer, and (internet) illiterate people are not excluded by default.

Regardless of the data collection mode, HIS unit response rates have been decreasing over the past decades [12-15]. Yet several meta-analyses found that web surveys perform even worse in terms of unit response rates than surveys conducted using other modes of data collection [16-19]. The higher unit response rates in F2F surveys can be attributed to the higher perceived survey legitimacy and to the persuasiveness of having someone at the doorstep [11]. More equal participation across all sociodemographic groups is also expected when using an F2F mode than when using a web mode. Laaksonen and Heiskanen [20] reported a considerably lower unit response rate of a web compared to F2F study organized in the Finnish general population (25% vs 50%, respectively). Moreover, they showed that being older, being a nonnative Finnish speaker, and being lower educated were associated with a lower probability of responding to a web survey, whereas these sociodemographic variables were not predictors of a lower response in their F2F survey.

Such direct comparisons between F2F and web surveys based on probability samples are rare. Web data collection (using postal mail recruitment) is more frequently compared to paper-and-pencil data collection in terms of coverage and unit nonresponse rates [21-25]. The aim of this study is to compare F2F data collection with web data collection using postal mail recruitment in terms of unit response and costs of a general population health survey using a sample drawn from a national population register. An additional objective is the assessment of the sociodemographic characteristics associated with unit nonresponse using the 2 modes. The hypothesis is that in 2018, our study period, the unit response rate on the web will still be lower than the F2F unit response rate, although this difference between unit response rates will probably vary among the various sociodemographic groups in the population. It is also expected that the costs will be higher in the F2F compared to the web study.

**Methods****Study Design**

Alongside the traditional F2F survey (ie, the Belgian Health Interview Survey by F2F edition 2018 [BHISF2F]), a web-based survey (ie, the Belgian Health Interview Survey by Web [BHISWEB]) was organized. For both studies, authorizations were received from the Belgian privacy commission and from the ethics committee of the Ghent University Hospital.

**Belgian Health Interview Survey by F2F**

A cross-sectional F2F study was conducted in Belgium with a target net sample size of 11,300. The survey was organized at

household level, and households were selected based on the national population register, using a multistage clustered sampling procedure. For every selected household, 3 replacement households matched on statistical sector (ie, a subdivision of a municipality), household size, and age of the reference person were also selected. In addition to this cluster, a substitute cluster of 4 households with no matched characteristics to the first cluster was created in case of nonparticipation of all first cluster households. Sample substitution was applied during data collection: nonparticipating households were substituted, if necessary several times, by replacement households.

The households selected for the BHISF2F received a postal advance letter stating that an interviewer would visit and containing information on the BHIS. In households with a maximum of 4 members, all members were asked to participate. For households with at least 5 members, only 4 members (selected according to a systematic approach) were asked to participate.

The gross of the BHISF2F questions were administered through a computer-assisted personal interview (CAPI), but questions on the most sensitive topics were included in a paper-and-pencil self-administered questionnaire. The latter was completed by (nonproxy) respondents aged 15 years and older during the interview session. The questionnaires were available in Dutch, French, German, and English.

Data collection took place from January 2018 to January 2019 and was organized in collaboration with our fieldwork partner, the Belgian Statistical Office (Statbel). More information can be found in [Multimedia Appendix 1](#) and in the methodological report of the BHISF2F [26].

### **Belgian Health Interview Survey by Web**

A cross-sectional web study was conducted in Belgium with a target net sample size of 1000, a number that was feasible based on the budget set aside for this study. The national population register was used as the sampling frame. Unlike the BHISF2F, the survey was organized at the individual level and only individuals aged 16 to 85 years were selected. Moreover, individuals living in collective or institutional households and individuals living in the German-speaking region of Belgium (East Belgium, <1% of the Belgian population) were excluded from the sampling frame. A multistage clustered sampling procedure, similar to the BHISF2F, was used to select individuals. For every selected individual, 9 replacement individuals who were comparable in terms of statistical sector, sex, and age were also selected. Matched sample substitution was applied during data collection: nonrespondents were replaced, if needed several times, by replacement individuals.

Selected individuals were invited through a postal letter, and one reminder letter was sent after 7 days. The access period of the web survey was 14 days. Participants received a €10 (US \$12) conditional incentive in the form of a gift voucher. The BHISWEB questionnaire was shorter than the BHISF2F questionnaire as the latter contained not only the EHIS questions but also some additional questions for national purposes. The EHIS questions corresponded with all variables requested by

Eurostat in the context of EHIS wave 3 [27]. The questionnaire was available in Dutch, French, and English.

Data collection was organized together with Statbel and took place from April to November 2018, with a break during July and August. More information on the design choices made in this study can be found in [Multimedia Appendix 1](#), [Multimedia Appendix 2](#), and elsewhere [28].

The BHISWEB and BHISF2F studies were based on 2 mutually exclusive samples. As was made clear in the earlier descriptions, the data collection mode as well as other related design features varied between the 2 surveys. Unit response rates were therefore compared between studies using different mode systems, rather than between studies using solely different modes. To assess the differences in these unit response rates, the gross samples of the BHISF2F and the BHISWEB must be made comparable. To improve the comparability the following steps were performed: (1) only individuals invited for the BHISF2F aged 16 to 85 years and not living in an institutionalized environment or in East Belgium were included; (2) only BHISF2F individuals invited for participation between April and November 2018 (excluding the holiday months) were included to correspond with the time frame of the BHISWEB; (3) a system of weighting to adjust for the differential sample selection used in the 2 studies (gross sample for BHISF2F was  $n=7698$  and gross sample for BHISWEB was  $n=6183$ ) and consequently to adjust for the differences in their age, sex, and region distribution was applied. Weights were assigned to the people invited to the BHISWEB to make this gross sample comparable to the BHISF2F gross sample in terms of the age, sex, and region distribution. The calculation of these weights was based on cross classified data on the BHISF2F gross sample in terms of age (16-40 years, 41-65 years, and 66 years and older), sex, and region (Flemish, Brussels-Capital, and Walloon regions). The people invited to the BHISF2F all received a weight of 1.

