BMJ Open

Patient-reported experiences with general practitioners: a randomized study of mail and web-based approaches following a national survey

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-036533
Article Type:	Original research
Date Submitted by the Author:	19-Dec-2019
Complete List of Authors:	Iversen, Hilde; Norwegian Institute of Public Health, Division of Health Services Holmboe, Olaf; Norwegian Institute of Public Health, Division of Health Services Bjertnaes, Oyvind; Norwegian Institute of Public Health, Division of Health Services
Keywords:	PRIMARY CARE, STATISTICS & RESEARCH METHODS, Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Patient-reported experiences with general practitioners: a randomized study of mail and web-based approaches following a national survey

Authors:

Hilde Hestad Iversen[1] (hilde.hestad.iversen@fhi.no) +47 46400425) (Corresponding author)

Olaf Holmboe[1] (olaf.holmboe@fhi.no)

Oyvind Bjertnaes[1] (oyvindandresen.bjertnaes@fhi.no)

Affiliations:

[1]Division of Health Services, Norwegian Institute of Public Health, PO Box 4404 Nydalen, 0403 Oslo, Norway

Keywords: data collection, general practitioners, patient satisfaction, physicians, random allocation, surveys and questionnaires

Word count: 2,340

Abstract

Objective

The aim of this study was to compare response rates, background characteristics, data quality and main study results for a survey of patient experiences with general practitioners (GPs) administered by mail and web-based approaches.

Design

Cross-sectional survey.

Setting

GPs in Norway.

Participants

Patients of family physicians in Norway.

Intervention

Based on a three-stage sampling design, 6999 patients of GPs aged 16 or older were randomized to one of two survey administration protocols: Group A, who were mailed an invitation with both a pen-and-paper and electronic response option (n=4,999) and Group B, who received an email invitation with electronic response option (n=2,000).

Main outcome measures

Response rates, background characteristics, data quality and main study results.

Results

The response rate was markedly higher for the mail survey (42.6%) than for the web-based survey (18.3%). A few of the background variables differed significantly between the two groups, but the data quality and patient-reported experiences were similar.

Conclusions

The response rate was 2.3-fold higher for the mail survey than for the web-based survey, but the two protocols yielded similar results for patient-reported experiences.

Article Summary

Strengths and Limitations

- The establishment of a national register with electronic contact information opens up new possibilities regarding electronic and web-based surveys
- A purely electronic protocol has not previously been explored in the national program for monitoring and reporting on health-care quality using patient experience surveys in Norway
- The results should be applicable to health systems with similar infrastructures and digital maturities, and also to countries working to establish regional or national digital infrastructures
- Future research needs to focus on initiatives for increasing response rates in web-based protocols, including sending multiple reminders using a combination of emails, messages on mobile phones, and other available platforms
- This study included adults evaluating their GPs, and so the results might not be generalizable to other health-care settings

Introduction

Norway introduced the regular general practitioner scheme in 2001, in which every inhabitant was assigned to an individual GP. GPs in Norway play a key role in the provision of health care, and are often the first contact point of patients with health services for most medical problems.[1] In 2018, The Ministry of Health and Care Services decided to evaluate the GP scheme, and part of this evaluation comprised a national patient experience survey.

The Norwegian Institute of Public Health (NIPH) is responsible for performing national patient experience surveys in Norway. Norway has a national program for monitoring and reporting on the quality of health care using patient experience surveys. The purpose of this program is to measure user experiences with health care systematically, with the obtained data used as a basis for interventions aimed at improving the quality of health care, health-care management, patient choice, and public accountability. The standard data-collection procedure in the national surveys is postdischarge mail surveys, which include a pen-and-paper questionnaire and an option to answer electronically. The results from two previous randomized studies and also studies of survey-mode preferences in different patient populations indicate that there is a rather modestly developed web mode preference overall.[2–9] However, the potential advantages of lower costs and shorter data-collection times are important arguments for performing further research into web-based surveys. Also, the expansion of Internet access and use may have changed the potential of the Internet to be an effective way to conduct such surveys. In Norway, the establishment of a national register with electronic contact information opens up new possibilities regarding electronic and web-based surveys, but so far this register has not been exploited in our national patient experience surveys.

The aim of the current study was to compare the standard mail survey mode of data collection with web-based data collection in Norway. The sample was randomized to one of two survey administration protocols: patients in Group A were mailed an invitation with both pen-and-paper and electronic response options, while those in Group B received an email invitation with an electronic response option only (using email addresses obtained from the national register). The response rates, data quality, background characteristics, and main study results were compared between the two groups.

Methods

Data

The sample consisted of patients aged 16 years and older registered with a GP in November 2018. The sampling plan had a three-stage design and aimed to produce a nationally representative sample. First, regular GP practices were randomly selected after stratification by the number of GPs at the practices and the municipality types. Second, all of the GPs were included in the selected practices that had up to four GPs, while four of them were randomly selected in the practices that had five or more GPs. Third, we randomly selected 14 adult patients from the list of patients of each GP.

This study included a total of 6,999 patients. Patients were randomized to 1 of 2 survey administration protocols: 4,999 patients to the main sample (Group A) and 2,000 patients to a subsample (Group B) (Fig. 1). Patients in Group A were mailed an invitation with both pen-and-paper and electronic response options, while those in Group B received an email

invitation with an electronic response option only. Two reminders were sent to nonrespondents in both samples using the same contact mode as for the first invitation.

Background data about the patients were obtained from public registries, including on gender, age, the number of years on the patient list of a GP, and the number of consultations during the past 24 months. Email addresses were collected from the national register for contact information, which is operated by the Agency for Public Management and eGovernment.

The Data Protection Officer at the NIPH recommended that the study be approved, and it was formally approved by the research director of the division for health services at the NIPH. The Norwegian Directorate of Health approved the use of data about nonrespondents in the nonresponse analysis, except those of patients who withdrew themselves from the study.

Measures

The Norwegian PEQ-GP (Patient Experiences with General Practitioner Questionnaire) was applied. This instrument was developed and validated according to the standard scientific procedures of the national patient-reported experience program in Norway.[5,10] A national validation study identified five scales that covered important aspects of the GP service relating to accessibility, evaluations of the GP and auxiliary staff, cooperation between the GP and other services, and patient enablement. We included 17 additional items that were relevant for evaluating the GP scheme. The questionnaire used in the randomized study consisted of 47 questions on 6 pages. Thirty-seven questions addressed experiences with the GP service, while ten were background questions. Most of the questions related to the user-reported experiences were answered in a 5-point response format ranging from “not at all” to “to a very large degree.” An additional page was included to allow the respondents to write

comments relating to experiences with their GPs and wishes regarding future changes to the GP scheme.

Statistical analysis

Items were assessed for levels of missing data, ceiling effects, and internal consistency. The internal consistency reliability of the five scales was assessed using the item-total correlation and Cronbach's alpha. We set the cutoff criterion for ceiling effects to 50%; that is, an item was considered acceptable if fewer than 50% of the respondents chose the most-favorable response option.[11,12]

Differences in respondent characteristics between Group A and Group B were tested using Pearson chi-square tests for categorical variables and independent-samples *t*-tests for continuous variables. Differences between the two groups regarding patient-reported experiences were tested using *t*-tests.

All of the statistical analyses were performed using SPSS (version 25.0).

Approval

The study was approved by the Data Protection Officer at the Norwegian Institute of Public Health. Return of the questionnaire represented patient consent in the study, which is the standard procedure in all patient experience surveys conducted by the Norwegian Institute of Public Health.

Patient and public involvement

The survey was about patients experiences with health care. Patients were included in the development process of the instrument, to secure that the questionnaire included the most important topics for patients.

Results

The overall response rate was 42.6% in Group A and 18.3% in Group B (Table 1). Most of the respondents (70.9%) in Group A answered on paper. The initial response rate was around 10% lower for Group B than for group A, with the remaining difference being related to reduced effects of both the first and second reminders.

The levels of missing data, proportion of responses in the "not applicable" option, ceiling effects, and internal consistency for the items are presented in Table 2. The levels of missing data ranged from 1.6% to 18.7% in Group A, and from 0.0% to 17.1% in Group B. The proportions of responses in the "not applicable" category ranged from 3.0% to 29.4% in Group A, and from 1.6% to 31.9% in Group B, and were higher in Group A than in Group B for all items except for two on the enablement scale and the items on the coordination and cooperation scale. All scales and items were below the ceiling-effect criterion of 50% in Group A, but two items exceed the criterion in Group B: one about whether the GP takes the patient seriously (52.2%) and the other about whether the GP communicates in a way that the patient can understand (56.0%). Cronbach's alpha values were similar in the two groups for four of the five indicators, but was lower (and below the criterion of 0.7) for the accessibility indicator in Group B. The remaining Cronbach's alpha values were above 0.7.

Table 3 compares the background characteristics of the respondents in the two survey administration protocols. The respondent age, time since previous contact, and education level differed significantly between the two groups, whereas there were no significant differences in gender, number of years on the list of the GP, number of consultations, number of diagnosis codes during the past 24 months, number of unique diagnosis codes the past 24 months, self-perceived physical health, self-perceived mental health, long-standing health problems, or geographic origin. The proportion of patients aged 30–49 years was higher in Group B than in Group A, and while Group A contained a higher proportion of patients who were aged ≥67 years. The web respondents were more likely to report that they had been in contact with their GP during the previous month. There was a significant tendency for those who responded to the email invitation to have a higher education level than those who responded to the mailed invitation: 61.5% of those in Group B reported being educated to the university level, compared to 47.0% in Group A.

Differences in patient-reported experiences between the two groups were small, varying from only 0.3 (GP is competent) to 3.5 (waiting time for urgent appointments is acceptable) on a scale from 0 to 100 (Table 4). There were no significant differences in the 5 indicators between the 2 groups, and only 1 of the 24 items was significantly different: the score for the item about the helpfulness and competence of other employees at GP practices was significantly higher in Group A than in Group B (p=0.046).

Discussion

This study compared response rates, background characteristics, data quality, and main study results between two randomized data-collection groups in a national survey of patient

experiences with GPs. Patients in Group A were mailed an invitation with both pen-and-paper and electronic response options, while those in Group B received an email invitation with an electronic response option only. The response rate was 2.3-fold higher for the mail protocol than for the web-based protocol, but the patient-reported experiences were similar in the two groups.

The results are consistent with literature reports that mail surveys achieve higher response rates than electronic and web-based approaches.[2–9] The current study of patient experiences with GPs is the first to explore a purely electronic protocol in the national program for monitoring and reporting on health-care quality using patient experience surveys in Norway. Web-based surveys have many advantages, including direct links to survey sites, ease of distribution, ease of receiving responses, and lower costs, but a major concern is that they exclude people without an email address as well as those with poor access to the Internet. The existence of a national register in Norway with electronic contact information presents a major opportunity for large-scale surveys of patient experiences, but as many as 15% of the patients in the electronic arm lacked a valid email address in the national register. Furthermore, only 18% of the contacted sample responded. A recent CAHPS survey produced corresponding results when comparing protocols based on web responses via an email invitation and mail.[7]

The response rate alone is a poor predictor of nonresponse bias, and previous studies have failed to find a consistent association between response rates and sample representativeness.[13,14] However, low response rates threaten the legitimacy of surveys in both the clinical and public domains, and reduce the ability of surveys to identify important differences in patient-reported experiences between providers and over time.[2,3] Future

research needs to focus on effective initiatives for increasing response rates in web-based protocols, including sending multiple reminders using a combination of emails, messages on mobile phones, and other available platforms. For example, the national infrastructure in Norway provides the possibility for secure digital mailboxes for all Norwegian inhabitants, which could be utilized for contacting digitally active patients. The current study showed that a lower education level and higher age were associated with a mail preference. Combined with the low response rates achieved for web-based protocols in this and other studies, future representative and high-quality surveys should include the opportunity to answer on pen-and-paper questionnaires. This could be implemented in a mixed-mode design that provides respondents with the option to choose how they want to respond, making it possible for patients without Internet access or sufficient computer skills to also participate.

The effects of background characteristics reported in the literature are inconsistent.[2–9] The results from two previous randomized studies showed similar background characteristics for respondents in different randomized groups.[2,3] However, the respondents in those surveys were all contacted by mail. In the current study we found that respondents invited by email were younger, more educated, and more likely to have had more-recent contact with their GP. We found no significant intergroup differences in the remaining nine background variables. Future research should assess how the national infrastructure in Norway could be used to tailor the mode of data collection to different groups, such as by providing a range of data-collection modes from purely electronic strategies (for respondents with high education levels) to a mail-based mixed mode (to older respondents and those with low education levels). The present and previous studies have revealed that patient-reported experiences are quite similar for different data-collection modes, but obviously the effects of the fragmented data-collection strategies remain to be determined.

A limitation of this study is that it only included adults evaluating their GPs in Norway, and so the results might not be generalizable to other health-care settings and countries. In particular, the national infrastructure and the digital maturity of the population in Norway might differ from the characteristics of other countries. However, the results should be applicable to health systems with similar infrastructures and digital maturities, and also to countries working to establish regional or national digital infrastructures.

Conclusions

Administering a survey of patient experiences with GPs using a web-based protocol produced results that were very similar to those obtained using the standard mail-mode data-collection procedure that is used in the national surveys, but had a much lower response rate. Furthermore, respondents in the digital group were younger, more educated, and had more-recent experiences with their GPs. Web-based surveys are faster and cheaper than standard mail surveys, but their low response rates threaten their legitimacy. Initiatives to increase response rates for web-based data collection and strategies for tailoring data collection to different groups should be key elements in future research.

Funding

This research received grant from the Norwegian Ministry of Health and Care Services.

Competing interests

On behalf of all authors, the corresponding author states that there is no competing interests.

Author contributions

The Ministry of Health and Care Services initiated the study. H.H.I. planned the study in consultation with O.A.B. and O.H. H.H.I. performed the statistical analyses with O.A.B. and O.H., and drafted the manuscript. O.A.B. and O.H. participated in the planning process, critically revised the manuscript draft and approved the final version of the manuscript. H.H.I. was the project manager for the survey. All authors read and approved the final manuscript.

Acknowledgements

We thank Inger Opedal Paulsrud from the Knowledge Centre at the NIPH for her help in conducting the survey, including performing administrative and technical tasks for the data collection.

Availability of data and material

The data set generated and/or analysed during this study is not publicly available due to the need to protect personal data.

