

ORIGINAL ARTICLE

# A Web-based survey among adults aged 40–54 years was time effective and yielded stable response patterns

Korbinian Weigl<sup>a,b,c,\*,1</sup>, Kaja Tikk<sup>a,b,1</sup>, Michael Hoffmeister<sup>a</sup>, Enrico N. De Toni<sup>d</sup>, Jochen Hampe<sup>e</sup>, Frank Kolligs<sup>b,d,f</sup>, Stefanie J. Klug<sup>g,h</sup>, Ulrich Mansmann<sup>i</sup>, Daniel Nasseh<sup>i</sup>, Jovana Radlovic<sup>j,k,l</sup>, Matthias Schwab<sup>j,l,m,n</sup>, Dirk Schweigler<sup>g</sup>, Anna-Magdalena Stephan<sup>e</sup>, Hermann Brenner<sup>a,b,o</sup>

<sup>a</sup>Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany

<sup>b</sup>German Cancer Consortium (DKTK), DKFZ, Heidelberg, Germany

<sup>c</sup>Medical Faculty, University of Heidelberg, Heidelberg, Germany

<sup>d</sup>Department of Medicine II, University of Munich, Munich, Germany

<sup>e</sup>Gastroenterology and Hepatology, Medical Clinic I, University Clinic Dresden, Technical University, Dresden, Germany

<sup>f</sup>Department of Medicine—Gastroenterology, Hepatology and Infectiology, HELIOS Clinic Berlin-Buch, Berlin, Germany

<sup>g</sup>Cancer Epidemiology, University Cancer Center, TU Dresden, Dresden, Germany

<sup>h</sup>Department of Sport and Health Sciences, Epidemiology, Technical University of Munich, Munich, Germany

<sup>i</sup>Institute for Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-University Munich, Munich, Germany

<sup>j</sup>Dr. Margarete Fischer-Bosch Institute of Clinical Pharmacology, University of Tuebingen, Stuttgart, Germany

<sup>k</sup>University of Tuebingen, Tuebingen, Germany

<sup>l</sup>German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Partner site Tuebingen, Germany

<sup>m</sup>Department of Clinical Pharmacology, University Hospital, Tuebingen, Germany

<sup>n</sup>Department of Biochemistry and Pharmacy, University of Tuebingen, Tuebingen, Germany

<sup>o</sup>Division of Preventive Oncology, DKFZ and National Center of Tumor Diseases (NCT), Heidelberg, Germany

Accepted 31 August 2018; Published online 6 September 2018

## Abstract

**Objective:** We want to present information about response patterns obtained by Web-based survey in a large-scale epidemiological study.

**Study Design and Setting:** Within the RAPS (Risk Adapted Prevention Strategies for colorectal cancer [CRC]) study, we invited 160,000 randomly selected persons aged 40–54 years in three large German cities from 2015 to 2016 to complete a Web-based questionnaire on CRC risk factors and screening (97 items, average time for completion 15 minutes). Invitation letters and up to two reminder letters were sent to each individual.

**Results:** A total of 21.4% of women and 18.0% of men completed the questionnaire. Overall cumulative response rates were 7.5%, 14.3%, and 19.6% after the initial invitation letter, and the first and second reminder, respectively, with prevalence of and associations of key epidemiological parameters (such as family history of cancer, previous colonoscopy, etc.) being remarkably stable across waves of responses. For example, the sex and age distribution of the sample did not change with additional answers gained from additional letters.

**Conclusion:** Web-based questionnaires are feasible, cost-effective, and time effective in the setting of large-scale epidemiological studies. Although response patterns were remarkably stable over several rounds of reminders with substantially increasing cumulative response rates, future research should address possibilities to further enhance response rates. © 2018 Elsevier Inc. All rights reserved.

**Keywords:** Response rates; Web-based survey; Epidemiology; Large-scale study; Online questionnaire; Colorectal cancer

**Funding:** This work was supported by the German Cancer Consortium (DKTK). J.R. and M.S. in part were supported by the Robert Bosch Stiftung, Stuttgart. The funding sources had no involvement in the study design; the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

**Conflict of interest:** None.

**Transcript profiling:** None.

**Writing assistance:** None.

<sup>1</sup> Authors share the first authorship and contributed equally to the preparation of the manuscript.

\* Corresponding author. Tel.: 0049 6221 42 1340; fax: 0049 6221 42 1302.

*E-mail address:* [k.weigl@dkfz.de](mailto:k.weigl@dkfz.de) (K. Weigl).