### **Analyses**

First, unit response rates were calculated using the weights as obtained via the method described earlier. For the BHISWEB, the unit response rate was the number of invited individuals having completed the first questions of the 3 first modules (ie, a set of questions related to the same topic) divided by the number of all invited individuals. This web response rate did not allow noncoverage due to having no internet or computer access to be disentangled from actual nonresponse. For the BHISF2F, the individual response rate and not the household response rate was calculated. More specifically, it was the number of selected individuals from the invited households who completed the first questions of the first 3 CAPI modules divided by the number of all selected individuals from the invited households. For these calculations, we did not differentiate between responses from primary selected individuals and responses from substitutes. We looked at the number of respondents in relation to the number of invited individuals (these individuals may be primary selected or substitutes). In addition to the response rate, the response rate ratio (ie, F2F unit response rate/web unit response rate) was calculated.

Demographic variables derived from the national register (sex, age group, number of household members, region of residence,

nationality, and urbanization rate) were used to calculate group-specific unit response rates for the 2 studies and response rate ratios. In addition, unit response rates and response rate ratios by level of education were calculated. Since the national register does not include information on the socioeconomic status of the invited individuals, data on the education level of invited BHISWEB and BHISF2F individuals were derived from a linkage with the Administrative Census 2011. The highest education level achieved was available in 7 categories according to the International Classification of Education and was recoded in 3 categories: low educational attainment (lower secondary education or less), intermediate educational attainment (higher secondary education and postsecondary nonhigher education), and high educational attainment (higher education). High levels of item-missingness were found for educational level (ie, this information was missing for 25.5% of all selected people). In order to get an idea of how the unit response varied with the substitution process, unit response rates per substitution wave are also presented in [Table 1](#): wave 1 concerns the unit response rate among the initially selected individuals; waves 2 to 4 and higher concern the unit response rate among the activated substitutes in each wave.

Second, logistic regression modeling was used to study the association between unit nonresponse and mode system, sociodemographic characteristics (sex, age group, education level, number of household members, region of residence, nationality, and urbanization rate), and substitution wave (basic model).

Third, to assess whether the effect of the sociodemographic characteristics on nonresponse depended on the mode system (effect modification), interaction terms were added to the basic logistic regression model. If the interaction term was significant, we stratified the regression analyses to calculate the effect of the sociodemographic characteristics by mode system (stratified models).

Due to the high level of item-missingness on the educational level variable, regression-based multiple imputation ( $m=20$ )

procedures were applied for the nonresponse analyses, presuming missingness at random. The SAS PROC MIANALYZE procedure was used for the multiple imputation. All analyses were conducted in SAS Enterprise Guide 7.1 (SAS Institute Inc). They were weighted and took into account the complex sampling designs (stratification in both studies and clustering at household level for the BHISF2F).

Last, a cost analysis was performed. The costs of the BHISWEB study were compared with the costs of an F2F study. A distinction was made between fixed costs (ie, costs regardless of the number of invitations) and variable costs (ie, costs depending on the number of invitations). The reported costs of the BHISWEB study corresponded exactly with expenditure related to BHISWEB. The costs for the F2F data collection could not be based entirely on the BHISF2F expenditure due to differences regarding the target sample size (eg, BHISF2F 11,300 vs BHISWEB 1000) and due to the specific BHISF2F financing, which included reduced tariffs from the printing company and fieldwork partner. Therefore, a cost estimation was performed for an F2F study with a target sample size of 1000 under financial conditions comparable to those in the BHISWEB study. In sum, this F2F data collection consisted of a postal advance letter and a collaboration with interviewers to contact and interview the selected individuals at home (via a CAPI including a paper-and-pencil self-administered questionnaire). Due to the higher perceived legitimacy and the persuasiveness of having someone on the doorstep, an incentive was not provided. This study was implemented with the same fieldwork partner as the BHISWEB survey. Therefore, the cost assessment for the F2F study was mainly based on their tender regarding fieldwork logistics, provision of fieldwork materials, and payment for the interviewers. Concretely, the following costs were included in the comparison: project management, information and communications technology, data warehousing, licenses for survey development, incentives, printing, packaging, postage, and interviewers.

**Table 1.** Weighted unit response rates of the Belgian Health Interview Survey by Web and the Belgian Health Interview Survey by F2F edition 2018.

Characteristics	BHISWEB <sup>a</sup> (n=6183), %	BHISF2F <sup>b</sup> (n=7698), %	Ratio <sup>c</sup> (F2F/web)
<b>Sex</b>			
Male	17.8	41.9	2.35
Female	18.2	44.3	2.43
<b>Age (years)</b>			
16-40	18.1	39.9	2.20
41-65	19.1	44.9	2.35
65+	15.4	45.1	2.93
<b>Education</b>			
Low	10.9	38.0	3.49
Middle	20.8	44.1	2.12
High	29.2	52.7	1.80
<b>Nationality</b>			
Belgian	19.7	43.7	2.22
European	11.4	40.7	3.57
Non-European	7.2	38.0	5.28
<b>Household size</b>			
1	12.6	40.5	3.21
2	19.3	44.3	2.30
3	17.9	42.0	2.35
≥4	20.3	44.2	2.18
<b>Region</b>			
Flemish	22.2	41.6	1.87
Brussels-Capital	12.2	41.8	3.43
Walloon	17.9	45.3	2.53
<b>Urbanicity</b>			
Urban	20.9	39.6	1.89
Suburban	18.9	46.3	2.45
Rural	21.3	53.9	2.53
Brussels-Capital	12.2	41.8	3.43
<b>Substitution wave</b>			
1	20.3	45.2	2.23
2	17.1	41.7	2.44
3	19.4	37.9	1.95
4	17.4	43.9	2.52
≥5	15.3	44.0	2.88

<sup>a</sup>BHISWEB: Belgian Health Interview Survey by Web.

<sup>b</sup>BHISF2F: Belgian Health Interview Survey by F2F edition 2018.

<sup>c</sup>Ratio: face-to-face response rate/web response rate.

## Results

In total, 16.3% (1010/6183) of invited individuals participated in the BHISWEB study and 43.1% (3316/7698) of invited

individuals participated in the BHISF2F study. This resulted in weighted unit response rates of 18.0% for the BHISWEB study and 43.1% for the BHISF2F study (unweighted response rates: BHISWEB 16.3%; BHISF2F 43.1%). The BHISF2F response rate was 2.39 times higher than the BHISWEB response rate.