References

1. Eide TB, Straand J, Rosvold EO. Patients' and GPs' expectations regarding healthcare-seeking behaviour: a Norwegian comparative study. BJGP Open. 2018;2(4):bjgpopen18X101615.
2. Bjertnaes OA, Iversen HH. User-experience surveys with maternity services: a randomized comparison of two data collection models. Int J Qual Health Care. 2012;24(4):433–8.
3. Bjertnaes O, Iversen HH, Skrivarhaug T. A randomized comparison of three data collection models for the measurement of parent experiences with diabetes outpatient care. BMC Med Res Methodol. 2018;18 (1):95.
4. Holmboe O, Bjertnæs OA. Patient experiences with Norwegian hospitals in 2015. Results following a national survey. Oslo: Norwegian Knowledge Centre for the Health Services, PasOpp-report 147; 2016. ISSN 1890-1565.
5. Holmboe O, Danielsen K, Iversen HH. Development of a method for the measurement of patient experiences with regular general practitioners. Oslo: Norwegian Knowledge Centre for the Health Services, PasOpp-report 1; 2015. ISSN 1890-1565.
6. Danielsen K, Holmboe O, Hansen SL, Førland O, Bjertnæs OA. Patient experiences with Norwegian primary care out-of-hours clinics. Oslo: Results from a survey among the "Vakttårn" clinics. Norwegian Knowledge Centre for the Health Services, PasOpp-report 1; 2012. ISSN 1890-1565.
7. Fowler FJ Jr, Cosenza C, Cripps LA, Edgman-Levitan S, Cleary PD. The effect of administration mode on CAHPS survey response rates and results: a comparison of mail and web-based approaches. Health Serv Res. 2019;54(3):714–21.
8. Bergeson SC, Gray J, Ehrmantraut LA, Laibson T, Hays RD. Comparing web-based with mail survey administration of the Consumer Assessment of Healthcare Providers

and Systems (CAHPS(®)) Clinician and Group survey. Prim Health Care. 2013;3:1000132.

9. Rodriguez HP, von Glahn 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(2):167–74.
10. Holmboe O, Iversen HH, Danielsen K, Bjertnaes O. The Norwegian patient experiences with GP questionnaire (PEQ-GP): reliability and construct validity following a national survey. BMJ Open. 2017;7(9):e016644.
11. Bjertnaes OA, Lyngstad I, Malterud K, Garratt A. The Norwegian EUROPEP questionnaire for patient evaluation of general practice: data quality, reliability and construct validity. Fam Pract. 2011;28(3):342–9.
12. Ruiz MA, Pardo A, Rejas J, Soto J, Villasante F, Aranguren JL. Development and validation of the "Treatment Satisfaction with Medicines Questionnaire" (SATMED-Q). Value Health. 2008;11(5):913–26.
13. Groves RM, Peytcheva E. The impact of nonresponse rates on nonresponse bias: a meta-analysis. Public Opin Q. 2008;72(2):167–89.
14. Groves RM. Nonresponse rates and nonresponse bias in household surveys. Public Opin Q. 2006;70(5):646–75.

Figure 1: CONSORT flow diagram

Table 1: Respondents before and after each reminder in the two randomized groups, and final response rates.

	Group A (*n*=4,760)	Group B (*n*=1,694)
Respondents before reminder:		
Electronic, *n*	272	117
Paper, *n*	560	-
Response rate, %	17.5	6.9
Respondents after first reminder:		
Electronic, *n*	171	126
Paper, *n*	533	-
Increase in response rate, %	14.8	7.4
Respondents after second reminder:		
Electronic, *n*	148	67
Paper, *n*	345	-
Increase in response rate, %	10.4	4.0
Total:		
Electronic, *n* (%)	591 (29.1)	310
Paper, *n* (%)	1438 (70.9)	
Response rate, %	42.6	18.3

Table 2: Comparison of missing data, ceiling effects, and internal consistency between the two randomized groups.

Scale and item[a]	Group A				Group B			
	Missing data (%)	Not applicable (%)	Ceiling effects (%)	Cronbach's alpha/item-total correlation	Missing data (%)	Not applicable (%)	Ceiling effects (%)	Cronbach's alpha/item-total correlation
GP				**0.924**				**0.935**
Do you feel that your GP takes you seriously?	5.0	3.4	48.0	0.762	0.0	2.9	52.2	0.813
Do you feel that your GP spends enough time with you?	4.9	3.2	28.9	0.702	0.0	3.2	32.7	0.745
Do you feel that your GP talks to you in a way you understand?	5.1	3.0	48.9	0.737	0.3	2.3	56.0	0.735
Do you feel that your GP is professionally competent?	5.2	4.0	41.8	0.752	0.6	2.9	43.1	0.805
Do you feel that your GP shows interest in your situation?	5.3	3.5	39.5	0.818	0.6	2.3	39.9	0.853
Do you feel that your GP includes you as much as you would like in decisions concerning you?	5.7	7.7	37.2	0.769	0.3	6.5	41.9	0.803
Does your GP provide you with sufficient information about your health problems and their treatment?	5.6	7.3	32.7	0.811	0.3	6.1	34.1	0.835
Does your GP provide you with sufficient information about the use and side effects of medication?	5.4	16.3	21.0	0.631	0.0	16.5	24.3	0.655
Does your GP refer you to further examinations or a specialist when you feel you need it?	5.1	10.6	43.2	0.646	0.3	10.6	49.3	0.633
Organization and auxiliary staff				**0.868**				**0.851**
Do you feel that your GP's practice is well organized?	5.2	4.1	26.6	0.681	1.0	1.6	25.8	0.615
Do you feel the other employees are helpful and competent?	4.9	3.6	36.1	0.813	0.3	2.9	31.7	0.806
Are you treated with courtesy and respect at the reception?	4.8	3.2	40.9	0.752	0.3	2.6	39.5	0.750
Accessibility				**0.774**				**0.688**
Was the waiting time for your last urgent appointment acceptable?	18.7	-	36.2	0.631	17.1	-	33.5	0.525
Is this waiting time for appointments that are not urgent acceptable?	12.3	-	18.0	0.631	6.8	-	11.4	0.525
Enablement				**0.906**				**0.925**
Does contact with your GP make you better able to understand your health problems?	1.6	16.6	19.5	0.803	1.0	15.8	26.0	0.836
Does contact with your GP make you better able to cope with your health problems?	1.6	19.7	16.3	0.852	0.6	20.3	21.2	0.875
Does contact with your GP better help you to stay healthy?	1.6	19.8	15.1	0.786	0.3	20.6	21.6	0.833
Coordination and cooperation				**0.875**				**0.876**
Do you feel that your GP is good at coordinating the range of health services available to you?	5.9	26.9	28.1	0.779	0.6	30.6	34.7	0.790
Do you feel that your GP cooperates well with other services you need?	5.7	29.4	29.0	0.779	0.6	31.9	29.7	0.790

[a]All items were scored on a 5-point response scale ranging from 1 ("not at all") to 5 ("to a very large degree")

Table 3: Comparison of respondent characteristics between the two randomized groups.

	Group A	Group B	*p*[a]
Gender, female	55.9 (1135)	59.7 (185)	0.216
Age group			<0.001
16–19 years	2.4 (49)	2.3 (7)	
20–29 years	7.8 (158)	9.4 (29)	
30–49 years	23.8 (482)	37.4 (116)	
50–66 years	34.2 (694)	31.6 (98)	
≥67 years	31.8 (646)	19.4 (60)	
Time on the list of the GP			0.526
<1 year	9.4 (191)	8.4 (26)	
1–2 years	19.4 (392)	21.3 (66)	
3–4 years	14.5 (293)	16.1 (50)	
5–10 years	20.4 (414)	22.3 (69)	
≥11 years	36.3 (735)	31.9 (99)	
Number of consultations during past 12 months			0.672
0	9.3 (186)	10.1 (31)	
1	15.7 (314)	15.3 (47)	
2–5	55.7 (1114)	52.1 (160)	
6–12	16.2 (323)	19.2 (59)	
≥13	3.1 (63)	3.3 (10)	
Number of diagnosis codes during past 24 months	13.8±13.5	12.6±10.7	0.083
Number of unique diagnosis codes during past 24 months	4.7±3.2	4.6±2.8	0.510
Time since previous contact			0.042
<1 month	36.5 (716)	42.1 (128)	
1–3 months	32.0 (628)	23.7 (72)	
4–6 months	13.5 (266)	15.1 (46)	
7–12 months	9.7 (191)	8.9 (27)	
>12 months	8.3 (163)	10.2 (31)	
Self-perceived physical health			0.951
Very poor	1.3 (27)	1.6 (5)	
Quite poor	5.3 (108)	5.2 (16)	
Both poor and good	23.8 (481)	22.4 (69)	
Quite good	48.3 (975)	50.3 (155)	
Very good	21.2 (429)	20.5 (63)	
Self-perceived mental health			0.475
Very poor	1.1 (22)	1.9 (6)	
Quite poor	3.0 (60)	3.9 (12)	
Both poor and good	15.5 (313)	15.6 (48)	
Quite good	41.7 (842)	38.0 (117)	
Very good	38.7 (781)	40.6 (125)	
Long-standing health problems			0.625
0	35.7 (708)	37.5 (115)	
1	32.9 (653)	34.9 (107)	
2	19.3 (383)	16.9 (52)	
≥3	12.1 (241)	10.7 (33)	
Education level			<0.001
Elementary school	15.6 (309)	7.1 (22)	
High school	37.4 (740)	31.4 (97)	
University, 0–4 years	25.6 (505)	35.3 (109)	
University, >4 years	21.4 (422)	26.2 (81)	
Geographic origin			0.205
Norway	88.6 (1756)	89.9 (276)	
Asia (incl. Turkey), Africa, or Latin America	4.8 (95)	3.3 (10)	
Eastern Europe (all countries, independent of EU membership)	3.5 (70)	2.3 (7)	
Western Europe, North America, or Oceania	3.0 (60)	4.6 (14)	

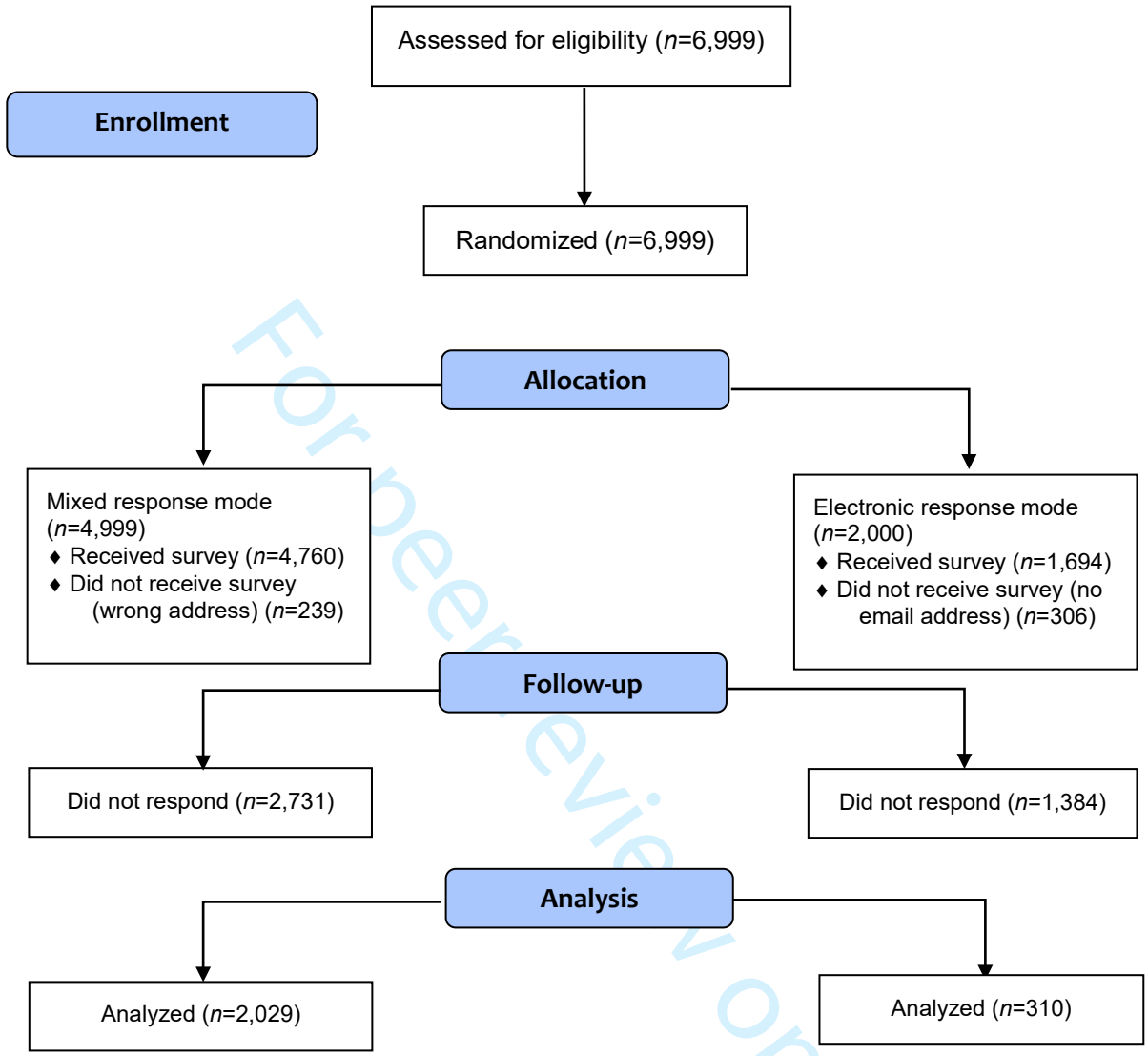
[a] Pearson chi-square tests for categorical variables and independent-samples *t*-tests for continuous variables. Data are *n* (%) or mean±SD values.

Table 4: Comparison of patient-reported experiences between the two randomized groups.

Scale and item[a]	Group A	Group B	*p*
GP	**78.3±16.8**	**78.8±17.8**	**0.651**
Do you feel that your GP takes you seriously?	83.1±19.8	84.1±20.3	0.429
Do you feel that your GP spends enough time with you?	73.8±22.8	73.3±25.8	0.707
Do you feel that your GP talks to you in a way you understand?	84.6±17.7	85.8±19.1	0.275
Do you feel that your GP is professionally competent?	82.1±18.0	81.8±19.4	0.747
Do you feel that your GP shows interest in your situation?	79.7±20.5	78.7±22.2	0.435
Do you feel that your GP includes you as much as you would like in decisions concerning you?	79.0±20.3	80.0±21.3	0.452
Does your GP provide you with sufficient information about your health problems and their treatment?	76.6±21.2	76.1±22.2	0.708
Does your GP provide you with sufficient information about the use and side effects of medication?	65.0±26.9	67.1±26.0	0.241
Does your GP refer you to further examination or a specialist when you feel you need it?	81.4±20.2	82.0±22.2	0.712
Organization and auxiliary staff	**78.2±17.7**	**77.1±18.2**	**0.322**
Do you feel that your GP's practice is well organized?	75.2±20.2	74.8±20.4	0.735
Do you feel the other employees are helpful and competent?	79.3±19.2	76.9±20.8	0.046
Are you treated with courtesy and respect at the reception?	80.9±19.7	79.9±20.7	0.430
Accessibility	**63.6±27.8**	**61.8±25.5**	**0.264**
Was the waiting time for your last urgent appointment acceptable?	69.5±30.6	69.1±30.4	0.828
Is this waiting time for appointments that are not urgent acceptable?	58.3±30.0	54.8±28.5	0.051
Enablement	**65.2±22.1**	**66.0±24.3**	**0.601**
Does contact with your GP make you better able to understand your health problems?	68.1±23.0	68.7±25.1	0.703
Does contact with your GP make you better able to cope with your health problems?	64.9±24.0	65.9±25.6	0.560
Does contact with your GP better help you to stay healthy?	62.7±25.0	64.5±26.7	0.303
Coordination and cooperation	**74.3±21.0**	**74.9±21.5**	**0.644**
Do you feel that your GP is good at coordinating the range of health services available to you?	75.0±21.2	77.7±20.5	0.079
Do you feel that your GP cooperates well with other services you need?	74.3±22.6	73.0±24.5	0.434

[a]All scales and items are scored from 0 to 100, where 100 is the best possible patient experience. Data are mean±SD values.