**What is new?****Key findings**

- Online questionnaires are an efficient tool for obtaining information in large-scale epidemiological studies.
- Reminder letters are worthwhile and may strongly increase cumulative response rates.
- Response rates varied by geographic location, sex, and age.
- Online questionnaires help speed up post-data-obtainment procedures such as data cleansing with built-in plausibility checks and lead to quicker provision of data for analysis.

**What this adds to what was known?**

- Online questionnaires do not only work when used as follow-up questionnaires but also when applied with initial contact.
- Repeat reminder letters substantially increased absolute response numbers while being cost-efficient.

**What is the implication and what should change now?**

- Epidemiological researchers should consider implementing online questionnaires as tools for data collection in studies investigating young to middle-aged populations at average risk.
- At least one reminder should be sent to increase overall response rate.

**1. Introduction**

Epidemiological studies need large number of participants and reasonable response rates. In population-based studies, initial contacts and first responses have traditionally most commonly been made by written letters or phone. In the Internet era, these media are more and more replaced by online communication, but Web-based surveys have not yet been widely used in large-scale epidemiological studies [1].

Generally, online questionnaires seem to have a number of advantages compared to paper-based alternatives. For example, they have previously been found to be returned quicker than postal questionnaires [2,3]. Another advantage of Web-based questionnaires compared to paper-based surveys is the possibility for researchers to immediately adjust questions if necessary without having to discard already printed copies. Furthermore, preliminary results can normally be easily obtained [4]. Although online questionnaires were thought to yield more measurement error than traditional data collection [4,5], more recent research indicates that the contrary seems to be true [1]. A further advantage is that built-in plausibility checks may immediately alert participants of

missing or inconsistent or implausible responses, which may increase data quality and completeness and substantially reduce the load of time-consuming contacts with participants and data cleaning. Importantly, there is evidence that individuals responding to Web-based surveys are comparable to those responding to traditional modes of data collection regarding a large variety of important characteristics (e.g., age, gender, income, education, and health status) [6–8].

Yet very little is known about response rates and response patterns that may be achieved with this approach in times of almost complete Internet coverage. Here, we report on response rates and response patterns and their relationship to observed prevalences and associations in a Web-based survey on family history (FH) of colorectal cancer (CRC) and related factors relevant to CRC screening for which 160,000 men and women aged 40–54 years in three large German cities were invited.

**2. Methods**

This survey was part I of the RAPS (*Risk Adapted Prevention Strategies for CRC*) study, a multiphase multicenter study conducted in the framework of the German Cancer Consortium. The full description of the RAPS study has been published elsewhere [9]. Briefly, in part I of this study, we aimed for recruiting 30,000 men and women aged 40–54 years in the catchment areas of three German study centers located in the cities of Dresden, Munich, and Stuttgart (approximately 540,000, 1,450,000, and 620,000 inhabitants, respectively). One of the cities, Dresden, is located in the Eastern part of Germany, the other two cities are located in the Western part of Germany. Potential participants were randomly selected from population registries of the three cities (Table 1). With an anticipated response rate of 25%, the initial plan was to select and contact 120,000 participants (40,000 per center). However, because of lower-than-expected response rates, the number of invited persons was later increased to 60,000 in Munich and Stuttgart.

Personal invitation letters were sent through regular mail between June 2015 and December 2016. Participants were invited to take part in a Web-based survey focusing on FH of CRC and related factors relevant to CRC screening. They were informed that those with a history of CRC in a first-degree relative will be invited for individual consultation in our study centers in part II of the study, but participation in the survey was encouraged independent of a FH. All participants (independent of FH of CRC) received a Web-based feedback on their individual CRC risk factors immediately after completion of the questionnaire. Access to the Web-based survey via computer or smartphone was facilitated using a personal access code, as was the online declaration of informed consent. The invitation letters included a QR code, which could be scanned and which directed to the online questionnaire. Respondents were identified via the personal access code, therefore an

anonymous completion of the questionnaire was not possible. People who did not respond to the first invitation were contacted two more times with reminders, each sent on average 3 weeks after the previous invitation.

In total, the online questionnaire consisted of 97 questions on the participants' history of diseases, uptake of health examinations, FH of CRC, and environmental and lifestyle factors. Out of all questions, only 13 had to be answered by all participants, and another 65 questions were only visible after previous filter questions applied and were hence not to be answered by every participant. The questionnaire took on average 15 minutes to complete (time was estimated based on extensive pretests, and actual time of participants was not recorded). Built-in plausibility and completeness checks ensured high quality and completeness of the submitted data. The study protocol including all data collection procedures was approved by the ethics committee of the Medical Faculty of the University of Heidelberg; the responsible ethics committees of the respective partner sites in Dresden, Munich, and Stuttgart; and the data protection commissioner of the German Cancer Research Center.