An overview of the unit response rates by sociodemographic characteristics and substitution wave is provided in [Table 1](#). Regardless of the sociodemographic subgroup in the population, the unit response rate was higher in the F2F study compared to the web study (ratio [F2F/web]>1). Nevertheless, the extent of this difference varied between sociodemographic subgroups. Especially for people aged 65 years and older, with low education levels, of non-Belgian nationality, living in a single-person household, or living in the Brussels-Capital region, the unit response rates were higher in the BHISF2F than in the BHISWEB study (ratio [F2F/web]>2.9). In both the BHISWEB and BHISF2F studies, unit response rates were highest among the initially selected individuals (20.3% in BHISWEB and 45.2% in BHISF2F) and not among the substitutes.

Logistic regression analysis showed that unit nonresponse was significantly higher in the web study compared to the F2F study (OR 3.53, 95% CI 3.18-3.91; [Table 2](#): basic model). The basic model also showed that some sociodemographic characteristics (ie, education level, nationality, household size, urbanicity) were significantly associated with nonresponse.

In addition, the model including interaction terms showed both significant and nonsignificant interaction terms between mode system and sociodemographic characteristics. For sex, household size, and urbanicity, no significant interaction effects were found, which means that the association between nonresponse and these characteristics did not differ by mode system. People living with 2 (OR 0.80, 95% CI 0.70-0.90) or at least 4 other household members (OR 0.78, 95% CI 0.69-0.90) were less likely to be nonrespondents than singles, as were people living in rural (OR 0.61, 95% CI 0.49-0.77) or suburban (OR 0.78, 95% CI 0.68-0.90) areas compared to people living in urban areas (ORs based on full model).

Significant interaction terms were found for age, education level, region, and nationality. Stratified models indicating the

results by mode system showed that after adjusting for all relevant covariates, there was an age effect in the F2F survey; younger people were more likely to be nonrespondents (OR 1.31, 95% CI 1.11-1.54) than older people, while no age effect was found in the web survey. The stratified analyses also showed a stronger education effect in the web study compared to the F2F study (web OR low vs high education level 2.71, 95% CI 2.21-3.39; F2F OR low vs high education level 1.70, 95% CI 1.48-1.95/web OR middle vs high education level 1.63, 95% CI 1.33-1.98; F2F OR middle vs high education level 1.44, 95% CI 1.25-1.65). Moreover, the nationality of invited people was associated with responding to a web survey but not to an F2F survey. People with another European (OR 1.60, 95% CI 1.20-2.13) or a non-European (OR 2.57, 95% CI 1.79-3.70) nationality were less likely to participate in the web survey than people with Belgian nationality. Last, a region effect was found in the web survey but not in the F2F survey. People living in the Brussels-Capital (OR 1.72, 95% CI 1.41-2.10) or Walloon (OR 1.47, 95% CI 1.15-1.87) region were less likely to participate in the web survey than people living in the Flemish Region.

Both fixed and variable costs were considerably lower for web than for F2F data collection ([Table 3](#)). The total cost per completed questionnaire was almost 3 times lower for web data collection (€41 [US \$48]) compared to the F2F data collection (€111 [US \$131]; [Table 3](#)). Two factors accounted for most of this cost difference: payment of the interviewers (for their completed interviews, training sessions attended, and transportation costs) and more expensive project management. The fieldwork follow-up in a web study is quite straightforward and based on automatic programming, but interviewers need extensive individual follow-up. Furthermore, a project manager is paid to deliver interviewer training and perform data checking and cleaning, which is more labor extensive in an F2F study than in a web study.

**Table 2.** Results of the nonresponse analyses (outcome=nonresponse), Belgian Health Interview Survey by Web, Belgian Health Interview Survey by F2F edition 2018.

	Global models		Stratified models	
	Basic <sup>a</sup> , OR <sup>b</sup> (95% CI)	Full <sup>c</sup> , OR (95% CI)	BHISWEB <sup>d</sup> , OR (95% CI)	BHISF2F <sup>e</sup> , OR (95% CI)
<b>Mode system (ref<sup>f</sup> F2F)</b>				
Web	3.53 (3.18-3.91)	N/A <sup>g</sup>	N/A	N/A
<b>Sociodemographic characteristics</b>				
<b>Sex (ref male)</b>				
Female	0.95 (0.89-1.02)	0.95 (0.89-1.02)	N/A	N/A
<b>Age (years; ref 65+)</b>				
16-40	1.12 (0.98-1.28)	N/A	0.86 (0.68-1.09)	1.31 (1.11-1.54)
41-65	1.02 (0.91-1.16)	N/A	0.90 (0.72-1.12)	1.11 (0.95-1.30)
<b>Education (ref high)</b>				
Low	2.00 (1.79-2.24)	N/A	2.71 (2.21-3.39)	1.70 (1.48-1.95)
Middle	1.51 (1.34-1.70)	N/A	1.63 (1.33-1.98)	1.44 (1.25-1.65)
<b>Nationality (ref Belgian)</b>				
European	1.22 (1.04-1.43)	N/A	1.60 (1.20-2.13)	1.08 (0.89-1.32)
Non-European	1.45 (1.20-1.75)	N/A	2.57 (1.79-3.70)	1.10 (0.86-1.41)
<b>Household size (ref 1)</b>				
2	0.80 (0.71-0.90)	0.80 (0.70-0.90)	N/A	N/A
3	0.86 (0.74-1.01)	0.87 (0.75-1.02)	N/A	N/A
≥4	0.79 (0.69-0.91)	0.78 (0.69-0.90)	N/A	N/A
<b>Region (ref Flemish)</b>				
Brussels-Capital	1.14 (0.99-1.31)	N/A	1.72 (1.41-2.10)	0.87 (0.72-1.05)
Walloon	1.09 (0.96-1.25)	N/A	1.47 (1.15-1.87)	0.93 (0.79-1.09)
<b>Urbanicity (ref urban)</b>				
Rural	0.65 (0.52-0.81)	0.61 (0.49-0.77)	N/A	N/A
Suburban	0.85 (0.75-0.97)	0.78 (0.68-0.90)	N/A	N/A
<b>Substitution wave (ref 1)</b>				
2	1.16 (1.02-1.33)	1.16 (1.02-1.32)	N/A	N/A
3	1.20 (1.03-1.40)	1.19 (1.02-1.38)	N/A	N/A
4	1.10 (0.93-1.31)	1.09 (0.92-1.30)	N/A	N/A
≥5	1.15 (0.98-1.36)	1.04 (0.88-1.24)	N/A	N/A

<sup>a</sup>Basic model: ORs based on logistic regression model with nonresponse as outcome and mode system and sociodemographic characteristics and substitution wave as independent variables.