Figure 1: CONSORT flow diagram



Reporting checklist for cross sectional study.

Based on the STROBE cross sectional guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the STROBE cross sectionalreporting guidelines, and cite them as:

von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies.

		Reporting Item	Page Number
Title and abstract			
Title	#1a	Indicate the study's design with a commonly used term in the title or the abstract	2
Abstract	#1b	Provide in the abstract an informative and balanced summary of what was done and what was found	2,3
Introduction			
Background / rationale	#2	Explain the scientific background and rationale for the investigation being reported	5-6
Objectives	#3	State specific objectives, including any prespecified hypotheses	2,6
Methods			
Study design	#4	Present key elements of study design early in the paper	2, 6/7
Setting	#5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6-8
Eligibility criteria	#6a	Give the eligibility criteria, and the sources and methods of selection of participants.	2, 6-8

	#7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7,8
Data sources / measurement	#8	For each variable of interest give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. Give information separately for for exposed and unexposed groups if applicable.	7,8
Bias	#9	Describe any efforts to address potential sources of bias	11-13
Study size	#10	Explain how the study size was arrived at	6,7
Quantitative variables	#11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	6-8
Statistical methods	#12a	Describe all statistical methods, including those used to control for confounding	7-8
Statistical methods	#12b	Describe any methods used to examine subgroups and interactions	6-8
Statistical methods	#12c	Explain how missing data were addressed	7,8,19
Statistical methods	#12d	If applicable, describe analytical methods taking account of sampling strategy	6
Statistical methods	#12e	Describe any sensitivity analyses	n/a
Results			
Participants	#13a	Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed. Give information separately for for exposed and unexposed groups if applicable.	6,7,17
Participants	#13b	Give reasons for non-participation at each stage	17
Participants	#13c	Consider use of a flow diagram	17
Descriptive data	#14a	Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders. Give information separately for exposed and unexposed groups if applicable.	9,10,20
Descriptive data	#14b	Indicate number of participants with missing data for each variable of interest	19
Outcome data	#15	Report numbers of outcome events or summary measures. Give information separately for exposed and unexposed groups if applicable.	18,19,21
Main results	#16a	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	n/a

Main results	#16b	Report category boundaries when continuous variables were categorized	n/a
Main results	#16c	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	#17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	n/a
Discussion			
Key results	#18	Summarise key results with reference to study objectives	10,11
Limitations	#19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.	12,13
Interpretation	#20	Give a cautious overall interpretation considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence.	10-13
Generalisability	#21	Discuss the generalisability (external validity) of the study results	12,13
Other Information			
Funding	#22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	13

BMJ Open

Patient-reported experiences with general practitioners: a randomized study of mail and web-based approaches following a national survey

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-036533.R1
Article Type:	Original research
Date Submitted by the Author:	18-Jun-2020
Complete List of Authors:	Iversen, Hilde; Norwegian Institute of Public Health, Division of Health Services Holmboe, Olaf; Norwegian Institute of Public Health, Division of Health Services Bjertnaes, Oyvind; Norwegian Institute of Public Health, Division of Health Services
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Health services research
Keywords:	PRIMARY CARE, STATISTICS & RESEARCH METHODS, Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Patient-reported experiences with general practitioners: a randomized study of mail and web-based approaches following a national survey

Authors:

Hilde Hestad Iversen[1] (hilde.hestad.iversen@fhi.no) +47 46400425) (Corresponding author)

Olaf Holmboe[1] (olaf.holmboe@fhi.no)

Oyvind Bjertnaes[1] (oyvindandresen.bjertnaes@fhi.no)

Affiliations:

[1]Division of Health Services, Norwegian Institute of Public Health, PO Box 4404 Nydalen, 0403 Oslo, Norway

Keywords: data collection, general practitioners, patient satisfaction, physicians, random allocation, surveys and questionnaires

Word count: 3,854

Abstract

Objective

The standard data-collection procedure in the national patient experience survey programme in Norway is post-discharge mail surveys, which include a pen-and-paper questionnaire and an option to answer electronically. A purely electronic protocol has not previously been explored, and the aim of this study was to compare response rates, background characteristics, data quality and main study results for a survey of patient experiences with general practitioners (GPs) administered by the standard mail data-collection procedure and a web-based approach.

Design

Cross-sectional survey.

Setting

In Norway every inhabitant is assigned to an individual GP, and the setting for the current survey is patients of GPs. Regular GP practices were randomly selected after stratification by the number of GPs at the practices and the municipality types.

Participants

The sample consisted of 6,999 patients aged 16 years and older registered with a GP in November 2018.

Intervention

Based on a three-stage sampling design, 6,999 patients of GPs aged 16 or older were randomized to one of two survey administration protocols: Group A, who were mailed an invitation with both a pen-and-paper and electronic response option (n=4,999) and Group B, who received an email invitation with electronic response option (n=2,000).

Main outcome measures

Response rates, background characteristics, data quality and main study results.

Results

The response rate was markedly higher for the mail survey (42.6%) than for the web-based survey (18.3%). A few of the background variables differed significantly between the two groups, but the data quality and patient-reported experiences were similar.

Conclusions

Web-based surveys are faster and less expensive than standard mail surveys, but their low response rates and coverage problems threaten their usefulness and legitimacy. Initiatives to increase response rates for web-based data collection, more non-response research and strategies for tailoring data collection to different groups should be key elements in future research.

Article Summary

Strengths and Limitations

- The establishment of a national register with electronic contact information opens up new possibilities regarding electronic and web-based surveys
- A purely electronic protocol has not previously been explored in the national program for monitoring and reporting on health-care quality using patient experience surveys in Norway
- The results should be applicable to health systems with similar infrastructures and digital maturities, and also to countries working to establish regional or national digital infrastructures
- Future research needs to focus on initiatives for increasing response rates in web-based protocols, including sending multiple reminders using a combination of emails, messages on mobile phones, and other available platforms
- This study included adults evaluating their GPs, and so the results might not be generalizable to other health-care settings

Introduction

Norway introduced the regular General Practitioner (GP) scheme in 2001. All inhabitants who are registered in the National Registry as living in Norway have the right to a GP/family doctor. Migrants eligible to stay in Norway for more than six months are entitled to enrol in the scheme. GPs in Norway play a key role in the provision of health care, and are often the first contact point of patients with health services for most medical problems.[1] In 2018, The Ministry of Health and Care Services decided to evaluate the GP scheme, and part of this evaluation comprised a national patient experience survey.

The Norwegian Institute of Public Health (NIPH) is responsible for performing national patient experience surveys in Norway. Norway has a national program for monitoring and reporting on the quality of health care using patient experience surveys. The purpose of this program is to measure user experiences with health care systematically, with the obtained data used as a basis for interventions aimed at improving the quality of health care, health-care management, patient choice, and public accountability. The standard data-collection procedure in the national surveys is post-discharge mail surveys, which include a pen-and-paper questionnaire and an option to answer electronically.

The results from previous studies of survey-mode preferences in different patient populations both in Norway and other countries indicate that there is a rather modestly developed web-mode preference.[2–11] In the national patient experience survey among patients visiting general practitioners in 2014 in Norway, only 18% of respondents answered electronically.[4]

A main limitation of previous studies has been the lack of e-mail addresses in the sample frame, and the implication that even the electronic group had to be invited by a postal invitation, adding to costs, and precluding the possibility of testing a comprehensive electronic data collection option. A purely electronic protocol has not previously been explored in the national program for monitoring and reporting on health-care quality using patient experience surveys in Norway.

The literature on the effects of background characteristics on the responses to different data collection methods are inconsistent.[2–10] Non-response bias has been studied in four patient populations in Norway through follow-up telephone interviews with non-respondents, [12–15] including non-respondents in a survey on patient experiences with GPs [15] The results have shown minor differences between the postal respondents from the national surveys and the postal non-respondents who have provided answers through follow-up interviews. In general, the impact of non-response bias in the large-scale surveys have been considered relatively small.

The use of Internet in the population is growing. In 2018, 90% of all Norwegian citizens used the Internet at a daily basis. [16] In all age groups under 60 years, between 90-99 percent reported to use the Internet daily, but corresponding results for those between 60-69 years was 81% and for those aged 70 years or more 67%. Seventeen percent of the citizens aged 70 years or more reported that they never used the Internet. In Norway, the establishment of a national register with electronic contact information opens new possibilities regarding electronic and Hilde -based surveys. A total of 88% of the population was registered in the national register for contact information in November 2018.[17] So far, this register has not been exploited in our national patient experience surveys.

Potential variations in the population coverage between paper- and web-based questionnaires and the risk of selection bias from using the Internet for questionnaire surveys are reduced, but a major concern with protocols that use only digital responses is leaving out people without available digital contact information. When comparing the standard mail survey mode of data collection with web-based data collection the characteristics of non-respondents and respondents in both groups should be explored.

The potential advantages of lower costs and shorter data-collection times are important arguments for performing further research into web-based surveys. Also, the expansion of Internet access and use may have changed the potential of the Internet to be an effective way to conduct such surveys. We considered these potential advantages and possibilities as important arguments for performing further research into web-based surveys.

The aim of the current study was to compare the standard mail survey mode of data collection with web-based data collection in Norway. The sample was randomized to one of two survey administration protocols: patients in Group A were mailed an invitation with both pen-and-paper and electronic response options, while those in Group B received an email invitation with an electronic response option only (using email addresses obtained from the national register). The response rates, data quality, background characteristics, and main study results were compared between the two groups.

Methods

Data

The sample consisted of patients aged 16 years and older registered with a GP in November 2018. The preconditions for the sampling frame were to report the results on a national level and to be able to estimate intraclass correlation coefficients on the GP practice level. We did not aim to benchmark at the GP level. With the patient sample size chosen, we explored how the ICC varied dependent on the number of GPs at the practice level and found that at least four GPs were needed per GP practice to reach an acceptable ICC, and that not much were gained by including more GPs per practice. The sampling plan had a three-stage design. First, regular GP practices were randomly selected after stratification by the number of GPs at the practices and the municipality types. Second, all the GPs were included in the selected practices that had up to four GPs, while four of them were randomly selected in the practices that had five or more GPs. Third, we randomly selected 14 adult patients from the list of patients of each GP.

This study included a total of 6,999 patients. Patients were randomized to 1 of 2 survey administration protocols: 4,999 patients to the main sample (Group A) and 2,000 patients to a subsample (Group B) (Fig. 1). The current study was the first to explore a purely electronic protocol in the national program of patient experience surveys in Norway. Also, we have not previously explored the quality of the email addresses collected from the national register for contact information. Considering the commission of achieving national representative results and the uncertainty regarding the responses from a purely electronic protocol, we evaluated the risk of randomizing the total sample in two groups as too high and chose to include fewer patients in the subsample.

Patients in Group A were mailed an invitation with both pen-and-paper and electronic response options. The invitation included a cover letter describing the purpose of the study, a paper questionnaire, a prepaid envelope and information and a login code to be able to respond electronically. The patients in Group B received an email invitation with an electronic response option only. The email invitation included information about the purpose of the study, a link to the online survey and a login code. Two reminders were sent to non-respondents in both samples using the same contact mode as for the first invitation. The first reminder was sent to both groups around three weeks after the first contact. The second reminder was sent around six weeks after the first contact. All reminders to Group A were sent by mail and included a new invitation, the paper questionnaire, the postage-paid envelope and the login code to enable electronic responses. Group B were sent a new email invitation with a link to the survey and a login code in both reminders.

Background data about the patients were obtained from public registries, including on gender, age, the number of years on the patient list of a GP, and the number of consultations during the past 24 months. Email addresses were collected from the national register for contact information, which is operated by the Agency for Public Management and eGovernment.

Measures

The Norwegian PEQ-GP (Patient Experiences with General Practitioner Questionnaire) was applied. This instrument was developed and validated according to the standard scientific procedures of the national patient-reported experience program in Norway.[5,10]

A national validation study identified five scales that covered important aspects of the GP service relating to accessibility, evaluations of the GP and auxiliary staff, cooperation

between the GP and other services, and patient enablement. We included 17 additional items that were relevant for evaluating the GP scheme. The questionnaire used in the randomized study consisted of 47 questions on 6 pages. Thirty-seven questions addressed experiences with the GP service, while ten were background questions. Most of the questions related to the user-reported experiences were answered in a 5-point response format ranging from "not at all" to "to a very large degree. Single item and index scores were transformed linearly from the 1 to 5 scale to a scale of 0–100. An additional page was included to allow the respondents to write comments relating to experiences with their GPs and wishes regarding future changes to the GP scheme.

Patient and Public Involvement

Patients were included in the development process of the instrument, to secure that the questionnaire included the most important topics for patients. To identify important topics, we assessed reviews of the literature and consulted a reference group comprising GPs, researchers and representatives from health authorities and patient organisations throughout the process of questionnaire development. The questionnaire was tested through cognitive interviews with patients. First, eight face-to-face interviews and nine telephone interviews were conducted. After an extensive revision, we conducted another 11 face-to-face interviews with patients. The revised version was tested in a pilot study.

Statistical analysis

The survey response rate by group was calculated as the proportion of eligible patients (ie, not those who had moved to a new house, died, or were otherwise ineligible) and who returned a completed survey (AAPOR response rate 4.0).[18]

Items were assessed for levels of missing data, ceiling effects, and internal consistency. The internal consistency reliability of the five scales was assessed using the item-total correlation and Cronbach's alpha. The item-total correlation coefficient quantifies the strength of an association between an item and the remainder of its indicator, with a coefficient of 0.4 considered acceptable. [19] Cronbach's alpha assesses the overall correlation between items within an indicator, and an alpha value of 0.7 is considered satisfactory.[19,20] We set the cut-off criterion for ceiling effects to 50%; that is, an item was considered acceptable if fewer than 50% of the respondents chose the most-favourable response option.[21,22]

Differences in respondent characteristics between Group A and Group B were tested using Pearson chi-square tests for categorical variables and independent-samples *t*-tests for continuous variables. Differences between the two groups regarding patient-reported experiences were tested using *t*-tests.

Differences in respondent characteristics between respondents and non-respondents in Group A and respondents and non-respondents in Group B were tested using Pearson chi-square tests for categorical variables and independent-samples t-tests for continuous variables. Variables available on non-respondents were gender, age, time on the list of the GP, number of consultations during the past 24 months and number of diagnosis during the past 24 months.

All the statistical analyses were performed using SPSS (version 25.0).

Approval

The Data Protection Officer at the NIPH recommended that the study be approved, and it was formally approved by the research director of the division for health services at the NIPH. The Norwegian Directorate of Health approved the use of data about non-respondents in the nonresponse analysis, except those of patients who withdrew themselves from the study. Return of the questionnaire represented patient consent in the study, which is the standard procedure in all patient experience surveys conducted by the Norwegian Institute of Public Health.

Results

The overall response rate was 42.6% in Group A and 18.3% in Group B (Table 1). 15% of the patients in the electronic arm lacked a valid email address in the national register, and 5% of the patients in the standard mail survey mode lacked a valid mailing address (Fig. 1). Most of the respondents (70.9%) in Group A answered on paper (Table 1). The initial response rate was around 10% lower for Group B than for group A, with the remaining difference being related to reduced effects of both the first and second reminders.