### 2.1. Statistical analysis

First, the random samples obtained from the population registries were compared to the overall population of the respective city with respect to age and sex. Response rates were then calculated, stratified by sex, age groups, and invitation letter (i.e., initial invitation, first reminder, second reminder). The number of responses obtained through the different invitation letters was furthermore plotted on a time axis.

To assess if and to what extent the composition of the study sample changed after each additional wave of respondents with respect to main characteristics (sex, age, FH of CRC), the distribution of these characteristics was assessed among cumulative numbers of respondents after the initial invitation letter and each reminder. Similarly, the associations of these characteristics with having had a previous colonoscopy, which were quantified by odds ratios (obtained from multiple logistic regression models including all three

characteristics), were assessed after consecutive inclusion of each wave of respondents.

### 3. Results

Table 1 shows the characteristics of the random samples retrieved from the population registries, in comparison to the general population. Overall, the obtained data sets represented very good random samples of the underlying population with regard to their sex and age distribution. A total of 28.3% of all eligible persons in the population of the three cities were invited to take part in the study. This proportion varied from approximately 18% in Munich to approximately 46% in Stuttgart.

Overall, 28,964 of 160,000 invited persons participated in the Web-based survey, which corresponds to an overall response rate of 18.1% (Table 2). A total of 7.8% of all invitations sent to the potential participants was not deliverable (e.g., because people had moved to a different place of living) and were returned to the coordinating study center in Heidelberg; this rate of undeliverable invitations was higher in male compared to female recipients (9.3% vs. 6.1%) and differed between the cities. The undeliverable fraction of invitations was almost twice as high among men living in Munich (11.3%) compared to men living in Dresden (5.8%). Furthermore, the proportion of undeliverable letters was higher when they were addressed to younger persons compared to older persons. Subtracting the proportion of undeliverable invitations from the total amount of invitations, the response rate among those who received the invitation was 19.6%. Another 0.8% of all online questionnaires were considered as not having been completed by the invited person, as the answers provided online showed a mismatch with the data obtained from registry offices with regard to age and/or sex, and were excluded from the data analysis.

The response rate differed among the three cities with the highest rate observed in Dresden (25.9%), whereas participation in Munich and Stuttgart was lower (17.3%

Table 1. Random sample drawings

Age group	Dresden						Munich			
	General population <sup>a</sup>		Random sample		Respondents		General population <sup>b</sup>		Random sample	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
40–44	14,359 (46.3%)	16,678 (53.7%)	5,708 (46.5%)	6,578 (53.5%)	1,373 (51.1%)	1,313 (48.9%)	52,159 (47.2%)	58,425 (52.8%)	9,117 (47.4%)	10,116 (52.6%)
45–49	16,189 (46.8%)	18,413 (53.2%)	6,453 (47.1%)	7,234 (52.9%)	1,717 (53.1%)	1,517 (46.9%)	55,782 (46.9%)	63,163 (53.1%)	10,157 (46.7%)	11,581 (53.3%)
50–54	17,693 (49.3%)	18,169 (50.7%)	6,950 (49.5%)	7,077 (50.5%)	2,091 (54.9%)	1,718 (45.1%)	50,662 (49.0%)	52,798 (51.0%)	9,213 (48.4%)	9,816 (51.6%)
Total	48,241 (47.5%)	53,260 (52.5%)	19,111 (47.8%)	20,889 (52.2%)	5,181 (53.3%)	4,548 (46.8%)	158,603 (47.6%)	174,386 (52.4%)	28,487 (47.5%)	31,513 (52.5%)

Values are expressed as *n* (%).

<sup>a</sup> As of 2015.

<sup>b</sup> As of 2014.

**Table 2.** Response rate of online questionnaire in %

Category	Dresden			Munich			Stuttgart			Overall		
	Total	Female	Male	Total	Female	Male	Total	Female	Male	Total	Female	Male
Overall <sup>a</sup>	24.5	27.3	22.0	15.6	17.3	14.1	16.3	17.9	14.8	18.1	20.0	16.3
Undeliverable	5.1	4.4	5.8	9.5	7.5	11.3	7.9	5.9	9.7	7.8	6.1	9.3
Overall <sup>b</sup>	25.9	28.6	23.3	17.3	18.7	16.0	17.7	19.0	16.4	19.6	21.4	18.0
In persons aged 40–44 <sup>b</sup>	24.4	26.6	22.5	16.7	18.1	15.5	16.7	17.7	15.9	18.7	20.0	17.4
In persons aged 45–49 <sup>b</sup>	25.3	28.4	22.6	17.6	18.8	16.5	17.3	18.3	16.4	19.4	21.0	18.0
In persons aged 50–54 <sup>b</sup>	27.6	30.3	24.8	17.5	19.2	15.7	18.9	20.9	17.0	20.7	22.9	18.6
After initial invitation letter <sup>b</sup>	10.1	11.5	8.8	6.7	7.5	6.0	6.5	7.0	6.1	7.5	8.3	6.7
After first reminder <sup>b</sup>	9.3	10.6	8.2	5.7	6.2	5.2	6.1	6.5	5.7	6.8	7.4	6.2
After second reminder <sup>b</sup>	6.6	6.6	6.4	4.9	5.0	4.8	5.1	5.5	4.7	5.4	5.6	5.2

Values are expressed in %.