<sup>b</sup>OR: odds ratio.

<sup>c</sup>Full model: ORs based on logistic regression model with nonresponse as outcome and mode system and sociodemographic characteristics, substitution wave, and significant interaction terms between mode system and sociodemographic characteristics as independent variables.

<sup>d</sup>BHISWEB (Belgian Health Interview Survey by Web): ORs based on logistic regression model with nonresponse as outcome and sociodemographic characteristics as independent variable for the BHISWEB study to show the stratified results for sociodemographic characteristics with significant interaction terms in the full model.

<sup>e</sup>BHISF2F (Belgian Health Interview Survey by F2F edition 2018): ORs based on logistic regression model with nonresponse as outcome and sociodemographic characteristics as independent variable for the BHISF2F study to show the stratified results for sociodemographic characteristics with significant interaction terms in the full model.

<sup>f</sup>ref: reference.

<sup>g</sup>N/A: not applicable.

**Table 3.** Cost figures<sup>a</sup> for web versus F2F data collection with a target sample size of 1000, Belgian Health Interview Survey by Web and F2F mode system.

	BHISWEB <sup>b</sup> (n=1010), €	F2F <sup>c</sup> mode system (n=1000), €
<b>Fixed costs</b>		
<b>Fieldwork logistics</b>		
Project management <sup>d</sup>	1641	35,040
ICT <sup>e</sup>	3282	3282
Data ware housing	4922	4922
Sampling	1641	1641
<b>Licenses<sup>f</sup></b>		
CAPISoftware <sup>g</sup>	N/A <sup>h</sup>	5375
Web survey software	8740	N/A
Total fixed costs	20,226	50,260
Fixed costs per completed questionnaire	20	50
<b>Variable costs</b>		
Incentive payments	10,212	N/A
<b>Printing, packaging and postage</b>		
Invitation letter (+folder)	5387	N/A
Reminder letter	4829	N/A
Incentive letter	880	N/A
Interviewer materials and advance letter (+folder)	N/A	4391
Interviewer laptops	N/A	5200
Interviewer payments	N/A	51,057
Total variable costs	21,308	60,648
Variable costs per completed questionnaire	21	61
<b>Total costs</b>		
Total fixed and variable costs	41,534	110,908
Costs per completed questionnaire	41	111

<sup>a</sup>Costs do not include salaries for researchers.

<sup>b</sup>BHISWEB: Belgian Health Interview Survey by Web.

<sup>c</sup>F2F: face-to-face.

<sup>d</sup>Project management includes costs for data control, follow-up, and training of interviewers; testing the programs; and salaries of project managers and administrative employees.

<sup>e</sup>Information and communication technology (ICT) includes the costs for developing an ICT infrastructure to organize the fieldwork.

<sup>f</sup>Costs associated with the training courses for developing the computer-assisted personal interview (CAPI) and web questionnaires were not considered since these development skills had already been acquired.

<sup>g</sup>CAPI: computer-assisted personal interview.

<sup>h</sup>N/A: not applicable.

## Discussion

### Principal Findings

In the context of the BHIS, web and F2F data collection were compared in terms of unit response, taking into account the different sociodemographic groups in the population and financial costs.

A response rate of 18.0% was obtained in the BHISWEB study and 43.1% in the BHISF2F study, making the web survey response almost 2.5 times lower compared to the F2F survey response. For all sociodemographic subgroups in the population, the unit response rate was higher in the F2F study compared to the web study. Nevertheless, the difference between web and F2F was more pronounced among people aged 65 years and older, with low education level, of non-Belgian nationality, living in a single-person household, and living in the



Brussels-Capital region. When taking into account only the individuals initially invited (not the substitutes of nonrespondents), the response rates were higher (20.3% in BHISWEB and 45.2% in BHISF2F). This can be explained by the fact that substitutes of hard-to-reach individuals are selected because they have similar sociodemographic characteristics and, consequently, they also have a higher chance on nonresponse [29].

Different sociodemographic characteristics were associated with nonresponding (or, conversely, with responding) in the BHISWEB compared to the BHISF2F survey. Having a non-Belgian nationality and living in the Brussels-Capital or Walloon regions were associated with a higher nonresponse rate in our web survey but not in our F2F survey. Age, on the other hand, was associated with nonresponse in the F2F study but not in the web study. In the F2F study, older people were less likely to be nonrespondents. In both studies, people with low or intermediate educational levels were less likely to respond than people with high educational levels, but this effect was stronger in the web study. The association between household size and urbanicity and nonresponse did not differ between the studies. Singles and people living in urban areas were less likely to respond to both studies.

The BHISWEB study has a considerable cost advantage over the F2F study; the total cost per completed questionnaire was almost 3 times lower (€1 [US \$48]) compared to the F2F data collection (€11 [US \$131]).

### Strengths and Limitations

A positive aspect of this study is the comparison of an F2F study and a web study using 2 random samples drawn from the national population register and the large amount of sociodemographic information available from both participating and nonparticipating sample members. The latter is not only due to the information obtained through the national register but also due to the efforts made by linking to the Administrative Census 2011 to obtain educational information. The use of this data is not perfect since there is a time delay of 7 years between our data collection and the last administrative census and it includes a considerable number of missing values regarding the highest level of educational attainment achieved. In order to address this item-missingness, multiple imputation procedures were applied under the missingness at random assumption.

One limitation is the fact that noncoverage due to having no internet or computer access could not be disentangled from actual nonresponse in our web survey. In Belgium, 87% of households with at least one household member aged between 16 and 74 years had access to the internet at the time of the study [30], so part of the unit nonresponse is in fact linked to noncoverage. A second limitation was the strict focus on unit response rates (and unit response rate differences between sociodemographic subgroups), although there is evidence that low response rates do not necessarily lead to large nonresponse bias (ie, the difference between the expected estimate based on the respondents and the true value in the population) [31]. A Danish interview-administered study also found that although the nonresponse rate was higher among people with low socioeconomic status, no significant association was found

between health status and nonresponse [32]. Proxy measures for health were used in this study: register data on hospital admission costs and dispensed prescription medicine costs. Moreover, increasing fieldwork efforts might increase the response rate, but this is not necessarily a cost-effective way of minimizing survey error [33]. Next to unit response, assessing factors related to questionnaire breakoff and item response would be of interest, as these are less commonly studied [34].