The levels of missing data, proportion of responses in the "not applicable" option, ceiling effects, and internal consistency for the items are presented in Table 2. The levels of missing data ranged from 1.6% to 18.7% in Group A, and from 0.0% to 17.1% in Group B. The proportions of responses in the "not applicable" category ranged from 3.0% to 29.4% in Group A, and from 1.6% to 31.9% in Group B, and were higher in Group A than in Group B for all items except for two on the enablement scale and the items on the coordination and cooperation scale. All scales and items were below the ceiling-effect criterion of 50% in

Group A, but two items exceed the criterion in Group B: one about whether the GP takes the patient seriously (52.2%) and the other about whether the GP communicates in a way that the patient can understand (56.0%). Cronbach's alpha values were similar in the two groups for four of the five indicators, but was lower (and below the criterion of 0.7) for the accessibility indicator in Group B. The remaining Cronbach's alpha values were above 0.7.

Table 3 compares the background characteristics of the respondents in the two survey administration protocols. The respondent age, time since previous contact, and education level differed significantly between the two groups, whereas there were no significant differences in gender, number of years on the list of the GP, number of consultations, number of diagnosis codes during the past 24 months, number of unique diagnosis codes the past 24 months, self-perceived physical health, self-perceived mental health, long-standing health problems, or geographic origin. The proportion of patients aged 30–49 years was higher in Group B than in Group A (37.4% compared to 23.8%). In group A, 31.8% of the patients were aged ≥67 years, a much higher proportion than in group B where the corresponding proportion was 19.4%. The respondents in Group B were more likely to report that they had been in contact with their GP during the previous month than respondents in Group A. There was a significant tendency for those who responded to the email invitation to have a higher education level than those who responded to the mailed invitation: 61.5% of those in Group B reported being educated to the university level, compared to 47.0% in Group A.

Significant differences were found between Group A and Group B within respondents and non-respondents with respect to gender and age (Table 4). Non-respondents tended to be more likely to be men and to be younger than respondents in both groups. Significant differences were also found for time on the list of the GP, number of consultations during the past 24

months and the two variables about number of diagnosis the last two years for Group A. Respondents tended to have been longer on the GPs list, and to have a higher number of consultations and diagnosis during the last two years. We found no additional differences between respondents and non-respondents in Group B.

Differences in patient-reported experiences between the two groups were small, varying from only 0.3 (GP is competent) to 3.5 (waiting time for appointments that are not urgent is acceptable) on a scale from 0 to 100 (Table 5). There were no significant differences in the 5 indicators between the 2 groups, and only 1 of the 24 items was significantly different: the score for the item about the helpfulness and competence of other employees at GP practices was significantly higher in Group A than in Group B (p=0.046).

Discussion

This study compared response rates, background characteristics, data quality, and main study results between two randomized data-collection groups in a national survey of patient experiences with GPs. Patients in Group A were mailed an invitation with both pen-and-paper and electronic response options, while those in Group B received an email invitation with an electronic response option only. The response rate was 2.3-fold higher for the mail protocol than for the web-based protocol, but the patient-reported experiences were similar in the two groups.

The current study of patient experiences with GPs is the first to explore a purely electronic protocol in the national program for monitoring and reporting on health-care quality using patient experience surveys in Norway. Web-based surveys have many advantages, including

direct links to survey sites, ease of distribution, ease of receiving responses, and lower costs, but a major concern is that they exclude people without an email address as well as those with poor access to the Internet. The existence of a national register in Norway with electronic contact information presents a major opportunity for large-scale surveys of patient experiences. The vast majority of the Norwegian population is included in the national register and uses the Internet at a daily basis, reducing potential variations in the population coverage between paper- and web-based questionnaires and the risk of selection bias from using the Internet for questionnaire surveys. However, as many as 15% of the patients in the electronic arm lacked a valid email address in the national register, the corresponding number we could not reach in the standard mail data-collection was 5%. Furthermore, only 18% of the contacted sample in the web-based approach responded.

The results are consistent with literature reports that mail surveys achieve higher response rates than electronic and web-based approaches.[2–11] A recent Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey produced corresponding results when comparing protocols based on web responses via an email invitation and mail.[7] The mail protocol yielded more than twice the response rate of the web approach. A study of patient experiences with individual physicians showed that response rates were higher by mail (51%) than web (15%).[9] In a study of patient experiences with outpatient clinics 14% responded to the web-based survey and 33% responded to the mail survey.[8] When considering the completeness of the responses, our results are in line previous studies we found that the web-based questionnaire had fewer missing values than the mail protocol. The levels of ceiling effects and internal consistency were similar in the two groups.

Despite the marked differences in response rates, the results showed minor differences in the level of patient reported experiences between the standard mail data-collection procedure and a web-based approach. There were no significant differences in the 5 indicators between the 2 groups, and only one single item score was significantly different between the two groups. These results are in line with other findings.[2,3,7,8,9] There might be several reasons for the high correlation between the response modes. The surveys were designed to be as similar as possible, including the invitation letter, the content, layout and structure of the questionnaire and the timing of the first contact and reminders. The invitations to the patients in Group A and Group B were sent the same week and non-respondents in both groups received two reminders.

The response rate alone is a poor predictor of nonresponse bias, and previous studies have failed to find a consistent association between response rates and sample representativeness.[23,24] However, low response rates threaten the legitimacy of surveys in both the clinical and public domains, and reduce the ability of surveys to identify important differences in patient-reported experiences between providers and over time.[2,3] Future research needs to focus on effective initiatives for increasing response rates in web-based protocols, including sending multiple reminders using a combination of emails, messages on mobile phones, and other available platforms. For example, the national infrastructure in Norway provides the possibility for secure digital mailboxes for all Norwegian inhabitants, which could be utilized for contacting digitally active patients.

The current study showed that a lower education level and higher age were associated with a mail preference. In the current study we found that respondents invited by email were younger, more educated, and more likely to have had more-recent contact with their GP. We

found no significant intergroup differences in the remaining nine background variables. There are several methods for assessing non-response bias, including comparison of respondents and non-respondents on background variables.[25] When we compared respondents with non-respondents, we found that men and younger patients were underrepresented as respondents in both groups. These differences are normally handled by non-response weighting, but such weights are only able to compensate for variables available in the sampling frame. We did not conduct further analysis of non-respondents, but previous follow-up studies of non-respondents in Norway indicate small additional bias.[12-15] However, none of these have included a purely digital protocol, which warrant future non-response research for digital protocols. The coverage challenges for the digital sampling frame should be part of this research, as 12% of the population was not registered in the register, and 15% of the registered persons lacked a valid email address. This coverage challenge is an additional weakness of purely digital approaches and should be compensated with other response options for those excluded.

The effects of background characteristics reported in the literature are inconsistent.[2–10] The results from two previous randomized studies showed similar background characteristics for respondents in different randomized groups.[2,3] However, the respondents in those surveys were all contacted by mail. Future research should assess how the national infrastructure in Norway could be used to tailor the mode of data collection to different groups, such as by providing a range of data-collection modes from purely electronic strategies (for respondents with high education levels) to a mail-based mixed mode (to older respondents and those with low education levels).

Combined with the low response rates achieved for web-based protocols in this and other studies, future representative and high-quality surveys should include the opportunity to answer on pen-and-paper questionnaires. This could be implemented in a mixed-mode design that provides respondents with the option to choose how they want to respond, making it possible for patients without Internet access or enough computer skills to also participate.

A limitation of this study is that it only included adults evaluating their GPs in Norway, and so the results might not be generalizable to other health-care settings and countries. In particular, the national infrastructure and the digital maturity of the population in Norway might differ from the characteristics of other countries. However, the results should be applicable to health systems with similar infrastructures and digital maturities, and to countries working to establish regional or national digital infrastructures.

Conclusions

Administering a survey of patient experiences with GPs using a web-based protocol produced results that were very similar to those obtained using the standard mail-mode data-collection procedure that is used in the national surveys but had a much lower response rate. Furthermore, respondents in the digital group were younger, more educated, and had more-recent experiences with their GPs. Men and younger patients were underrepresented as respondents in both groups. Web-based surveys are faster and cheaper than standard mail surveys, but their low response rates threaten their legitimacy. Initiatives to increase response rates for web-based data collection and strategies for tailoring data collection to different groups should be key elements in future research.

Funding

This research received grant from the Norwegian Ministry of Health and Care Services.

Competing interests

On behalf of all authors, the corresponding author states that there are no competing interests.

Author contributions

The Ministry of Health and Care Services initiated the study. H.H.I. planned the study in consultation with O.B. and O.H. H.H.I. performed the statistical analyses with O.B. and O.H. and drafted the manuscript. O.B. and O.H. participated in the planning process, critically revised the manuscript draft and approved the final version of the manuscript. H.H.I. was the project manager for the survey. All authors read and approved the final manuscript.

Acknowledgements

We thank Inger Opedal Paulsrud from the NIPH for her help in conducting the survey, including performing administrative and technical tasks for the data collection. We would also like to thank Sylvia Drange Sletten, Marius Troøyen and Sindre Møgster Braaten from the NIPH for help developing the system for web-based data collection.

Availability of data and material

The data set generated and/or analysed during this study is not publicly available due to the need to protect personal data.

References

1. Eide TB, Straand J, Rosvold EO. Patients' and GPs' expectations regarding healthcare-seeking behaviour: a Norwegian comparative study. BJGP Open. 2018;2(4):bjgpopen18X101615.
2. Bjertnaes OA, Iversen HH. User-experience surveys with maternity services: a randomized comparison of two data collection models. Int J Qual Health Care. 2012;24(4):433–8.
3. Bjertnaes O, Iversen HH, Skrivarhaug T. A randomized comparison of three data collection models for the measurement of parent experiences with diabetes outpatient care. BMC Med Res Methodol. 2018;18 (1):95.
4. Holmboe O, Bjertnæs OA. Patient experiences with Norwegian hospitals in 2015. Results following a national survey. Oslo: Norwegian Knowledge Centre for the Health Services, PasOpp-report 147; 2016. ISSN 1890-1565.
5. Holmboe O, Danielsen K, Iversen HH. Development of a method for the measurement of patient experiences with regular general practitioners. Oslo: Norwegian Knowledge Centre for the Health Services, PasOpp-report 1; 2015. ISSN 1890-1565.
6. Danielsen K, Holmboe O, Hansen SL, Førland O, Bjertnæs OA. Patient experiences with Norwegian primary care out-of-hours clinics. Oslo: Results from a survey among the "Vakttårn" clinics. Norwegian Knowledge Centre for the Health Services, PasOpp-report 1; 2012. ISSN 1890-1565.
7. Fowler FJ Jr, Cosenza C, Cripps LA, Edgman-Levitan S, Cleary PD. The effect of administration mode on CAHPS survey response rates and results: a comparison of mail and web-based approaches. Health Serv Res. 2019;54(3):714–21.
8. Bergeson SC, Gray J, Ehrmantraut LA, Laibson T, Hays RD. Comparing web-based with mail survey administration of the Consumer Assessment of Healthcare Providers

and Systems (CAHPS(®)) Clinician and Group survey. Prim Health Care. 2013;3:1000132.

9. Rodriguez HP, von Glahn 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(2):167–74.
10. Holmboe O, Iversen HH, Danielsen K, Bjertnaes O. The Norwegian patient experiences with GP questionnaire (PEQ-GP): reliability and construct validity following a national survey. BMJ Open. 2017;7(9):e016644.
11. Iversen HH, Bjertnaes O, Helland Y, Skrivarhaug T. The Adolescent Patient Experiences of Diabetes Care Questionnaire (APEQ-DC): Reliability and Validity in a Study Based on Data from the Norwegian Childhood Diabetes Registry. Patient Relat Outcome Meas. 2019;10:405-16.
12. Guldvog B, Hofoss D, Pettersen KI, Ebbesen J, Ronning OM: PS-RESKVA - pasienttilfredshet i sykehus [PS-RESKVA -patient satisfaction in hospitals]. Tidsskr Nor Laegeforen. 1998, 118: 386-391.
13. Dahle KA, Iversen H, Sjetne IS, Oltedal S: Do non-respondents' characteristics and experiences differ from those of respondents? Poster at the the 25th International Conference by the Society for Quality in Health Care (ISQua). 2008, Copenhagen.
14. Iversen H, Dahle KA, Holmboe O: Using telephone interviews to explore potential non-response bias in a national postal survey. Poster at the the 26th International Conference by the Society for Quality in Health Care (ISQua). 2009, Dublin.
15. Danielsen K, Iversen H, Holmboe O. Telephone interviews to assess non-response bias in a postal user experience survey: Kirsten Danielsen. European Journal of Public Health. 2015;25(suppl_3).

16. Kompetanse Norge. Befolkningens bruk av digitale verktøy. Available from: https://www.kompetansenorge.no/statistikk-og-analyse/grunnleggende-digital-ferdigheter/befolkingens-bruk-av-digitale-verktoy/#_[Accessed 10 June 2020].
17. Digitaliseringsdirektoratet. Befolkningens bruk av digitale verktøy. Kontakt- og reservasjonsregisteret. Available from: https://samarbeid.difi.no/statistikk/kontakt-og-reservasjonsregisteret_[Accessed 10 June 2020].
18. The American Association for Public Opinion Research. 2016. Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 9th edition. AAPOR. https://www.aapor.org/Standards-Ethics/Standard-Definitions-(1).aspx. [Accessed 10 June 2020].
19. Nunnally JC, Bernstein IH. Psychometric Theory. 3rd edn. New York: McGraw-Hill; 1994.
20. Kline RB. Principles and Practice of Structural Equation Modeling. New York: Guildford; 2005.
21. Bjertnaes OA, Lyngstad I, Malterud K, Garratt A. The Norwegian EUROPEP questionnaire for patient evaluation of general practice: data quality, reliability and construct validity. Fam Pract. 2011;28(3):342–9.
22. Ruiz MA, Pardo A, Rejas J, Soto J, Villasante F, Aranguren JL. Development and validation of the "Treatment Satisfaction with Medicines Questionnaire" (SATMED-Q). Value Health. 2008;11(5):913–26.
23. Groves RM, Peytcheva E. The impact of nonresponse rates on nonresponse bias: a meta-analysis. Public Opin Q. 2008;72(2):167–89.
24. Groves RM. Nonresponse rates and nonresponse bias in household surveys. Public Opin Q. 2006;70(5):646–75.

25. Halbesleben JRB, Whitman MV. Evaluating Survey Quality in Health Services Research: A Decision Framework for Assessing Nonresponse Bias. Health Services Research. 2013;48(3):913-30.

Figure 1: CONSORT flow diagram

Table 1: Respondents before and after each reminder in the two randomized groups, and final response rates.

	Group A (*n*=4,760)	Group B (*n*=1,694)
Respondents before reminder:		
Electronic, *n*	272	117
Paper, *n*	560	-
Response rate, %	17.5	6.9
Respondents after first reminder:		
Electronic, *n*	171	126
Paper, *n*	533	-
Increase in response rate, %	14.8	7.4
Respondents after second reminder:		
Electronic, *n*	148	67
Paper, *n*	345	-
Increase in response rate, %	10.4	4.0
Total:		
Electronic, *n* (%)	591 (29.1)	310
Paper, *n* (%)	1438 (70.9)	
Response rate, %	42.6	18.3

Table 2: Comparison of missing data, ceiling effects, and internal consistency between the two randomized groups.