<sup>a</sup> Response rate with regard to originally drawn random sample of 160,000 persons.

<sup>b</sup> Response rate in relation to deliverable letters.

and 17.7%, respectively). The response rate in total was lowest for persons aged 40–44 years (18.7%) and increased slightly with increasing age (19.4% and 20.7% for age groups 45–49 years and >50 years, respectively).

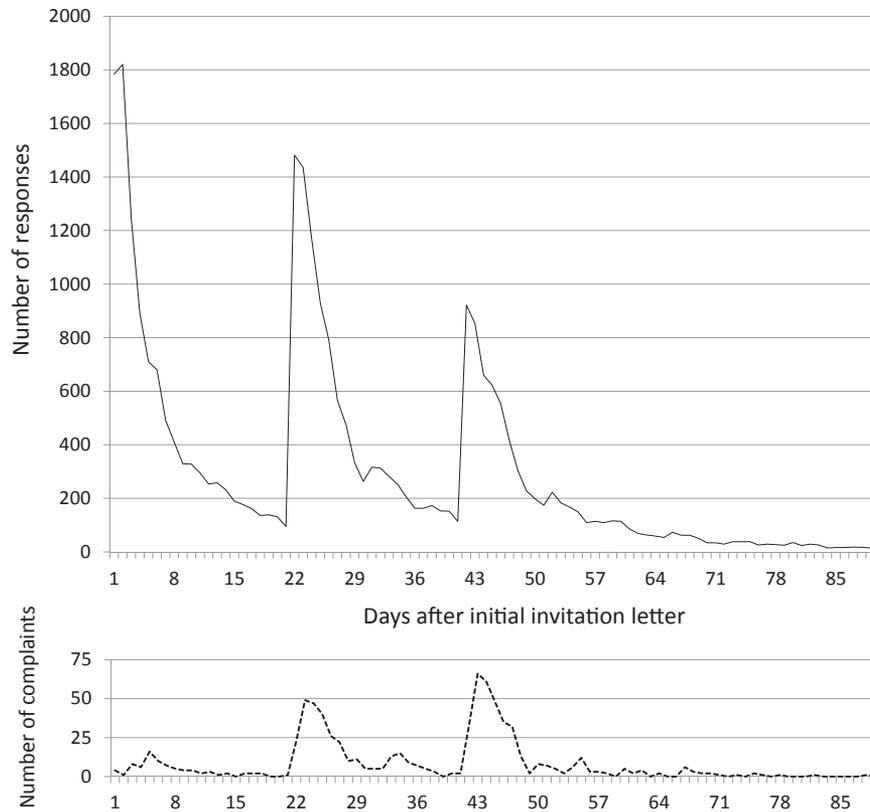
The proportion of respondents with non-German nationality was distinctly smaller than in the underlying populations (Dresden: 2.3% vs. 5.7%, Munich: 12.8% vs. 28.6%, Stuttgart: 10.6% vs. 23.6%).

A maximum of three individual letters were sent to each potential participant (initial invitation letter and in case of no reply a maximum of two reminders). Although the observed response rate of 7.5% to the initial invitation was rather low, this substantially increased with each round of reminder leading to a cumulative response of 14.3% and 19.6% after the first and second reminder, respectively, that is, each reminder prompted almost as many (6.8% and 5.4%) responses as the initial invitation letter. A higher response rate in women compared to men could be observed throughout all cities, all age groups, and all waves of invitations. The online questionnaire was accessed but not completed by additional 1.1% ( $n = 1,814$ ) of all invited persons. A hard copy of the questionnaire was requested by

0.5% ( $n = 845$ ) of all invited persons. Of those, 62% ( $n = 522$ ) completed the questionnaire, constituting approximately 1.8% of all completed questionnaires in total. The proportion of non-German citizens was higher in persons requesting a hard copy than in participants of the online questionnaire (13.8% vs. 8.4%). Respondents who requested a hard copy of the questionnaire were further more likely to have a FH of CRC (12.6% vs. 9.4%) compared to respondents of the online questionnaire.