### Comparison With Prior Work

The response rate difference of 26 percentage points between F2F versus web was higher than the mean difference of 12 percentage points reported in the recent meta-analyses of Daikeler et al [19], who compared response rates between the web and other survey modes. Nonetheless, our web survey was organized in the general population among newly recruited individuals (no panel members), and postal mail instead of email invitations were used. These factors are known to contribute to a higher response rate difference between web surveys and surveys organized using other modes [19]. For all sociodemographic groups, a higher response was found in the F2F study versus web study, which indicates that F2F data collection is still the most appropriate way to achieve acceptable response rates among all sociodemographic groups [7].

In addition, we found that the association between sociodemographic characteristics and nonresponse varied between the BHISWEB and BHISF2F surveys. In line with the results of other studies [6,20], we found, for example, that people of non-Belgian origin participate less in web surveys. This can be explained by their lower internet access rates [35-37]. Moreover, a web questionnaire is self-administered, which means that respondents should be not only internet literate but also capable of reading and fully understanding questions in the official national language. This is not the case when using F2F data collection; respondents must only be able to understand the questions posed by the interviewers, but these interviewers can clarify and repeat questions when needed. Moreover, interviewers can motivate people who are less fluent in the official national language to participate by highlighting the importance of the study in simple language and by referring to the help they will offer during the interview process. This can explain why our results showed that nationality was not associated with a lower F2F response.

Second, young people had higher nonresponse rates in our F2F survey than those in older age groups. This age effect was not found in the web study. Additional analysis showed that the higher F2F nonresponse was related only to their higher noncontact rates because refusal rates did not differ between different age groups. Previous studies also found higher noncontact rates among younger age groups [38,39]. This is attributed to the fact that this working-age population group is less likely to be at home when an interviewer contacts them than people older than 65 years, who are most often retired [38]. In a web survey, these interviewer contacts at home are not required, which may explain why no age effect was found in the web survey. Moreover, younger people have a high probability of meeting the necessary conditions (eg, internet

access) and having the skills to participate in a web survey [5,9,35,36,40].

Third, people with low education levels are less likely to participate in surveys, regardless of the data collection modes used [6,13,20,23,41-43]. Our unit nonresponse analyses by mode system also confirmed this since lower educated people participated less in both studies. Nevertheless, this socioeconomic difference was stronger in the web survey than in the F2F survey. Reasons for this could be lower internet access rates [9,36,37] and less frequent internet use among low educated people [40]. Moreover, this socioeconomic group may have a greater need for interviewers to explain the importance of the survey and to motivate them to participate.

A considerable cost advantage of web versus F2F data collection was reported in this paper. Most other cost comparison studies have also found a major cost advantage when using the web versus other modes (F2F, telephone, paper). A cost comparison conducted in the framework of a cross-sectional parental survey on the mental health of children showed that web survey data collection (using postal mail recruitment and including one reminder letter) was 4 times cheaper than F2F data collection [6]. Substantial cost advantages (half of the cost) were also reported for web data collection (using mail invitations) compared to paper-and-pencil data collection in the context of a parental survey on children's health status [22]. Sinclair et al [21] reported that a web survey (using postal mail invitations) offered a considerable cost advantage compared to a telephone survey organized in the context of a community-based health survey. Nevertheless, their paper-and-pencil survey using mail recruitment had lower costs than their web survey due to the high costs associated with web survey development. By using email instead of postal mail invitations, the cost advantage of a web survey over other modes would even be more pronounced. In most countries, emails cannot be sent to a random sample drawn from the general population because email addresses are not available for researchers to contact potential respondents. Denmark is an exception since a large proportion of Danish citizens have a mandatory digital mailbox that is used for communication with public authorities and to request survey participation [42].

## Conclusion

The use of F2F data collection should be preferred over the use of single-mode web data collection for population-based HISs. This recommendation is based on (1) the considerably lower unit response rates of the web survey compared to the F2F survey (18% vs 43%), (2) markedly lower response rates from some specific sociodemographic groups in the web survey versus F2F survey (ie, older people, low educated people, people of foreign nationalities, and people living alone), and (3) the nonresponse analyses which showed that certain sociodemographic groups—people with low education level, of non-Belgian nationalities, and living in the Brussels-Capital

or Walloon regions—were more disadvantaged in a web study compared to an F2F study.

Lower response rates can induce more bias in measured HIS indicators because there is a greater chance that web respondents and nonrespondents show differences not only in terms of these sociodemographic characteristics but also in terms of their health status and health behavior characteristics. Moreover, if response rates show greater differences between various sociodemographic groups, this will affect comparisons between different sociodemographic groups. Single-mode web data collection would therefore better be restricted to specific target groups with universal internet access (eg, university students, online panel members) and should preferably not be used for surveys organized in the general population. This study did, however, show that web data collection offered a considerable cost advantage compared to F2F data collection.

## Recommendations and Future Prospects

In order to benefit from this cost advantage to some extent without increasing the risk of nonresponse bias, the web mode could be integrated in a mixed-mode design. This approach, in which some respondents complete the questionnaire on the web and other respondents (ie, those unwilling/incapable to participate online) use another mode, is already being tested and used in multiple European HISs [44-46]. A specific mixed-mode methodology, push-to-web, could, for example, be considered: people would first be invited by postal mail to participate online and they would then be contacted by an interviewer only in case of nonparticipation [47].

Based on the results of this study, experimenting with adaptive survey designs in which different sample members are assigned to different data collection modes can also be recommended. A potential strategy could be to invite only sociodemographic groups more eager to participate online for a web HIS while continuing to focus on increasing F2F response rates for the other sociodemographic groups. Tailoring the HIS data collection to different sociodemographic groups could reduce the nonresponse bias without increasing the costs.

When considering mode changes, it might also be useful to experiment with adapted recruitment procedures. Elements to take into account in future research could include the incentives for participation, the number and type of reminders, the use of tailored invitation letters based on age group, and the allocation of experienced interviewers to work with difficult to reach subpopulations.