Scale and item[a]	Group A				Group B			
	Missing data (%)	Not applicable (%)	Ceiling effects (%)	Cronbach's alpha/item-total correlation	Missing data (%)	Not applicable (%)	Ceiling effects (%)	Cronbach's alpha/item-total correlation
GP				**0.924**				**0.935**
Do you feel that your GP takes you seriously?	5.0	3.4	48.0	0.762	0.0	2.9	52.2	0.813
Do you feel that your GP spends enough time with you?	4.9	3.2	28.9	0.702	0.0	3.2	32.7	0.745
Do you feel that your GP talks to you in a way you understand?	5.1	3.0	48.9	0.737	0.3	2.3	56.0	0.735
Do you feel that your GP is professionally competent?	5.2	4.0	41.8	0.752	0.6	2.9	43.1	0.805
Do you feel that your GP shows interest in your situation?	5.3	3.5	39.5	0.818	0.6	2.3	39.9	0.853
Do you feel that your GP includes you as much as you would like in decisions concerning you?	5.7	7.7	37.2	0.769	0.3	6.5	41.9	0.803
Does your GP provide you with sufficient information about your health problems and their treatment?	5.6	7.3	32.7	0.811	0.3	6.1	34.1	0.835
Does your GP provide you with sufficient information about the use and side effects of medication?	5.4	16.3	21.0	0.631	0.0	16.5	24.3	0.655
Does your GP refer you to further examinations or a specialist when you feel you need it?	5.1	10.6	43.2	0.646	0.3	10.6	49.3	0.633
Organization and auxiliary staff				**0.868**				**0.851**
Do you feel that your GP's practice is well organized?	5.2	4.1	26.6	0.681	1.0	1.6	25.8	0.615
Do you feel the other employees are helpful and competent?	4.9	3.6	36.1	0.813	0.3	2.9	31.7	0.806
Are you treated with courtesy and respect at the reception?	4.8	3.2	40.9	0.752	0.3	2.6	39.5	0.750
Accessibility				**0.774**				**0.688**
Was the waiting time for your last urgent appointment acceptable?	18.7	-	36.2	0.631	17.1	-	33.5	0.525
Is this waiting time for appointments that are not urgent acceptable?	12.3	-	18.0	0.631	6.8	-	11.4	0.525
Enablement				**0.906**				**0.925**
Does contact with your GP make you better able to understand your health problems?	1.6	16.6	19.5	0.803	1.0	15.8	26.0	0.836
Does contact with your GP make you better able to cope with your health problems?	1.6	19.7	16.3	0.852	0.6	20.3	21.2	0.875
Does contact with your GP better help you to stay healthy?	1.6	19.8	15.1	0.786	0.3	20.6	21.6	0.833
Coordination and cooperation				**0.875**				**0.876**
Do you feel that your GP is good at coordinating the range of health services available to you?	5.9	26.9	28.1	0.779	0.6	30.6	34.7	0.790
Do you feel that your GP cooperates well with other services you need?	5.7	29.4	29.0	0.779	0.6	31.9	29.7	0.790

[a]All items were scored on a 5-point response scale ranging from 1 ("not at all") to 5 ("to a very large degree")

Table 3: Comparison of respondent characteristics between the two randomized groups.

	Group A	Group B	*p*[a]
Gender, female	55.9 (1135)	59.7 (185)	0.216
Age group			<0.001
16–19 years	2.4 (49)	2.3 (7)	
20–29 years	7.8 (158)	9.4 (29)	
30–49 years	23.8 (482)	37.4 (116)	
50–66 years	34.2 (694)	31.6 (98)	
≥67 years	31.8 (646)	19.4 (60)	
Time on the list of the GP			0.526
<1 year	9.4 (191)	8.4 (26)	
1–2 years	19.4 (392)	21.3 (66)	
3–4 years	14.5 (293)	16.1 (50)	
5–10 years	20.4 (414)	22.3 (69)	
≥11 years	36.3 (735)	31.9 (99)	
Number of consultations during past 12 months			0.672
0	9.3 (186)	10.1 (31)	
1	15.7 (314)	15.3 (47)	
2–5	55.7 (1114)	52.1 (160)	
6–12	16.2 (323)	19.2 (59)	
≥13	3.1 (63)	3.3 (10)	
Number of diagnosis codes during past 24 months	13.8±13.5	12.6±10.7	0.083
Number of unique diagnosis codes during past 24 months	4.7±3.2	4.6±2.8	0.510
Time since previous contact			0.042
<1 month	36.5 (716)	42.1 (128)	
1–3 months	32.0 (628)	23.7 (72)	
4–6 months	13.5 (266)	15.1 (46)	
7–12 months	9.7 (191)	8.9 (27)	
>12 months	8.3 (163)	10.2 (31)	
Self-perceived physical health			0.951
Very poor	1.3 (27)	1.6 (5)	
Quite poor	5.3 (108)	5.2 (16)	
Both poor and good	23.8 (481)	22.4 (69)	
Quite good	48.3 (975)	50.3 (155)	
Very good	21.2 (429)	20.5 (63)	
Self-perceived mental health			0.475
Very poor	1.1 (22)	1.9 (6)	
Quite poor	3.0 (60)	3.9 (12)	
Both poor and good	15.5 (313)	15.6 (48)	
Quite good	41.7 (842)	38.0 (117)	
Very good	38.7 (781)	40.6 (125)	
Long-standing health problems			0.625
0	35.7 (708)	37.5 (115)	
1	32.9 (653)	34.9 (107)	
2	19.3 (383)	16.9 (52)	
≥3	12.1 (241)	10.7 (33)	
Education level			<0.001
Elementary school	15.6 (309)	7.1 (22)	
High school	37.4 (740)	31.4 (97)	
University, 0–4 years	25.6 (505)	35.3 (109)	
University, >4 years	21.4 (422)	26.2 (81)	
Geographic origin			0.205
Norway	88.6 (1756)	89.9 (276)	
Asia (incl. Turkey), Africa, or Latin America	4.8 (95)	3.3 (10)	
Eastern Europe (all countries, independent of EU membership)	3.5 (70)	2.3 (7)	
Western Europe, North America, or Oceania	3.0 (60)	4.6 (14)	

[a] Pearson chi-square tests for categorical variables and independent-samples *t*-tests for continuous variables. Data are *n* (%) or mean±SD values.

Table 4: Comparison of respondent characteristics between the two randomized groups.

	Group A			Group B		
	Respondents	Non-respondents	*p*[a]	Respondents	Non-respondents	*p*[a]
Gender, female	55.9 (1135)	46.1 (1259)	<0.001	59.7 (185)	49.8 (689)	0.002
Age group			<0.001			<0.001
16–19 years	2.4 (49)	7.3 (200)		2.3 (7)	6.4 (88)	
20–29 years	7.8 (158)	20.7 (566)		9.4 (29)	17.6 (243)	
30–49 years	23.8 (482)	39.5 (1080)		37.4 (116)	38.8 (537)	
50–66 years	34.2 (694)	21.2 (579)		31.6 (98)	26.0 (360)	
≥67 years	31.8 (646)	11.2 (306)		19.4 (60)	11.3 (156)	
Time on the list of the GP			<0.001			0.739
<1 year	9.4 (191)	10.5 (288)		8.4 (26)	9.6 (133)	
1–2 years	19.4 (392)	24.3 (664)		21.3 (66)	23.3 (322)	
3–4 years	14.5 (293)	15.3 (419)		16.1 (50)	13.7 (189)	
5–10 years	20.4 (414)	20.9 (570)		22.3 (69)	22.3 (309)	
≥11 years	36.3 (735)	28.9 (790)		31.9 (99)	31.1 (431)	
Number of consultations during past 24 months	10.8±11.3	7.6±10.7	<0.001	9.6±9.2	8.4±11.5	0.077
Number of diagnosis codes during past 24 months	13.8±13.5	11.3±14.1	<0.001	12.6±10.7	12.2±14.9	0.611
Number of unique diagnosis codes during past 24 months	4.7±3.2	4.1±3.1	<0.001	4.6±2.8	4.2±3.2	0.107

[a] Pearson chi-square tests for categorical variables and independent-samples *t*-tests for continuous variables. Data are *n* (%) or mean±SD values.

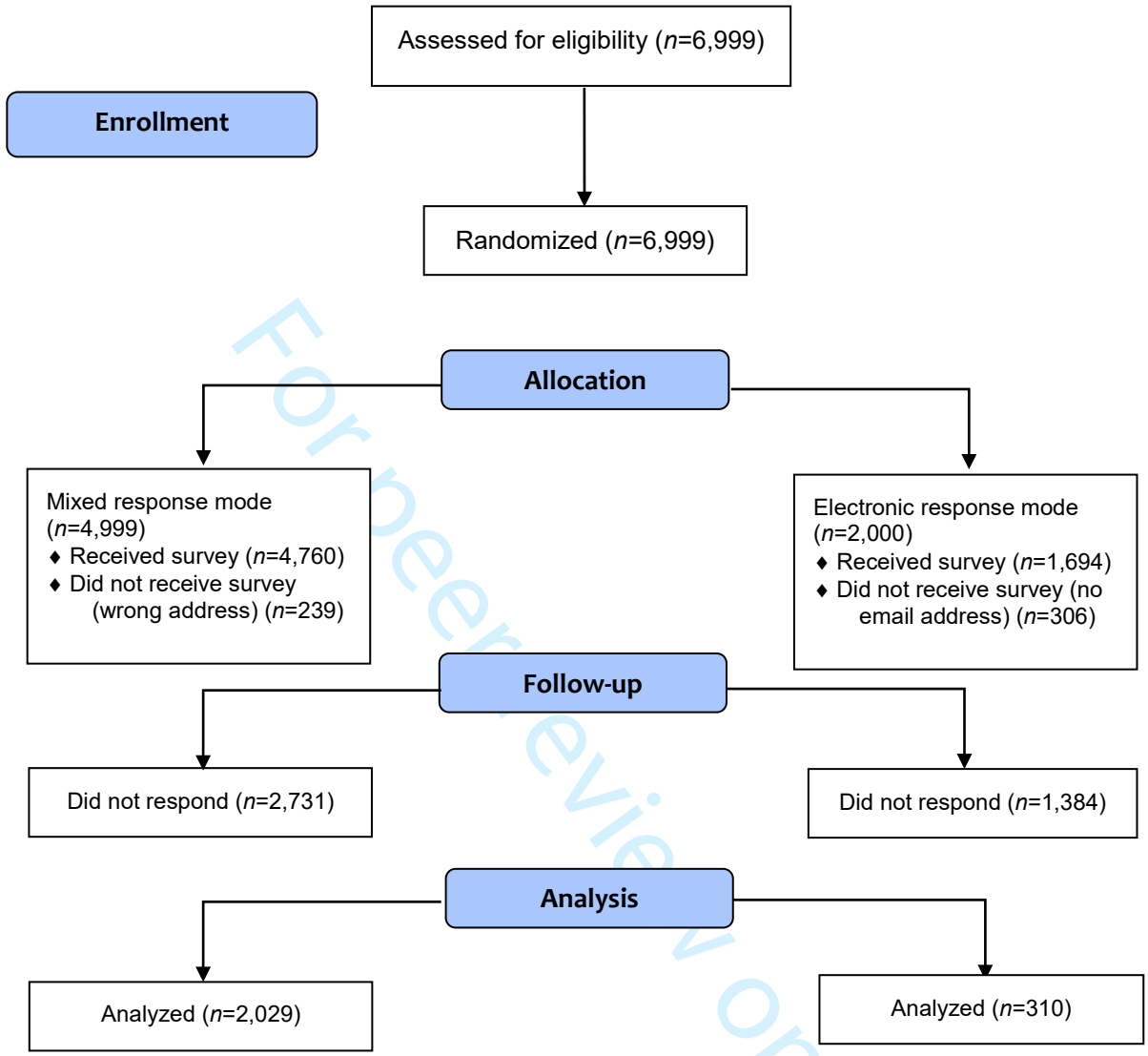
Table 5: Comparison of patient-reported experiences between the two randomized groups.

Scale and item[a]	Group A	Group B	*P*
GP	**78.3±16.8**	**78.8±17.8**	**0.651**
Do you feel that your GP takes you seriously?	83.1±19.8	84.1±20.3	0.429
Do you feel that your GP spends enough time with you?	73.8±22.8	73.3±25.8	0.707
Do you feel that your GP talks to you in a way you understand?	84.6±17.7	85.8±19.1	0.275
Do you feel that your GP is professionally competent?	82.1±18.0	81.8±19.4	0.747
Do you feel that your GP shows interest in your situation?	79.7±20.5	78.7±22.2	0.435
Do you feel that your GP includes you as much as you would like in decisions concerning you?	79.0±20.3	80.0±21.3	0.452
Does your GP provide you with sufficient information about your health problems and their treatment?	76.6±21.2	76.1±22.2	0.708
Does your GP provide you with sufficient information about the use and side effects of medication?	65.0±26.9	67.1±26.0	0.241
Does your GP refer you to further examination or a specialist when you feel you need it?	81.4±20.2	82.0±22.2	0.712
Organization and auxiliary staff	**78.2±17.7**	**77.1±18.2**	**0.322**
Do you feel that your GP's practice is well organized?	75.2±20.2	74.8±20.4	0.735
Do you feel the other employees are helpful and competent?	79.3±19.2	76.9±20.8	0.046
Are you treated with courtesy and respect at the reception?	80.9±19.7	79.9±20.7	0.430
Accessibility	**63.6±27.8**	**61.8±25.5**	**0.264**
Was the waiting time for your last urgent appointment acceptable?	69.5±30.6	69.1±30.4	0.828
Is this waiting time for appointments that are not urgent acceptable?	58.3±30.0	54.8±28.5	0.051
Enablement	**65.2±22.1**	**66.0±24.3**	**0.601**
Does contact with your GP make you better able to understand your health problems?	68.1±23.0	68.7±25.1	0.703
Does contact with your GP make you better able to cope with your health problems?	64.9±24.0	65.9±25.6	0.560
Does contact with your GP better help you to stay healthy?	62.7±25.0	64.5±26.7	0.303
Coordination and cooperation	**74.3±21.0**	**74.9±21.5**	**0.644**
Do you feel that your GP is good at coordinating the range of health services available to you?	75.0±21.2	77.7±20.5	0.079
Do you feel that your GP cooperates well with other services you need?	74.3±22.6	73.0±24.5	0.434

[a]All scales and items are scored from 0 to 100, where 100 is the best possible patient experience. Data are mean±SD values

.

Figure 1: CONSORT flow diagram



Reporting checklist for cross sectional study.

Based on the STROBE cross sectional guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the STROBE cross sectionalreporting guidelines, and cite them as:

von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies.