The number of online questionnaire responses according to the time elapsed after the initial invitation letters are shown by the solid line in the upper panel of Fig. 1. The peaks on days 22 and 42 represent the time points when the participants had received the reminder letters. As expected, the response rate decreased steadily with every additional day and became very small approximately 3 weeks after the dispatch of each letter. Fig. 1 also shows the number of complaints (by phone calls to the study centers) according to time from the initial invitation letter (bottom area with the dotted line). Altogether, the study centers received 778 complaints (i.e., from about 0.5% of invited persons). Like the responses, the complaints arrived in waves after the delivery of invitation letters, but their

Munich		Stuttgart					
Respondents		General population <sup>b</sup>		Random sample		Respondents	
Female	Male	Female	Male	Female	Male	Female	Male
1,495 (52.5%)	1,354 (47.5%)	20,349 (47.8%)	22,202 (52.2%)	9,125 (47.1%)	10,237 (52.9%)	1,466 (50.7%)	1,427 (49.3%)
1,748 (51.1%)	1,670 (48.9%)	21,976 (47.4%)	24,118 (52.3%)	9,645 (47.3%)	10,737 (52.7%)	1,654 (51.3%)	1,573 (48.7%)
1,649 (54.1%)	1,398 (45.9%)	20,808 (48.3%)	22,293 (51.7%)	9,731 (48.0%)	10,524 (52.0%)	1,932 (54.5%)	1,616 (45.5%)
4,892 (52.5%)	4,422 (47.5%)	63,133 (47.9%)	68,613 (52.1%)	28,502 (47.5%)	31,498 (52.5%)	5,052 (52.3%)	4,616 (47.7%)



**Fig. 1.** Number of online questionnaire responses and number of complaints according to days after initial invitation letter. While the number of days between the letters varied, the first reminder was most often sent 22 days after the initial invitation letter and the second reminder another 20 days after the first reminder. For simplicity reasons, we neglected answers given thereafter for this figure. Please note the different scales in the y-axis of the two graphs.

numbers increased with every subsequent letter. Those who complained primarily requested the deletion of personal data or expressed data protection concerns and doubts about the legitimacy of the study. Of all complaints, about 30% were filed in Dresden, about 23% in Munich, and 48% in Stuttgart.

Main characteristics of study participants according to response waves are shown in Table 3. The overall distribution of the sample according to gender and age did not change with additional answers gained from each

additional letter. The proportion of participants who stated to have a FH of CRC (mentioned as the main topic of the study in all invitation letters) was 11% in the initial response rate and decreased only slightly with each reminder (9.7% after first reminder, 9.4% after second reminder). The proportion of participants who stated to already have undergone a previous colonoscopy remained almost constant across response waves (29.6% after the initial invitation, 28.6% and 28.4% after the first and second reminders, respectively).

**Table 3.** Main characteristics of study participants according to invitation waves

Characteristic	Group	Initial invitation	+ 1st reminder	+ 2nd reminder
		n (%)	n (%)	n (%)
Sex	Female	5,900 (53.8)	11,163 (53.5)	15,125 (52.7)
	Male	5,059 (46.2)	9,702 (46.5)	13,586 (47.3)
Age group	40–44 years	3,241 (29.6)	6,125 (29.4)	8,428 (29.4)
	45–49 years	3,832 (35.0)	7,215 (34.6)	9,879 (34.4)
	50–54 years	3,886 (35.5)	7,525 (36.1)	10,404 (36.2)
Family history in a first-degree relative		1,212 (11.0)	2,029 (9.7)	2,705 (9.4)
Previous colonoscopy	Total	3,248 (29.6)	5,968 (28.6)	8,148 (28.4)
	Without family history	2,560 (26.3)	4,841 (25.7)	6,674 (25.7)
	With family history	688 (56.7)	1,127 (55.5)	1,474 (54.5)

Table 4 shows the associations of the main study population characteristics with having undergone a previous colonoscopy and their development across the various response waves. Age and having a FH of CRC were strongly associated with having had a previous colonoscopy with odds ratios again remaining almost constant across the various response waves.

#### 4. Discussion

The present study gives an overview on the response rate of an online questionnaire used in the RAPS study, a large population-based epidemiological study. One invitation letter and up to two subsequent reminder letters resulted in a cumulative response rate of approximately 20% in a population-based setting, with higher response rates among women and residents from a city located in Eastern Germany compared to two cities located in Western Germany. These response rates were achieved without any follow-up of nonresponders through phone. The method proved to be immensely time-saving and cost-saving in the invited population compared to surveys requesting the return of paper-based questionnaires.