When evaluating new designs, it might be worthwhile to focus not only on unit response rates but also on other less ambiguous indicators of nonresponse bias. These could include the calculation of R-indicators (R stands for representativeness), indicators that measure the similarity between the respondents to a survey and the sample or the population under investigation [48].

## Acknowledgments

The study was funded by the internal budget of Sciensano. The authors wish to thank the Belgian Statistical Office for organizing the fieldwork.

## Authors' Contributions

EB conducted the literature searches and summaries of previous related work, undertook the statistical analyses, interpreted the results, wrote the initial version of the manuscript, and conducted the revisions. S Driessens developed the web and computer-assisted personal interview questionnaires. EB, S Demarest, and RC conceptualized the study. JVDH and FB contributed to the methods used for the statistical analyses. S Demarest, RC, and GVH were responsible for supervision. S Demarest, RC, S Driessens, FB, LG, JVDH, and GVH substantively revised the earlier drafts of the manuscript. All authors read and approved the final manuscript.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

Sampling, sample substitution, recruitment procedure, and proxy interviewing.

[\[DOCX File, 65 KB-Multimedia Appendix 1\]](#)

## Multimedia Appendix 2

Checklist for Reporting Results of Internet E-Surveys (CHERRIES).

[\[DOCX File, 19 KB-Multimedia Appendix 2\]](#)

## References

1. Demarest S, Van der Heyden J, Charafeddine R, Driessens S, Gisle L, Tafforeau J. Methodological basics and evolution of the Belgian health interview survey 1997-2008. *Arch Public Health* 2013;71(1):24 [FREE Full text] [doi: [10.1186/0778-7367-71-24](https://doi.org/10.1186/0778-7367-71-24)] [Medline: [24047278](https://pubmed.ncbi.nlm.nih.gov/24047278/)]
2. De Leeuw D. To mix or not to mix data collection modes in surveys. *J Off Stat* 2005;21(2):233-255 [FREE Full text]
3. Loosveldt G. Face-to-face interviews. In: De Leeuw ED, Hox JJ, Dillman DA, editors. *International Handbook of Survey Methodology*. New York: Taylor & Francis Group; 2008:210-220.
4. De Leeuw ED. Choosing the method of data collection. In: De Leeuw ED, Hox JJ, Dillman DA, editors. *International Handbook of Survey Methodology*. New York: Taylor & Francis Group; 2008:113-135.
5. Callegaro M, Manfreda K, Vehovar V. *Web Survey Methodology*. London: Sage; 2015.
6. Heiervang E, Goodman R. Advantages and limitations of web-based surveys: evidence from a child mental health survey. *Soc Psychiatry Psychiatr Epidemiol* 2011;46(1):69-76. [doi: [10.1007/s00127-009-0171-9](https://doi.org/10.1007/s00127-009-0171-9)] [Medline: [19921078](https://pubmed.ncbi.nlm.nih.gov/19921078/)]
7. Blair J, Czaja R, Blair E. *Designing Surveys: A Guide to Decisions and Procedures*. 3rd edition. Thousand Oaks: SAGE Publications; 2013.
8. Tourangeau R, Rips L, Rasinski K. *The Psychology of Survey Response*. Cambridge: Cambridge University Press; 2000.
9. Mohorko A, De Leeuw E, Hox J. Internet coverage and coverage bias in Europe: developments across countries and over time. *J Off Stat* 2013;29(4):609-622. [doi: [10.2478/jos-2013-0042](https://doi.org/10.2478/jos-2013-0042)]
10. Schnell R, Noack M, Torregroza S. Differences in general health of internet users and non-users and implications for the use of web surveys. *Surv Res Methods* 2017;11(2):105-123. [doi: [10.18148/srm/2017.v11i2.6803](https://doi.org/10.18148/srm/2017.v11i2.6803)]
11. Roberts C. Mixing modes of data collection in surveys: a methodological review. 2007. URL: <http://eprints.ncrm.ac.uk/418/1/MethodsReviewPaperNCRM-008.pdf> [accessed 2021-12-10]
12. Aromaa A, Koponen P, Tafforeau J, Vermeire C, HIS/HES Core Group. Evaluation of health interview surveys and health examination surveys in the European Union. *Eur J Public Health* 2003;13(3 Suppl):67-72. [doi: [10.1093/eurpub/13.suppl\\_1.67](https://doi.org/10.1093/eurpub/13.suppl_1.67)] [Medline: [14533752](https://pubmed.ncbi.nlm.nih.gov/14533752/)]
13. Galea S, Tracy M. Participation rates in epidemiologic studies. *Ann Epidemiol* 2007;17(9):643-653. [doi: [10.1016/j.annepidem.2007.03.013](https://doi.org/10.1016/j.annepidem.2007.03.013)] [Medline: [17553702](https://pubmed.ncbi.nlm.nih.gov/17553702/)]
14. Morton SMB, Bandara DK, Robinson EM, Carr PEA. In the 21st Century, what is an acceptable response rate? *Aust N Z J Public Health* 2012;36(2):106-108. [doi: [10.1111/j.1753-6405.2012.00854.x](https://doi.org/10.1111/j.1753-6405.2012.00854.x)] [Medline: [22487341](https://pubmed.ncbi.nlm.nih.gov/22487341/)]
15. Tolonen H, Helakorpi S, Talala K, Helasoja V, Martelin T, Prättälä R. 25-year trends and socio-demographic differences in response rates: Finnish adult health behaviour survey. *Eur J Epidemiol* 2006;21(6):409-415. [doi: [10.1007/s10654-006-9019-8](https://doi.org/10.1007/s10654-006-9019-8)] [Medline: [16804763](https://pubmed.ncbi.nlm.nih.gov/16804763/)]
16. Manfreda KL, Bosnjak M, Berzelak J, Haas I, Vehovar V. Web surveys versus other survey modes: a meta-analysis comparing response rates. *Int J Market Res* 2008;50(1):79-104. [doi: [10.1177/147078530805000107](https://doi.org/10.1177/147078530805000107)]
17. Shih T, Fan X. Response rates and mode preferences in web-mail mixed-mode surveys: a meta-analysis. *Int J Internet Sci* 2007;2(1):59-82 [FREE Full text]
18. Blumenberg C, Barros AJD. Response rate differences between web and alternative data collection methods for public health research: a systematic review of the literature. *Int J Public Health* 2018;63(6):765-773. [doi: [10.1007/s00038-018-1108-4](https://doi.org/10.1007/s00038-018-1108-4)] [Medline: [29691594](https://pubmed.ncbi.nlm.nih.gov/29691594/)]