		Reporting Item	Page Number
Title and abstract			
Title	#1a	Indicate the study's design with a commonly used term in the title or the abstract	2
Abstract	#1b	Provide in the abstract an informative and balanced summary of what was done and what was found	2,3
Introduction			
Background / rationale	#2	Explain the scientific background and rationale for the investigation being reported	5-6
Objectives	#3	State specific objectives, including any prespecified hypotheses	2,6
Methods			
Study design	#4	Present key elements of study design early in the paper	2, 6/7
Setting	#5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6-8
Eligibility criteria	#6a	Give the eligibility criteria, and the sources and methods of selection of participants.	2, 6-8

	#7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7,8
Data sources / measurement	#8	For each variable of interest give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. Give information separately for for exposed and unexposed groups if applicable.	7,8
Bias	#9	Describe any efforts to address potential sources of bias	11-13
Study size	#10	Explain how the study size was arrived at	6,7
Quantitative variables	#11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	6-8
Statistical methods	#12a	Describe all statistical methods, including those used to control for confounding	7-8
Statistical methods	#12b	Describe any methods used to examine subgroups and interactions	6-8
Statistical methods	#12c	Explain how missing data were addressed	7,8,19
Statistical methods	#12d	If applicable, describe analytical methods taking account of sampling strategy	6
Statistical methods	#12e	Describe any sensitivity analyses	n/a
Results			
Participants	#13a	Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed. Give information separately for for exposed and unexposed groups if applicable.	6,7,17
Participants	#13b	Give reasons for non-participation at each stage	17
Participants	#13c	Consider use of a flow diagram	17
Descriptive data	#14a	Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders. Give information separately for exposed and unexposed groups if applicable.	9,10,20
Descriptive data	#14b	Indicate number of participants with missing data for each variable of interest	19
Outcome data	#15	Report numbers of outcome events or summary measures. Give information separately for exposed and unexposed groups if applicable.	18,19,21
Main results	#16a	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	n/a

Main results	#16b	Report category boundaries when continuous variables were categorized	n/a
Main results	#16c	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	#17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	n/a
Discussion			
Key results	#18	Summarise key results with reference to study objectives	10,11
Limitations	#19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.	12,13
Interpretation	#20	Give a cautious overall interpretation considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence.	10-13
Generalisability	#21	Discuss the generalisability (external validity) of the study results	12,13
Other Information			
Funding	#22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	13

BMJ Open

Patient-reported experiences with general practitioners: a randomized study of mail and web-based approaches following a national survey

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-036533.R2
Article Type:	Original research
Date Submitted by the Author:	30-Aug-2020
Complete List of Authors:	Iversen, Hilde; Norwegian Institute of Public Health, Division of Health Services Holmboe, Olaf; Norwegian Institute of Public Health, Division of Health Services Bjertnaes, Oyvind; Norwegian Institute of Public Health, Division of Health Services
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Health services research
Keywords:	PRIMARY CARE, STATISTICS & RESEARCH METHODS, Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Patient-reported experiences with general practitioners: a randomized study of mail and web-based approaches following a national survey

Authors:

Hilde Hestad Iversen[1] (hilde.hestad.iversen@fhi.no) +47 46400425) (Corresponding author)

Olaf Holmboe[1] (olaf.holmboe@fhi.no)

Oyvind Bjertnaes[1] (oyvindandresen.bjertnaes@fhi.no)

Affiliations:

[1]Division of Health Services, Norwegian Institute of Public Health, PO Box 222 Skøyen, 0213 Oslo, Norway

Keywords: data collection, general practitioners, patient satisfaction, physicians, random allocation, surveys and questionnaires

Word count: 3,917

Abstract

Objective

The standard data-collection procedure in the Norwegian national patient experience survey programme is post-discharge mail surveys, which include a pen-and-paper questionnaire with the option to answer electronically. A purely electronic protocol has not previously been explored in Norway. The aim of this study was to compare response rates, background characteristics, data quality and main study results for a survey of patient experiences with general practitioners (GPs) administered by the standard mail data-collection procedure and a web-based approach.

Design

Cross-sectional survey.

Setting

General practitioner offices in Norway.

Participants

The sample consisted of 6,999 patients aged 16 years and older registered with a GP in November 2018.

Intervention

Based on a three-stage sampling design, 6,999 patients of GPs aged 16 or older were randomized to one of two survey administration protocols: Group A, who were mailed an invitation with both a pen-and-paper including an electronic response option (n=4,999) and Group B, who received an email invitation with electronic response option (n=2,000).

Main outcome measures

Response rates, background characteristics, data quality and main study results.

Results

The response rate was markedly higher for the mail survey (42.6%) than for the web-based survey (18.3%). A few of the background variables differed significantly between the two groups, but the data quality and patient-reported experiences were similar.

Conclusions

Web-based surveys are faster and less expensive than standard mail surveys, but their low response rates and coverage problems threaten their usefulness and legitimacy. Initiatives to increase response rates for web-based data collection and strategies for tailoring data collection to different groups should be key elements in future research.

Article Summary

Strengths and Limitations

- The current study is the first to explore a purely electronic protocol in the national program of patient experience surveys in Norway
- No previous surveys in the national program have tested coverage and the quality of the email addresses in the national register for contact information
- The study did not use other available digital contact methods than email addresses, and the generalizability to health systems with different infrastructures and digital maturities is uncertain
- The study included adults evaluating their GPs, and the results might not be generalizable to other patient groups and health-care settings

Introduction

Norway introduced the regular General Practitioner (GP) scheme in 2001. All inhabitants registered in the National Registry as Norwegian residents have the right to a GP/family doctor. Migrants eligible to stay in Norway for more than six months are entitled to enrol in the scheme. GPs in Norway play a key role in the provision of health care, and are often the first point of contact to acquire health services for most medical problems.[1] In 2018, The Ministry of Health and Care Services decided to evaluate the GP scheme, and part of this evaluation comprised a national patient experience survey.

The Norwegian Institute of Public Health (NIPH) is responsible for conducting national patient experience surveys in Norway. Norway has a national program for monitoring and reporting on the quality of health care using patient experience surveys. The purpose of this program is to measure user experiences with health care systematically, with the obtained data used as a basis for interventions aimed at improving the quality of health care, health-care management, patient choice, and public accountability. The standard data-collection procedure in the national surveys is post-discharge mail surveys, which include a pen-and-paper questionnaire and an option to answer electronically.

The results from previous studies of survey-mode preferences in different patient populations both in Norway and other countries indicate that web mode surveys have lower response rates than other modes.[2–11] In the national patient experience survey among patients visiting general practitioners in 2014 in Norway, only 18% of respondents answered electronically.[4] However, the potential advantages of lower costs and shorter data-collection periods are important arguments for performing further research into web-based surveys. Also, the expansion of

Internet access and use may have changed the potential of the Internet as an effective way to conduct such surveys.

When comparing the standard mail survey mode of data collection with web-based data collection the characteristics of non-respondents and respondents in both groups should be explored. The literature on the effects of background characteristics on the responses to different data collection methods are inconsistent.[2–10] Non-response bias has been studied in four patient populations in Norway through follow-up telephone interviews with non-respondents,[12–15] including non-respondents in a survey on patient experiences with GPs.[15] The results have shown minor differences between the postal respondents from the national surveys and the postal non-respondents who have provided answers through follow-up interviews. In general, the impact of non-response bias in the large-scale surveys has been considered relatively small.

The use of Internet in the general population is growing. In 2018, 90% of all Norwegian citizens used the Internet on a daily basis. [16] In all age groups under 60 years, between 90-99 percent reported using the Internet daily, but corresponding results for those between 60-69 years was 81% and 67% for those aged 70 years or more. Seventeen percent of the citizens aged 70 years or more reported that they never used the Internet. Potential differences in population coverage between paper- and web-based questionnaires and the risk of selection bias from using the Internet for questionnaire surveys has been reduced, but a major concern with protocols that use only digital responses continues to be leaving out people without available digital contact information.

A purely electronic protocol for patient experience surveys has not previously been explored in the national program for monitoring and reporting on health-care quality in Norway. A main limitation of previous studies has been the lack of e-mail addresses in the sample frame, with the implication that even the electronic group had to be invited by a postal invitation, adding to costs, and precluding the possibility of testing a comprehensive electronic data collection option. The establishment of a national register with electronic contact information opens new possibilities regarding electronic and web-based surveys. A total of 88% of the population was registered in the national register for contact information in November 2018.[17] So far, this register has not been utilized in our national patient experience surveys.

The aim of the current study was to compare the standard mail survey mode of data collection with exclusively web-based data collection in Norway. The sample was randomized to one of two survey administration protocols: patients in Group A were mailed an invitation with both pen-and-paper and electronic response options, while those in Group B received an email invitation with an electronic response option only (using email addresses obtained from the national register). The response rates, data quality, background characteristics, and main study results were compared between the two groups.

Methods

Data

The sample consisted of patients aged 16 years and older registered with a GP in November 2018. The preconditions for the sampling frame were to report the results on a national level and to be able to estimate intraclass correlation coefficients on the GP practice level. With the patient sample size chosen, we explored how the intraclass correlation coefficient (ICC)

varied dependent on the number of GPs at the practice level and found that at least four GPs were needed per GP practice to reach an acceptable ICC, and that not much was gained by including more GPs per practice. The sampling plan had a three-stage design. First, regular GP practices were randomly selected after stratification by the number of GPs per practice and municipality type. Second, all the GPs were included in the selected practices that had up to four GPs, while four of them were randomly selected in the practices that had five or more GPs. Third, we randomly selected 14 adult patients from the list of patients for each GP.

This study included a total of 6,999 patients. Patients were randomized to 1 of 2 survey administration protocols: 4,999 patients to the main sample (Group A) and 2,000 patients to a subsample (Group B) (Fig. 1). The current study was the first to explore a purely electronic protocol in the national program of patient experience surveys in Norway. The quality of the patient contact information collected from the national register was also previously unexplored. Considering the commission of achieving nationally representative results and the uncertainty regarding the responses from a purely electronic protocol, we evaluated the risk of randomizing the total sample in two groups as too high and chose to include fewer patients in the subsample.

Patients in Group A were mailed an invitation with both pen-and-paper and electronic response options. The invitation included a cover letter describing the purpose of the study, a paper questionnaire, a prepaid envelope and information and a login code to be able to respond electronically. The patients in Group B received an email invitation with an electronic response option only. The email invitation included information about the purpose of the study, a link to the online survey and a login code. Two reminders were sent to non-respondents in both samples using the same contact mode as the first invitation. The first

reminder was sent to both groups around three weeks after the first contact. The second reminder was sent around six weeks after the first contact. All reminders to Group A were sent by mail and included a new invitation, the paper questionnaire, the postage-paid envelope and the login code to enable electronic responses. Group B were sent a new email invitation with a link to the survey and a login code in both reminders.

Background data about the patients were obtained from public registries, including gender, age, the number of years on the patient list of a GP, and the number of consultations during the past 24 months. Email addresses were collected from the national register for contact information, which is operated by the Agency for Public Management and eGovernment.

Measures

The Norwegian PEQ-GP (Patient Experiences with General Practitioner Questionnaire) was applied. This instrument was developed and validated according to the standard scientific procedures of the national patient-reported experience program in Norway.[5,10]

A national validation study identified five scales that covered important aspects of the GP service relating to accessibility, evaluations of the GP and auxiliary staff, cooperation between the GP and other services, and patient enablement. We included 17 additional items that were relevant for evaluating the GP scheme. The questionnaire used in the randomized study consisted of 47 questions on 6 pages. Thirty-seven questions addressed experiences with the GP service, while ten were background questions. Most of the questions related to the user-reported experiences were answered in a 5-point response format ranging from “not at all” to “to a very large degree”. Single item and index scores were transformed linearly from the 1 to 5 scale to a scale of 0–100 An additional page was included to allow the

respondents to write comments related to experiences with their GPs and suggestions for future changes to the GP scheme.

Patient and Public Involvement

Patients were included in the development process of the instrument, securing the inclusion of the most important topics for patients. To identify important topics, we assessed reviews of the literature and consulted a reference group comprising GPs, researchers and representatives from health authorities and patient organisations throughout the process of questionnaire development. Cognitive interviews with patients were used to test the questionnaire. First, eight face-to-face interviews and nine telephone interviews were conducted. After an extensive revision, we conducted another 11 face-to-face interviews with patients. The revised version was tested in a pilot study.

Statistical analysis

The survey response rate by group was calculated as the proportion of eligible patients (i.e. those who had not changed address, died, or were otherwise ineligible) who returned a completed survey (AAPOR response rate 4.0).[18]

Items were assessed for levels of missing data, ceiling effects, and internal consistency. The internal consistency reliability of the five scales was assessed using the item-total correlation and Cronbach's alpha. The item-total correlation coefficient quantifies the strength of an association between an item and the remainder of its indicator, with a coefficient of 0.4 considered acceptable. [19] Cronbach's alpha assesses the overall correlation between items within an indicator, and an alpha value of 0.7 is considered satisfactory.[19,20] We set the cut-off

criterion for ceiling effects to 50%; that is, an item was considered acceptable if fewer than 50% of the respondents chose the most-favourable response option.[21,22]

Differences in respondent characteristics between Group A and Group B were tested using Pearson chi-square tests for categorical variables and independent-samples *t*-tests for continuous variables. Differences between the two groups regarding patient-reported experiences were tested using *t*-tests.

Differences in respondent characteristics between respondents and non-respondents in Group A and respondents and non-respondents in Group B were tested using Pearson chi-square tests for categorical variables and independent-samples t-tests for continuous variables. Variables available on non-respondents were gender, age, time on the list of the GP, number of consultations during the past 24 months and number of diagnosis during the past 24 months.

All the statistical analyses were performed using SPSS (version 25.0).

Approval

The Data Protection Officer at the NIPH recommended that the study be approved, and it was formally approved by the research director of the division for health services at the NIPH. The Norwegian Directorate of Health approved the use of data from non-respondents in the nonresponse analysis, except those of patients who withdrew themselves from the study. Return of the questionnaire represented patient consent in the study, which is the standard procedure in all patient experience surveys conducted by the Norwegian Institute of Public Health.

Results

The overall response rate was 42.6% in Group A and 18.3% in Group B (Table 1). 15% of the patients in the electronic arm lacked a valid email address in the national register, and 5% of the patients in the standard mail survey mode lacked a valid mailing address (Fig. 1). Most of the respondents (70.9%) in Group A answered on paper (Table 1). The initial response rate was around 10% lower for Group B than for group A, with the remaining difference being related to reduced effects of both the first and second reminders.

The levels of missing data, proportion of responses in the "not applicable" option, ceiling effects, and internal consistency for the items are presented in Table 2. The levels of missing data ranged from 1.6% to 18.7% in Group A, and from 0.0% to 17.1% in Group B. The proportions of responses in the "not applicable" category ranged from 3.0% to 29.4% in Group A, and from 1.6% to 31.9% in Group B, and were higher in Group A than in Group B for all items except for two on the enablement scale and the items on the coordination and cooperation scale. All scales and items were below the ceiling-effect criterion of 50% in Group A, but two items exceed the criterion in Group B: one about whether the GP takes the patient seriously (52.2%) and the other about whether the GP communicates in a way that the patient can understand (56.0%). Cronbach's alpha values were similar in the two groups for four of the five indicators, but was lower (and below the criterion of 0.7) for the accessibility indicator in Group B. The remaining Cronbach's alpha values were above 0.7.