This study included several novel approaches with respect to the organization and implementation of large-scale population-based epidemiological studies. First, the questionnaire targeted to assess CRC risk factors and factors related to CRC screening in a large population-based sample was designed on an online platform. The overall response rate of approximately 20% seems low. However, as to our knowledge, no study has previously evaluated response rates in a large-scale epidemiological study using an online platform at first contact, our response rate is difficult to evaluate. Existing research on that topic concluded that response rates in surveys are generally decreasing [10] and are lower for Internet vs. other modes [11–14]. One previous study reported an impressive response rate for their online questionnaire of 41% (with additional 31% of invited participants completing the paper-based alternative) [15]. However, the online questionnaire was implemented for a follow-up survey to an already existing study and not for first contact. Another study offered both an online questionnaire and a paper-based version to their targeted population of military personnel and achieved an overall response rate of 37% [6]. Of these, approximately 55% used the online questionnaire (~20.2% of all invited persons), but it remains uncertain how many persons would have participated via online questionnaire if a paper-based version would not have been made available. A study that used a similar approach to ours to midwives in Australia and New Zealand yielded a response rate of less than 5%, but the authors reported difficulties with the various responsible councils regarding recruitment [16].

We expected higher participation rates among younger persons compared to older ones, but indeed, the opposite

was true. Most likely, although younger participants are probably more computer literate due to higher habituation, the topic of the questionnaire (i.e., CRC and FH thereof) may have been more present and hence more interesting for older persons. It can therefore be assumed that the obtained response rates in the different age strata are at least to some degree linked to the survey content. The low proportion of participants of non-German nationality compared to the underlying populations in the three cities might be to a large extent explained by language barriers because the invitation letters and the online questionnaire itself were written exclusively in German.

Although our approach was very efficient in recruiting very large number of participants with very limited resources, the potential of selection bias requires careful consideration due to the overall response rate of approximately 20%, which is lower than the response rates considered satisfactory in most conventional epidemiological surveys. In particular, it is unclear to what extent use of an online questionnaire itself might be a source of selection bias. One previous study found that participants completing an online questionnaire were younger, higher educated, and more supportive of research compared with participants from a telephone survey [17]. However, this study, as well as most other studies examining Internet vs. paper-based questionnaires, was performed more than 10 years ago and it can be assumed that Internet access and media literacy have significantly improved in the meantime. Specifically addressing our main research question, there seems to be no obvious reason why use of online questionnaires in general might differ between persons with and without FH of CRC.

A more plausible source of selection bias could though be that participants with a FH of CRC might be more willing to participate in a study on this topic, regardless of the type of data collection. In line with this assumption, the proportion of study participants with a FH decreased with each wave of respondents. However, the change in the overall proportion of participants with a FH was very modest and even mostly negligible when including the third wave of respondents after the second reminder. The same was true for the associations of age, sex, and FH with a previous colonoscopy, which hardly changed across consecutive rounds of additional recruitments. The high proportion of participants who have undergone colonoscopy is thereby fully in line with previous estimates in the German population [18]. Of note, although the response (i.e., completed questionnaires) after the initial invitation was relatively low (i.e., 7.5%), the resulting sample already resembled the study population obtained after two more reminders with respect to main characteristics and their associations with colonoscopy use which underlines the similarity of the composition of the study population throughout the invitation process. Although these patterns do not support the existence of major selection bias, some overrepresentation of more

**Table 4.** Associations of study population characteristics with previous colonoscopies

Study population characteristics	Initial invitation			
	Previous colonoscopy			
	Yes, <i>N</i> (%)	No, <i>N</i> (%)	OR <sup>a</sup>	95% CI
Sex				
Female	1,754 (29.7)	4,146 (70.3)	1.00	Ref.
Male	1,494 (29.5)	3,565 (70.5)	1.01	0.93, 1.10
Age				
40–44 years	702 (21.7)	2,539 (78.3)	1.00	Ref.
45–49 years	1,127 (29.4)	2,705 (70.6)	1.47	1.32, 1.65
50–54 years	1,419 (36.5)	2,467 (63.5)	2.03	1.82, 2.26
FH				
No	2,560 (26.3)	7,187 (73.7)	1.00	Ref.
Yes	688 (56.8)	524 (43.2)	3.61	3.19, 4.09

Abbreviations: CI, confidence interval; FH, family history; OR, odds ratio; Ref., reference.

<sup>a</sup> All models adjusted for sex, age, and family history.

health conscious and of more literate participants still has to be assumed. This is for example indicated by the higher proportion of persons with a FH of CRC in respondents actively requesting a hard copy of the questionnaire and seems furthermore plausible considering the incentive offered (a personalized feedback on CRC risk factors). However, similar to studies using paper-based questionnaire only, it is impossible to measure with certainty how well the sample represents the underlying population.