19. Daikeler J, Bosnjak M, Lozar MK. Web versus other survey modes: an updated and extended meta-analysis comparing response rates. *J Surv Stat Methodol* 2020;8(3):513-519. [doi: [10.1093/jssam/smz008](https://doi.org/10.1093/jssam/smz008)]
20. Laaksonen S, Heiskanen M. Comparison of three modes for a crime victimization survey. *J Survey Stat Methodol* 2014;2(4):459-483. [doi: [10.1093/jssam/smu018](https://doi.org/10.1093/jssam/smu018)]
21. Sinclair M, O'Toole J, Malawaraarachchi M, Leder K. Comparison of response rates and cost-effectiveness for a community-based survey: postal, internet and telephone modes with generic or personalised recruitment approaches. *BMC Med Res Methodol* 2012;12:132 [FREE Full text] [doi: [10.1186/1471-2288-12-132](https://doi.org/10.1186/1471-2288-12-132)] [Medline: [22938205](https://pubmed.ncbi.nlm.nih.gov/22938205/)]
22. Hohwü L, Lyshol H, Gissler M, Jonsson SH, Petzold M, Obel C. Web-based versus traditional paper questionnaires: a mixed-mode survey with a Nordic perspective. *J Med Internet Res* 2013;15(8):e173 [FREE Full text] [doi: [10.2196/jmir.2595](https://doi.org/10.2196/jmir.2595)] [Medline: [23978658](https://pubmed.ncbi.nlm.nih.gov/23978658/)]
23. Dannetun E, Tegnell A, Giesecke J. Parents' attitudes towards hepatitis B vaccination for their children. A survey comparing paper and web questionnaires, Sweden 2005. *BMC Public Health* 2007;7:86 [FREE Full text] [doi: [10.1186/1471-2458-7-86](https://doi.org/10.1186/1471-2458-7-86)] [Medline: [17511891](https://pubmed.ncbi.nlm.nih.gov/17511891/)]
24. Kongsved SM, Basnov M, Holm-Christensen K, Hjollund NH. Response rate and completeness of questionnaires: a randomized study of Internet versus paper-and-pencil versions. *J Med Internet Res* 2007;9(3):e25 [FREE Full text] [doi: [10.2196/jmir.9.3.e25](https://doi.org/10.2196/jmir.9.3.e25)] [Medline: [17942387](https://pubmed.ncbi.nlm.nih.gov/17942387/)]
25. Källmén H, Sinadinovic K, Berman AH, Wennberg P. Risky drinking of alcohol in Sweden: a randomized population survey comparing web- and paper-based self-reports. *Nordic Stud Alcohol Drugs* 2017;28(2):123-130. [doi: [10.2478/v10199-011-0013-4](https://doi.org/10.2478/v10199-011-0013-4)]
26. Demarest S, Berete F, Charafeddine R, Van der Heyden J. Gezondheidsenquête 2018: Methodologie. Brussel: Sciensano; 2019. URL: [https://his.wiv-isp.be/nl/Gedeelde%20documenten/Method\\_NL\\_2018.pdf](https://his.wiv-isp.be/nl/Gedeelde%20documenten/Method_NL_2018.pdf) [accessed 2022-01-06]
27. Eurostat. European Health Interview Survey (EHIS wave 3) Methodological Manual. Luxembourg: Publication Office of the European Union; 2018. URL: <https://ec.europa.eu/eurostat/documents/3859598/8762193/KS-02-18-240-EN-N.pdf/5fa53ed4-4367-41c4-b3f5-260ced9ff2f6> [accessed 2021-12-10]
28. Braekman E, Charafeddine R, Demarest S, Drieskens S, Tafforeau J, Van der Heyden J, et al. Is the European Health Interview Survey online yet? Response and net sample composition of a web-based data collection. *Eur J Public Health* 2020;30(3):567-573. [doi: [10.1093/eurpub/ckz206](https://doi.org/10.1093/eurpub/ckz206)] [Medline: [31697353](https://pubmed.ncbi.nlm.nih.gov/31697353/)]
29. Demarest S, Molenberghs G, Van der Heyden J, Gisle L, Van Oyen H, de Waleffe S, et al. Sample substitution can be an acceptable data-collection strategy: the case of the Belgian Health Interview Survey. *Int J Public Health* 2017;62(8):949-957. [doi: [10.1007/s00038-017-0976-3](https://doi.org/10.1007/s00038-017-0976-3)] [Medline: [28470413](https://pubmed.ncbi.nlm.nih.gov/28470413/)]
30. Beschikbaarheid van internet in het huishouden per gewest 2006-2018. Statbel. 2019. URL: <https://bestat.statbel.fgov.be/bestat/crosstable.xhtml?view=d136ea24-9d50-421e-871a-7b807d163236> [accessed 2021-12-10]
31. Peytcheva E, Groves R. Using variation in response rates of demographic subgroups as evidence of nonresponse bias in survey estimates. *J Off Stat* 2009;25(2):193-201.
32. Ekholm O, Gundgaard J, Rasmussen NKR, Hansen EH. The effect of health, socio-economic position, and mode of data collection on non-response in health interview surveys. *Scand J Public Health* 2010;38(7):699-706. [doi: [10.1177/1403494810382474](https://doi.org/10.1177/1403494810382474)] [Medline: [20851845](https://pubmed.ncbi.nlm.nih.gov/20851845/)]
33. Sturgis P, Williams J, Brunton-Smith I, Moore J. Fieldwork effort, response rate, and the distribution of survey outcomes: a multilevel meta-analysis. *Public Opin Q* 2017;81(2):523-542. [doi: [10.1093/poq/nfw055](https://doi.org/10.1093/poq/nfw055)]
34. Blumenberg C, Zugna D, Popovic M, Pizzi C, Barros AJD, Richiardi L. Questionnaire breakoff and item nonresponse in web-based questionnaires: multilevel analysis of person-level and item design factors in a birth cohort. *J Med Internet Res* 2018;20(12):e11046 [FREE Full text] [doi: [10.2196/11046](https://doi.org/10.2196/11046)] [Medline: [30530454](https://pubmed.ncbi.nlm.nih.gov/30530454/)]
35. Leenheer J, Scherpenzeel A. Does it pay off to include non-internet households in an internet panel? *Int J Internet Sci* 2013;8(1):17-29 [FREE Full text]
36. Sterrett D, Malato D, Benz J, Tompson T, English N. Assessing changes in coverage bias of web surveys in the United States. *Public Opin Q* 2017;81(S1):338-356. [doi: [10.1093/poq/nfx002](https://doi.org/10.1093/poq/nfx002)]
37. Couper MP, Kapteyn A, Schonlau M, Winter J. Noncoverage and nonresponse in an Internet survey. *Soc Sci Res* 2007;36(1):131-148. [doi: [10.1016/j.ssresearch.2005.10.002](https://doi.org/10.1016/j.ssresearch.2005.10.002)]
38. Durrant G, Steele F. Multilevel modelling of refusal and non-contact in household surveys: evidence from six UK Government surveys. *J R Stat Soc Ser A Stat Soc* 2009;172(2):361-381. [doi: [10.1111/j.1467-985x.2008.00565.x](https://doi.org/10.1111/j.1467-985x.2008.00565.x)]
39. 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. *Int J Public Health* 2014;59(1):197-206. [doi: [10.1007/s00038-013-0460-7](https://doi.org/10.1007/s00038-013-0460-7)] [Medline: [23619721](https://pubmed.ncbi.nlm.nih.gov/23619721/)]
40. ICT-gebruik bij individuen. Statbel. URL: <https://statbel.fgov.be/nl/themas/huishoudens/ict-gebruik-huishoudens#figures> [accessed 2021-12-10]
41. Gajic A, Cameron D, Hurley J. The cost-effectiveness of cash versus lottery incentives for a web-based, stated-preference community survey. *Eur J Health Econ* 2012;13(6):789-799. [doi: [10.1007/s10198-011-0332-0](https://doi.org/10.1007/s10198-011-0332-0)] [Medline: [21691841](https://pubmed.ncbi.nlm.nih.gov/21691841/)]