Table 3 compares the background characteristics of the respondents in the two survey administration protocols. Respondent age, time since previous contact, and education level

differed significantly between the two groups, yet there were no significant differences in gender, number of years on the list of the GP, number of consultations, number of diagnosis codes during the past 24 months, number of unique diagnosis codes the past 24 months, self-perceived physical health, self-perceived mental health, long-standing health problems, or geographic origin. The proportion of patients aged 30–49 years was higher in Group B than in Group A (37.4% compared to 23.8%). In Group A, 31.8% of the patients were aged ≥67 years, a much higher proportion than in Group B where the corresponding proportion was 19.4%. The respondents in Group B were more likely to report that they had been in contact with their GP during the previous month than respondents in Group A. There was a significant tendency for those who responded to the email invitation to have a higher education level than those who responded to the mailed invitation: 61.5% of those in Group B reported being educated to the university level, compared to 47.0% in Group A.

Significant differences were found between Group A and Group B within respondents and non-respondents with respect to gender and age (Table 4). In both groups, non-respondents tended to be more likely to be men and to be younger than respondents. Significant differences were also found for time on the list of the GP, number of consultations during the past 24 months and the two variables about number of diagnosis the last two years for Group A. Respondents tended to have been longer on the GPs list, and to have a higher number of consultations and diagnosis during the last two years. We found no additional differences between respondents and non-respondents in Group B.

Differences in patient-reported experiences between the two groups were small, varying from only 0.3 (GP is competent) to 3.5 (waiting time for appointments that are not urgent is acceptable) on a scale from 0 to 100 (Table 5). There were no significant differences in the 5

indicators between the 2 groups, and only 1 of the 24 items was significantly different: the score for the item about the helpfulness and competence of other employees at GP practices was significantly higher in Group A than in Group B (p=0.046).

Discussion

This study compared response rates, background characteristics, data quality, and main study results between two randomized data-collection groups in a national survey of patient experiences with GPs. Patients in Group A were mailed an invitation with both pen-and-paper and electronic response options, while those in Group B received an email invitation with an electronic response option only. The response rate was 2.3-fold higher for the mail protocol than for the web-based protocol, but the patient-reported experiences were similar in the two groups.

The current study of patient experiences with GPs is the first to explore a purely electronic protocol in the national program for monitoring and reporting on health-care quality using patient experience surveys in Norway. Web-based surveys have many advantages, including direct links to survey sites, ease of distribution, ease of receiving responses, and lower costs, but a major concern is that they exclude those without an email address or with poor access to the Internet. The existence of a national register in Norway with electronic contact information presents a major opportunity for large-scale surveys of patient experiences. The vast majority of the Norwegian population is included in the national register and use the Internet on a daily basis, reducing potential variations in the population coverage between paper- and web-based questionnaires and the risk of selection bias from using the Internet for questionnaire surveys. However, as many as 15% of the patients in the electronic arm lacked

a valid email address in the national registry, the corresponding number we could not reach in the standard mail data-collection was 5%. Furthermore, only 18% of the contacted sample in the web-based approach responded.

The results are consistent with a number of previous studies reporting that mail surveys achieve higher response rates than electronic and web-based approaches.[2–11] A recent Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey produced similar results when comparing protocols based on web responses via an email invitation and mail.[7] The mail protocol yielded more than twice the response rate of the web approach. A study of patient experiences with individual physicians showed that response rates were higher using mail (51%) than web (15%).[9] In a study of patient experiences with outpatient clinics 14% responded to the web-based survey and 33% responded to the mail survey.[8] When considering the completeness of the responses, we found that the web-based questionnaire had fewer missing values than the mail protocol, which is in line with previous studies. The levels of ceiling effects and internal consistency were similar in the two groups.

Despite the marked differences in response rates, the results showed minor differences in the level of patient reported experiences between the standard mail data-collection procedure and a web-based approach. There were no significant differences in the 5 indicators between the 2 groups, and only one single item score was significantly different between the two groups. These results are in line with other findings, that have shown only marginal differences in patient experiences and satisfaction between patients in web-based and other modes.[2,3,7,8,9] There might be several reasons for the high correlation between the response modes. The surveys were designed to be as similar as possible, including the invitation letter, the content, layout and structure of the questionnaire and the timing of the first contact and reminders. The

invitations to the patients in Group A and Group B were sent the same week and non-respondents in both groups received two reminders.

The response rate alone is a poor predictor of nonresponse bias, and previous studies have failed to find a consistent association between response rates and sample representativeness.[23,24] However, low response rates threaten the legitimacy of surveys in both the clinical and public domains, and reduce the ability of surveys to identify important differences in patient-reported experiences between providers and over time.[2,3] Future research needs to focus on effective initiatives for increasing response rates in web-based protocols, including sending multiple reminders using a combination of emails, messages on mobile phones, and other available platforms. For example, the national infrastructure in Norway provides the possibility for secure digital mailboxes for all Norwegian inhabitants, which could be utilized for contacting digitally active patients.

The current study showed that a lower education level and higher age were associated with a mail preference. In the current study, we found that respondents invited by email were younger, more educated, and more likely to have had more-recent contact with their GP. We found no significant intergroup differences in the remaining nine background variables. There are several methods for assessing non-response bias, including comparison of respondents and non-respondents on background variables.[25] When we compared respondents with non-respondents, we found that men and younger patients were underrepresented as respondents in both groups. These differences are normally handled by non-response weighting, but such weights are only able to compensate for variables available in the sampling frame. We did not conduct further analysis of non-respondents, but previous follow-up studies of non-respondents in Norway indicate small additional bias.[12-15] However, none of these have

included a purely digital protocol, which warrant future non-response research for digital protocols. The coverage challenges for the digital sampling frame should be part of this research, as 12% of the population was not registered in the register, and 15% of the registered persons lacked a valid email address. This coverage challenge is an additional weakness of purely digital approaches and should be compensated with other response options for those excluded.

The effects of background characteristics reported in the literature are inconsistent.[2–10] The results from two previous randomized studies showed similar background characteristics for respondents in different randomized groups.[2,3] However, the respondents in those surveys were all contacted by mail. Future research should assess how the national infrastructure in Norway could be used to tailor the mode of data collection to different groups, such as by providing a range of data-collection modes from purely electronic strategies (for respondents with high education levels) to a mail-based mixed mode (to older respondents and those with low education levels).

Combined with the low response rates achieved for web-based protocols in this and other studies, future representative and high-quality surveys should include the opportunity to answer on pen-and-paper questionnaires. This could be implemented in a mixed-mode design that provides respondents with the option to choose how they want to respond, making it possible for patients without Internet access or enough computer skills to also participate.

A limitation of this study is that it only included adults evaluating their GPs in Norway, and so the results might not be generalizable to other health-care settings and countries. In particular, the national infrastructure and the digital maturity of the population in Norway

might differ from the characteristics of other countries. However, the results should be applicable to health systems with similar infrastructures and digital maturities, and to countries working to establish regional or national digital infrastructures. The survey was not linked to a specific contact with the GP or GP office, or actual use e.g. the last six months, which might have resulted in lower response rates and implies that we were unable to make any assumptions about specific contacts. Differences in respondent characteristics between respondents and non-respondents in both groups were tested, but not differences in patient reported experiences since we lacked a follow-up study of non-respondents. However, the impact of non-response bias in previous large-scale surveys have been relatively small.[12–15]

Conclusions

Administering a survey of patient experiences with GPs using a web-based protocol produced results that were very similar to those obtained using the standard mail-mode data-collection procedure used in the national surveys but had a much lower response rate. Furthermore, respondents in the digital group were younger, more educated, and had more-recent experiences with their GPs. Men and younger patients were underrepresented as respondents in both groups. Web-based surveys are faster and cheaper than standard mail surveys, but their low response rates threaten their legitimacy. Initiatives to increase response rates for web-based data collection and strategies for tailoring data collection to different groups should be key elements in future research.

Funding

This research received grant from the Norwegian Ministry of Health and Care Services (17/4584-12).

Competing interests

On behalf of all authors, the corresponding author states that there are no competing interests.

Author contributions

The Ministry of Health and Care Services initiated the study. H.H.I. planned the study in consultation with O.B. and O.H. H.H.I. performed the statistical analyses with O.B. and O.H. and drafted the manuscript. O.B. and O.H. participated in the planning process, critically revised the manuscript draft and approved the final version of the manuscript. H.H.I. was the project manager for the survey. All authors read and approved the final manuscript.

Acknowledgements

We thank Inger Opedal Paulsrud from the NIPH for her help in conducting the survey, including performing administrative and technical tasks for the data collection. We would also like to thank Sylvia Drange Sletten, Marius Troøyen and Sindre Møgster Braaten from the NIPH for help developing the system for web-based data collection.

Availability of data and material

The data set generated and/or analysed during this study is not publicly available due to the need to protect personal data.

References

1. Eide TB, Straand J, Rosvold EO. Patients' and GPs' expectations regarding healthcare-seeking behaviour: a Norwegian comparative study. BJGP Open. 2018;2(4):bjgpopen18X101615.
2. Bjertnaes OA, Iversen HH. User-experience surveys with maternity services: a randomized comparison of two data collection models. Int J Qual Health Care. 2012;24(4):433–8.
3. Bjertnaes O, Iversen HH, Skrivarhaug T. A randomized comparison of three data collection models for the measurement of parent experiences with diabetes outpatient care. BMC Med Res Methodol. 2018;18 (1):95.
4. Holmboe O, Bjertnæs OA. Patient experiences with Norwegian hospitals in 2015. Results following a national survey. Oslo: Norwegian Knowledge Centre for the Health Services, PasOpp-report 147; 2016. ISSN 1890-1565.
5. Holmboe O, Danielsen K, Iversen HH. Development of a method for the measurement of patient experiences with regular general practitioners. Oslo: Norwegian Knowledge Centre for the Health Services, PasOpp-report 1; 2015. ISSN 1890-1565.
6. Danielsen K, Holmboe O, Hansen SL, Førland O, Bjertnæs OA. Patient experiences with Norwegian primary care out-of-hours clinics. Oslo: Results from a survey among the "Vakttårn" clinics. Norwegian Knowledge Centre for the Health Services, PasOpp-report 1; 2012. ISSN 1890-1565.
7. Fowler FJ Jr, Cosenza C, Cripps LA, Edgman-Levitan S, Cleary PD. The effect of administration mode on CAHPS survey response rates and results: a comparison of mail and web-based approaches. Health Serv Res. 2019;54(3):714–21.
8. Bergeson SC, Gray J, Ehrmantraut LA, Laibson T, Hays RD. Comparing web-based with mail survey administration of the Consumer Assessment of Healthcare Providers

and Systems (CAHPS(®)) Clinician and Group survey. Prim Health Care. 2013;3:1000132.

9. Rodriguez HP, von Glahn 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(2):167–74.
10. Holmboe O, Iversen HH, Danielsen K, Bjertnaes O. The Norwegian patient experiences with GP questionnaire (PEQ-GP): reliability and construct validity following a national survey. BMJ Open. 2017;7(9):e016644.
11. Iversen HH, Bjertnaes O, Helland Y, Skrivarhaug T. The Adolescent Patient Experiences of Diabetes Care Questionnaire (APEQ-DC): Reliability and Validity in a Study Based on Data from the Norwegian Childhood Diabetes Registry. Patient Relat Outcome Meas. 2019;10:405-16.
12. Guldvog B, Hofoss D, Pettersen KI, Ebbesen J, Ronning OM: PS-RESKVA - pasienttilfredshet i sykehus [PS-RESKVA -patient satisfaction in hospitals]. Tidsskr Nor Laegeforen. 1998, 118: 386-391.
13. Dahle KA, Iversen H, Sjetne IS, Oltedal S: Do non-respondents' characteristics and experiences differ from those of respondents? Poster at the the 25th International Conference by the Society for Quality in Health Care (ISQua). 2008, Copenhagen.
14. Iversen H, Dahle KA, Holmboe O: Using telephone interviews to explore potential non-response bias in a national postal survey. Poster at the the 26th International Conference by the Society for Quality in Health Care (ISQua). 2009, Dublin.
15. Danielsen K, Iversen H, Holmboe O. Telephone interviews to assess non-response bias in a postal user experience survey: Kirsten Danielsen. European Journal of Public Health. 2015;25(suppl_3).

16. Kompetanse Norge. Befolkningens bruk av digitale verktøy. Available from: https://www.kompetansenorge.no/statistikk-og-analyse/grunnleggende-digital-ferdigheter/befolkingens-bruk-av-digitale-verktoy/#_[Accessed 10 June 2020].
17. Digitaliseringsdirektoratet. Befolkningens bruk av digitale verktøy. Kontakt- og reservasjonsregisteret. Available from: https://samarbeid.difi.no/statistikk/kontakt-og-reservasjonsregisteret_[Accessed 10 June 2020].
18. The American Association for Public Opinion Research. 2016. Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 9th edition. AAPOR. https://www.aapor.org/Standards-Ethics/Standard-Definitions-(1).aspx. [Accessed 10 June 2020].
19. Nunnally JC, Bernstein IH. Psychometric Theory. 3rd edn. New York: McGraw-Hill; 1994.
20. Kline RB. Principles and Practice of Structural Equation Modeling. New York: Guildford; 2005.
21. Bjertnaes OA, Lyngstad I, Malterud K, Garratt A. The Norwegian EUROPEP questionnaire for patient evaluation of general practice: data quality, reliability and construct validity. Fam Pract. 2011;28(3):342–9.
22. Ruiz MA, Pardo A, Rejas J, Soto J, Villasante F, Aranguren JL. Development and validation of the "Treatment Satisfaction with Medicines Questionnaire" (SATMED-Q). Value Health. 2008;11(5):913–26.
23. Groves RM, Peytcheva E. The impact of nonresponse rates on nonresponse bias: a meta-analysis. Public Opin Q. 2008;72(2):167–89.
24. Groves RM. Nonresponse rates and nonresponse bias in household surveys. Public Opin Q. 2006;70(5):646–75.

25. Halbesleben JRB, Whitman MV. Evaluating Survey Quality in Health Services Research: A Decision Framework for Assessing Nonresponse Bias. Health Services Research. 2013;48(3):913-30.

Figure 1: CONSORT flow diagram

Table 1: Respondents before and after each reminder in the two randomized groups, and final response rates.

	Group A (*n*=4,760)	Group B (*n*=1,694)
Respondents before reminder:		
Electronic, *n*	272	117
Paper, *n*	560	-
Response rate, %	17.5	6.9
Respondents after first reminder:		
Electronic, *n*	171	126
Paper, *n*	533	-
Increase in response rate, %	14.8	7.4
Respondents after second reminder:		
Electronic, *n*	148	67
Paper, *n*	345	-
Increase in response rate, %	10.4	4.0
Total:		
Electronic, *n* (%)	591 (29.1)	310
Paper, *n* (%)	1438 (70.9)	
Response rate, %	42.6	18.3

Table 2: Comparison of missing data, ceiling effects, and internal consistency between the two randomized groups.