Since the cumulative response rate substantially increased with each reminder, the question arises whether additional reminders would have been worthwhile. Although they most likely would have increased the overall (cumulative) response rate indeed, our data suggest that the number of complaints would also increase rapidly with each round of reminder. Given the minimal changes in the distribution of participants with respect to the main study characteristics and the associations of those study characteristics with previous use of colonoscopy, the gain in participation with further rounds of reminders, be it by letter or phone, would have to be carefully balanced against the additional costs and inconvenience.

Research on the cost-effectiveness of Web-based surveys is scarce, with one study claiming lower costs for postal surveys [19], and other studies having inconclusive results [20,21] or suggesting lower costs for Web-based or mixed-methods approaches [22–26]. It is reasonable to assume that financial savings for implementing Web-based rather than paper-based questionnaires increase with increasing study size. The expenses for each individual participating after the initial invitation letter (including printing costs and stamps) amounted to 5.71€ (6.12\$). Each additional reminder increased the cumulative costs per participant approximately by 31%, but overall costs per study participant still remained modest. In fact, printing

the questionnaire, sending it to every person, and covering return mail costs instead would not have been financially feasible for this study.

To increase time efficiency and cost-efficiency, reminder letters were delivered by e-Post, an electronic postal service that allows the automatic processing and simultaneous delivery of thousands of letters. A further major aspect of cost-effectiveness is avoidance of the need of data entry and minimization of data cleaning and plausibility checks of entered data as the latter can be implemented directly in the online database. This also tremendously speeds up availability of data for analyses after completion of data collection.

In our study, we paid utmost care with respect to data security in close collaboration with several data protection officers before launching the study, and we carefully communicated data security aspects to the invitees. Nevertheless, contacting 160,000 persons from the general population can and will evoke some unintended reactions. Altogether, we received approximately 780 complaints (e.g., regarding legitimate entitlement to draw random samples). This number represents only a very small fraction (0.5%) of the total number of persons whom we contacted. Including a statement with the legislative basis for drawing the random samples from the population registry offices in the invitation letters which we have done is likely to have been essential for keeping this proportion low.

In summary, Web-based data collection proved to be feasible and economic in this very large-scale epidemiological study. With further increasing Internet literacy, access, and use in most populations, this mode of data collection might become a standard economic tool especially for very large-scale epidemiological surveys. Further research should address possibilities to optimize appeal and acceptance of this mode of data collection and to maximize both data quality and response proportions.

+ 1st reminder				+ 2nd reminder			
Previous colonoscopy				Previous colonoscopy			
Yes, N (%)	No, N (%)	OR <sup>a</sup>	95% CI	Yes, N (%)	No, N (%)	OR <sup>a</sup>	95% CI
3,160 (28.3)	8,003 (72.7)	1.00	Ref.	4,267 (28.2)	10,858 (71.8)	1.00	Ref.
2,808 (28.9)	6,894 (72.1)	1.05	0.99, 1.12	3,881 (28.6)	9,705 (71.4)	1.04	0.98, 1.09
1,291 (21.1)	4,834 (78.9)	1.00	Ref.	1,740 (20.7)	6,688 (79.4)	1.00	Ref.
1,996 (27.7)	5,219 (72.3)	1.40	1.29, 1.52	2,739 (27.7)	7,140 (72.3)	1.45	1.35, 1.55
2,681 (35.6)	4,844 (64.4)	2.03	1.87, 2.19	3,669 (35.3)	6,735 (64.7)	2.05	1.91, 2.19
4,841 (25.7)	13,995 (74.3)	1.00	Ref.	6,674 (25.7)	19,332 (74.3)	1.00	Ref.
1,127 (55.5)	902 (44.5)	3.54	3.22, 3.89	1,474 (54.5)	1,231 (45.5)	3.39	3.13, 3.68

## Acknowledgments

The authors especially thank all study participants. The authors also thank Drs. Utz Benscheid, Isabel Lerch, Volker Herrmann, Bernd Rasemann, Katrin Kollaske, Ulrike Seifert, Miriam Hofmann, Helga Lex, Caroline Tetzner, and Bernd Borstel for their excellent technical assistance and their contributions in data collection, monitoring, and documentation and thank Kai Ravesloot for the help with programming the online questionnaire.