42. Ebert JF, Huibers L, Christensen B, Christensen MB. Paper- or web-based questionnaire invitations as a method for data collection: cross-sectional comparative study of differences in response rate, completeness of data, and financial cost. *J Med Internet Res* 2018 Jan 23;20(1):e24 [FREE Full text] [doi: [10.2196/jmir.8353](https://doi.org/10.2196/jmir.8353)] [Medline: [29362206](https://pubmed.ncbi.nlm.nih.gov/29362206/)]
43. Demarest S, Van der Heyden J, Charafeddine R, Tafforeau J, Van Oyen H, Van Hal G. Socio-economic differences in participation of households in a Belgian national health survey. *Eur J Public Health* 2013;23(6):981-985. [doi: [10.1093/eurpub/cks158](https://doi.org/10.1093/eurpub/cks158)] [Medline: [23183496](https://pubmed.ncbi.nlm.nih.gov/23183496/)]
44. Jensen HAR, Ekholm O, Davidsen M, Christensen AI. The Danish health and morbidity surveys: study design and participant characteristics. *BMC Med Res Methodol* 2019;19(1):91 [FREE Full text] [doi: [10.1186/s12874-019-0733-9](https://doi.org/10.1186/s12874-019-0733-9)] [Medline: [31053088](https://pubmed.ncbi.nlm.nih.gov/31053088/)]
45. Lange C, Finger JD, Allen J, Born S, Hoebel J, Kuhnert R, et al. Implementation of the European health interview survey (EHIS) into the German health update (GEDA). *Arch Public Health* 2017;75:40 [FREE Full text] [doi: [10.1186/s13690-017-0208-6](https://doi.org/10.1186/s13690-017-0208-6)] [Medline: [28936356](https://pubmed.ncbi.nlm.nih.gov/28936356/)]
46. Mauz E, Hoffmann R, Houben R, Krause L, Kamtsiuris P, Gößwald A. Mode equivalence of health indicators between data collection modes and mixed-mode survey designs in population-based health interview surveys for children and adolescents: methodological study. *J Med Internet Res* 2018;20(3):e64 [FREE Full text] [doi: [10.2196/jmir.7802](https://doi.org/10.2196/jmir.7802)] [Medline: [29506967](https://pubmed.ncbi.nlm.nih.gov/29506967/)]
47. Lynn P. Evaluating push-to-web methodology for mixed-mode surveys using address-based samples. *Surv Res Methods* 2020;14(1):19-30. [doi: [10.18148/SRM/2020.V14I1.7591](https://doi.org/10.18148/SRM/2020.V14I1.7591)]
48. Schouten B, Cobben F, Bethlehem J. Indicators for the representativeness of survey response. *Surv Methodol* 2009;35(1):110-113.

## Abbreviations

**BHIS:** Belgian Health Interview Survey  
**BHISF2F:** Belgian Health Interview Survey by F2F edition 2018  
**BHISWEB:** Belgian Health Interview Survey by Web  
**CAPI:** computer-assisted personal interview  
**EHIS:** European Health Interview Survey  
**Eurostat:** European Statistical Office  
**F2F:** face-to-face  
**OR:** odds ratio  
**Statbel:** Belgian Statistical Office

*Edited by R Kukafka, G Eysenbach; submitted 06.12.20; peer-reviewed by M Callegaro, M Jordan-Marsh; comments to author 16.03.21; revised version received 30.03.21; accepted 05.10.21; published 07.01.22*

*Please cite as:*

Braekman E, Demarest S, Charafeddine R, Drieskens S, Berete F, Gisle L, Van der Heyden J, Van Hal G  
*Unit Response and Costs in Web Versus Face-To-Face Data Collection: Comparison of Two Cross-sectional Health Surveys*  
*J Med Internet Res* 2022;24(1):e26299  
URL: <https://www.jmir.org/2022/1/e26299>  
doi: [10.2196/26299](https://doi.org/10.2196/26299)  
PMID:

©Elise Braekman, Stefaan Demarest, Rana Charafeddine, Sabine Drieskens, Finaba Berete, Lydia Gisle, Johan Van der Heyden, Guido Van Hal. Originally published in the *Journal of Medical Internet Research* (<https://www.jmir.org>), 07.01.2022. 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 work, first published in the *Journal of Medical Internet Research*, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.