## References

- [1] van Gelder MM, Bretveld RW, Roelvelde N. Web-based questionnaires: the future in epidemiology? *Am J Epidemiol* 2010;172:1292–8.
- [2] Akl EA, Maroun N, Klocke RA, Montori V, Schunemann HJ. Electronic mail was not better than postal mail for surveying residents and faculty. *J Clin Epidemiol* 2005;58:425–9.
- [3] Kroth PJ, McPherson L, Leverence R, Pace W, Daniels E, Rhyne RL, et al. Combining web-based and mail surveys improves response rates: a PBRN study from PRIME net. *Ann Fam Med* 2009;7:245–8.
- [4] Wyatt JC. When to use web-based surveys. *J Am Med Inform Assoc* 2000;7:426–9.
- [5] Lozar Manfreda K, Batagelj Z, Vehovar V. Design of web survey questionnaires: three basic experiments. *J Comput Mediat Commun* 2002;7.
- [6] Smith B, Smith TC, Gray GC, Ryan MA. When epidemiology meets the Internet: Web-based surveys in the millennium cohort study. *Am J Epidemiol* 2007;166:1345–54.
- [7] Fleming CM, Bowden M. Web-based surveys as an alternative to traditional mail methods. *J Environ Manage* 2009;90:284–92.
- [8] West R, Gilsenan A, Coste F, Zhou X, Brouard R, Nonnemaker J, et al. The ATTEMPT cohort: a multi-national longitudinal study of predictors, patterns and consequences of smoking cessation; introduction and evaluation of internet recruitment and data collection methods. *Addiction* 2006;101:1352–61.
- [9] Tikk K, Weigl K, Hoffmeister M, Igel S, Schwab M, Hampe J, et al. Study protocol of the RaPS study: novel risk adapted prevention strategies for people with a family history of colorectal cancer. *BMC Cancer* 2018;18:720.
- [10] Galea S, Tracy M. Participation rates in epidemiologic studies. *Ann Epidemiol* 2007;17:643–53.
- [11] McCluskey S, Topping AE. Increasing response rates to lifestyle surveys: a pragmatic evidence review. *Perspect Public Health* 2011;131:89–94.
- [12] Leece P, Bhandari M, Sprague S, Swionkowski MF, Schemitsch EH, Tornetta P, et al. Internet versus mailed questionnaires: a controlled comparison (2). *J Med Internet Res* 2004;6:e39.
- [13] Balter KA, Balter O, Fondell E, Lagerros YT. Web-based and mailed questionnaires: a comparison of response rates and compliance. *Epidemiology* 2005;16:577–9.
- [14] Lozar Manfreda K, Bosnjak M, Berzelak J, Haas I, Vehovar RV. Web surveys versus other survey modes: a meta-analysis comparing response rates. *Int J Market Res* 2008;50:79–104.
- [15] Ekman A, Dickman PW, Klint A, Weiderpass E, Litton JE. Feasibility of using web-based questionnaires in large population-based epidemiological studies. *Eur J Epidemiol* 2006;21:103–11.
- [16] Turner C, Bain C, Schluter PJ, Yorkston E, Bogossian F, McClure R, et al. Cohort profile: the nurses and midwives e-cohort study—a novel electronic longitudinal study. *Int J Epidemiol* 2009;38:53–60.
- [17] Fricker S, Galesic M, Tourangeau R, Yan T. An experimental comparison of web and telephone surveys. *Public Opin Q* 2005;69:370–92.
- [18] Chen C, Stock C, Hoffmeister M, Brenner H. Public health impact of colonoscopy use on colorectal cancer mortality in Germany and the United States. *Gastrointest Endosc* 2018;87:213–221.e2.
- [19] 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.
- [20] Rodriguez HP, von Glaahn T, Rogers WH, Chang H, Fanjiang G, Safran DG. Evaluating patients' experiences with individual physicians: a randomized trial of mail, internet, and interactive voice response telephone administration of surveys. *Med Care* 2006;44:167–74.
- [21] Ekman A, Litton JE. New times, new needs; e-epidemiology. *Eur J Epidemiol* 2007;22:285–92.
- [22] Zuidgeest M, Hendriks M, Koopman L, Spreuwenberg P, Rademakers J. A comparison of a postal survey and mixed-mode survey using a questionnaire on patients' experiences with breast care. *J Med Internet Res* 2011;13:e68.

- [23] 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:e173.
- [24] Kypri K, Gallagher SJ, Cashell-Smith ML. An internet-based survey method for college student drinking research. *Drug Alcohol Depend* 2004;76:45–53.
- [25] McAlindon T, Formica M, Kabbara K, LaValley M, Lehmer M. Conducting clinical trials over the internet: feasibility study. *BMJ* 2003;327:484–7.
- [26] Adams J, White M. Health behaviours in people who respond to a web-based survey advertised on regional news media. *Eur J Public Health* 2008;18:335–8.