Scale and item[a]	Group A				Group B			
	Missing data (%)	Not applicable (%)	Ceiling effects (%)	Cronbach's alpha/item-total correlation	Missing data (%)	Not applicable (%)	Ceiling effects (%)	Cronbach's alpha/item-total correlation
GP				**0.924**				**0.935**
Do you feel that your GP takes you seriously?	5.0	3.4	48.0	0.762	0.0	2.9	52.2	0.813
Do you feel that your GP spends enough time with you?	4.9	3.2	28.9	0.702	0.0	3.2	32.7	0.745
Do you feel that your GP talks to you in a way you understand?	5.1	3.0	48.9	0.737	0.3	2.3	56.0	0.735
Do you feel that your GP is professionally competent?	5.2	4.0	41.8	0.752	0.6	2.9	43.1	0.805
Do you feel that your GP shows interest in your situation?	5.3	3.5	39.5	0.818	0.6	2.3	39.9	0.853
Do you feel that your GP includes you as much as you would like in decisions concerning you?	5.7	7.7	37.2	0.769	0.3	6.5	41.9	0.803
Does your GP provide you with sufficient information about your health problems and their treatment?	5.6	7.3	32.7	0.811	0.3	6.1	34.1	0.835
Does your GP provide you with sufficient information about the use and side effects of medication?	5.4	16.3	21.0	0.631	0.0	16.5	24.3	0.655
Does your GP refer you to further examinations or a specialist when you feel you need it?	5.1	10.6	43.2	0.646	0.3	10.6	49.3	0.633
Organization and auxiliary staff				**0.868**				**0.851**
Do you feel that your GP's practice is well organized?	5.2	4.1	26.6	0.681	1.0	1.6	25.8	0.615
Do you feel the other employees are helpful and competent?	4.9	3.6	36.1	0.813	0.3	2.9	31.7	0.806
Are you treated with courtesy and respect at the reception?	4.8	3.2	40.9	0.752	0.3	2.6	39.5	0.750
Accessibility				**0.774**				**0.688**
Was the waiting time for your last urgent appointment acceptable?	18.7	-	36.2	0.631	17.1	-	33.5	0.525
Is this waiting time for appointments that are not urgent acceptable?	12.3	-	18.0	0.631	6.8	-	11.4	0.525
Enablement				**0.906**				**0.925**
Does contact with your GP make you better able to understand your health problems?	1.6	16.6	19.5	0.803	1.0	15.8	26.0	0.836
Does contact with your GP make you better able to cope with your health problems?	1.6	19.7	16.3	0.852	0.6	20.3	21.2	0.875
Does contact with your GP better help you to stay healthy?	1.6	19.8	15.1	0.786	0.3	20.6	21.6	0.833
Coordination and cooperation				**0.875**				**0.876**
Do you feel that your GP is good at coordinating the range of health services available to you?	5.9	26.9	28.1	0.779	0.6	30.6	34.7	0.790
Do you feel that your GP cooperates well with other services you need?	5.7	29.4	29.0	0.779	0.6	31.9	29.7	0.790

[a]All items were scored on a 5-point response scale ranging from 1 ("not at all") to 5 ("to a very large degree")

Table 3: Comparison of respondent characteristics between the two randomized groups.

	Group A	Group B	*p*[a]
Gender, female	55.9 (1135)	59.7 (185)	0.216
Age group			<0.001
16–19 years	2.4 (49)	2.3 (7)	
20–29 years	7.8 (158)	9.4 (29)	
30–49 years	23.8 (482)	37.4 (116)	
50–66 years	34.2 (694)	31.6 (98)	
≥67 years	31.8 (646)	19.4 (60)	
Time on the list of the GP			0.526
<1 year	9.4 (191)	8.4 (26)	
1–2 years	19.4 (392)	21.3 (66)	
3–4 years	14.5 (293)	16.1 (50)	
5–10 years	20.4 (414)	22.3 (69)	
≥11 years	36.3 (735)	31.9 (99)	
Number of consultations during past 12 months			0.672
0	9.3 (186)	10.1 (31)	
1	15.7 (314)	15.3 (47)	
2–5	55.7 (1114)	52.1 (160)	
6–12	16.2 (323)	19.2 (59)	
≥13	3.1 (63)	3.3 (10)	
Number of diagnosis codes during past 24 months	13.8±13.5	12.6±10.7	0.083
Number of unique diagnosis codes during past 24 months	4.7±3.2	4.6±2.8	0.510
Time since previous contact			0.042
<1 month	36.5 (716)	42.1 (128)	
1–3 months	32.0 (628)	23.7 (72)	
4–6 months	13.5 (266)	15.1 (46)	
7–12 months	9.7 (191)	8.9 (27)	
>12 months	8.3 (163)	10.2 (31)	
Self-perceived physical health			0.951
Very poor	1.3 (27)	1.6 (5)	
Quite poor	5.3 (108)	5.2 (16)	
Both poor and good	23.8 (481)	22.4 (69)	
Quite good	48.3 (975)	50.3 (155)	
Very good	21.2 (429)	20.5 (63)	
Self-perceived mental health			0.475
Very poor	1.1 (22)	1.9 (6)	
Quite poor	3.0 (60)	3.9 (12)	
Both poor and good	15.5 (313)	15.6 (48)	
Quite good	41.7 (842)	38.0 (117)	
Very good	38.7 (781)	40.6 (125)	
Long-standing health problems			0.625
0	35.7 (708)	37.5 (115)	
1	32.9 (653)	34.9 (107)	
2	19.3 (383)	16.9 (52)	
≥3	12.1 (241)	10.7 (33)	
Education level			<0.001
Elementary school	15.6 (309)	7.1 (22)	
High school	37.4 (740)	31.4 (97)	
University, 0–4 years	25.6 (505)	35.3 (109)	
University, >4 years	21.4 (422)	26.2 (81)	
Geographic origin			0.205
Norway	88.6 (1756)	89.9 (276)	
Asia (incl. Turkey), Africa, or Latin America	4.8 (95)	3.3 (10)	
Eastern Europe (all countries, independent of EU membership)	3.5 (70)	2.3 (7)	
Western Europe, North America, or Oceania	3.0 (60)	4.6 (14)	

[a] Pearson chi-square tests for categorical variables and independent-samples *t*-tests for continuous variables. Data are *n* (%) or mean±SD values.

Table 4: Comparison of respondent characteristics between the two randomized groups.

	Group A			Group B		
	Respondents	Non-respondents	*p*[a]	Respondents	Non-respondents	*p*[a]
Gender, female	55.9 (1135)	46.1 (1259)	<0.001	59.7 (185)	49.8 (689)	0.002
Age group			<0.001			<0.001
16–19 years	2.4 (49)	7.3 (200)		2.3 (7)	6.4 (88)	
20–29 years	7.8 (158)	20.7 (566)		9.4 (29)	17.6 (243)	
30–49 years	23.8 (482)	39.5 (1080)		37.4 (116)	38.8 (537)	
50–66 years	34.2 (694)	21.2 (579)		31.6 (98)	26.0 (360)	
≥67 years	31.8 (646)	11.2 (306)		19.4 (60)	11.3 (156)	
Time on the list of the GP			<0.001			0.739
<1 year	9.4 (191)	10.5 (288)		8.4 (26)	9.6 (133)	
1–2 years	19.4 (392)	24.3 (664)		21.3 (66)	23.3 (322)	
3–4 years	14.5 (293)	15.3 (419)		16.1 (50)	13.7 (189)	
5–10 years	20.4 (414)	20.9 (570)		22.3 (69)	22.3 (309)	
≥11 years	36.3 (735)	28.9 (790)		31.9 (99)	31.1 (431)	
Number of consultations during past 24 months	10.8±11.3	7.6±10.7	<0.001	9.6±9.2	8.4±11.5	0.077
Number of diagnosis codes during past 24 months	13.8±13.5	11.3±14.1	<0.001	12.6±10.7	12.2±14.9	0.611
Number of unique diagnosis codes during past 24 months	4.7±3.2	4.1±3.1	<0.001	4.6±2.8	4.2±3.2	0.107

[a] Pearson chi-square tests for categorical variables and independent-samples *t*-tests for continuous variables. Data are *n* (%) or mean±SD values.

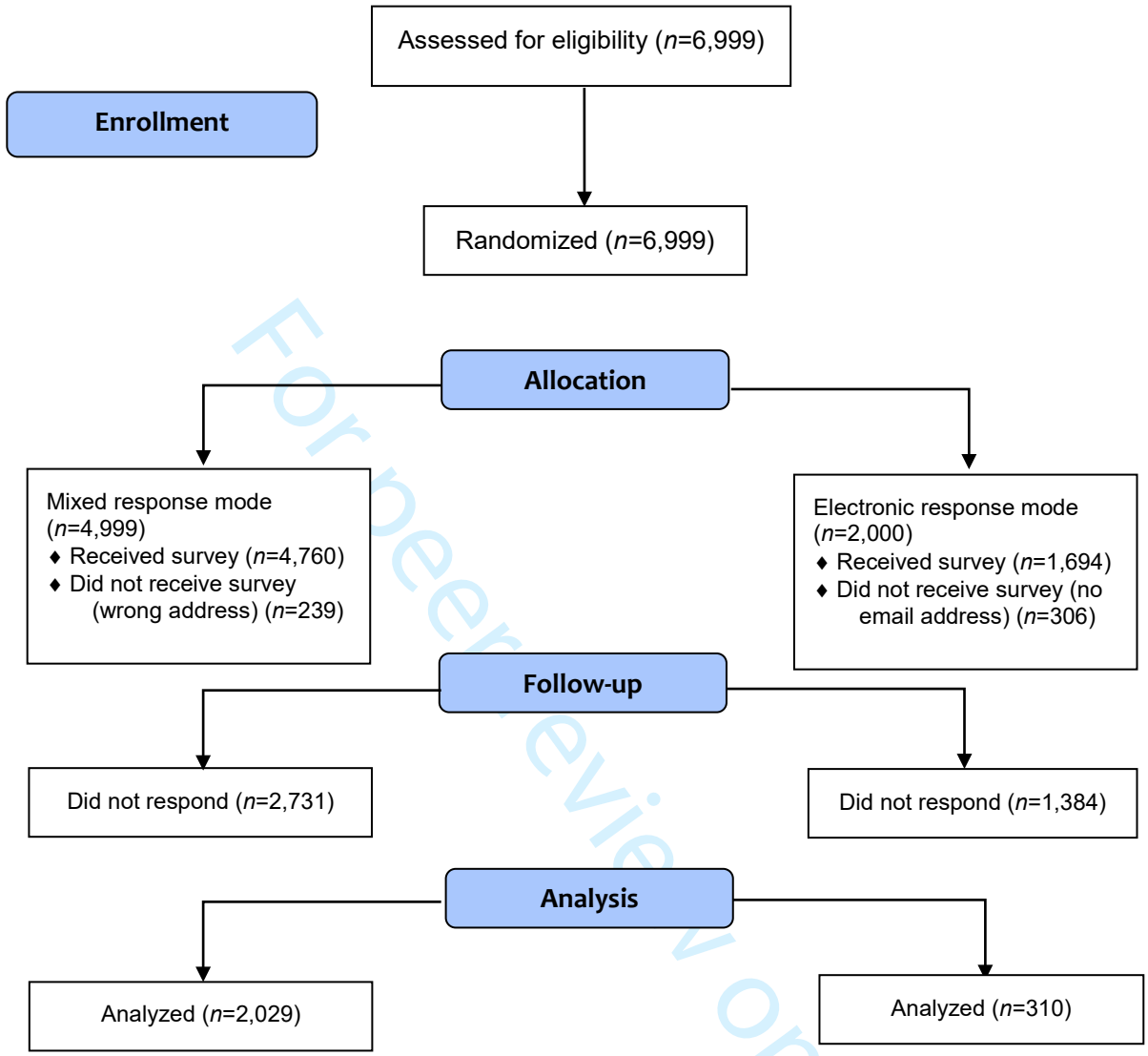
Table 5: Comparison of patient-reported experiences between the two randomized groups.

Scale and item[a]	Group A	Group B	*P*
GP	**78.3±16.8**	**78.8±17.8**	**0.651**
Do you feel that your GP takes you seriously?	83.1±19.8	84.1±20.3	0.429
Do you feel that your GP spends enough time with you?	73.8±22.8	73.3±25.8	0.707
Do you feel that your GP talks to you in a way you understand?	84.6±17.7	85.8±19.1	0.275
Do you feel that your GP is professionally competent?	82.1±18.0	81.8±19.4	0.747
Do you feel that your GP shows interest in your situation?	79.7±20.5	78.7±22.2	0.435
Do you feel that your GP includes you as much as you would like in decisions concerning you?	79.0±20.3	80.0±21.3	0.452
Does your GP provide you with sufficient information about your health problems and their treatment?	76.6±21.2	76.1±22.2	0.708
Does your GP provide you with sufficient information about the use and side effects of medication?	65.0±26.9	67.1±26.0	0.241
Does your GP refer you to further examination or a specialist when you feel you need it?	81.4±20.2	82.0±22.2	0.712
Organization and auxiliary staff	**78.2±17.7**	**77.1±18.2**	**0.322**
Do you feel that your GP's practice is well organized?	75.2±20.2	74.8±20.4	0.735
Do you feel the other employees are helpful and competent?	79.3±19.2	76.9±20.8	0.046
Are you treated with courtesy and respect at the reception?	80.9±19.7	79.9±20.7	0.430
Accessibility	**63.6±27.8**	**61.8±25.5**	**0.264**
Was the waiting time for your last urgent appointment acceptable?	69.5±30.6	69.1±30.4	0.828
Is this waiting time for appointments that are not urgent acceptable?	58.3±30.0	54.8±28.5	0.051
Enablement	**65.2±22.1**	**66.0±24.3**	**0.601**
Does contact with your GP make you better able to understand your health problems?	68.1±23.0	68.7±25.1	0.703
Does contact with your GP make you better able to cope with your health problems?	64.9±24.0	65.9±25.6	0.560
Does contact with your GP better help you to stay healthy?	62.7±25.0	64.5±26.7	0.303
Coordination and cooperation	**74.3±21.0**	**74.9±21.5**	**0.644**
Do you feel that your GP is good at coordinating the range of health services available to you?	75.0±21.2	77.7±20.5	0.079
Do you feel that your GP cooperates well with other services you need?	74.3±22.6	73.0±24.5	0.434

[a]All scales and items are scored from 0 to 100, where 100 is the best possible patient experience. Data are mean±SD values

.

Figure 1: CONSORT flow diagram



Reporting checklist for cross sectional study.

Based on the STROBE cross sectional guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the STROBE cross sectionalreporting guidelines, and cite them as:

von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies.

		Reporting Item	Page Number
Title and abstract			
Title	#1a	Indicate the study's design with a commonly used term in the title or the abstract	2
Abstract	#1b	Provide in the abstract an informative and balanced summary of what was done and what was found	2,3
Introduction			
Background / rationale	#2	Explain the scientific background and rationale for the investigation being reported	5-6
Objectives	#3	State specific objectives, including any prespecified hypotheses	2,6
Methods			
Study design	#4	Present key elements of study design early in the paper	2, 6/7
Setting	#5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6-8
Eligibility criteria	#6a	Give the eligibility criteria, and the sources and methods of selection of participants.	2, 6-8

	#7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7,8
Data sources / measurement	#8	For each variable of interest give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. Give information separately for for exposed and unexposed groups if applicable.	7,8
Bias	#9	Describe any efforts to address potential sources of bias	11-13
Study size	#10	Explain how the study size was arrived at	6,7
Quantitative variables	#11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	6-8
Statistical methods	#12a	Describe all statistical methods, including those used to control for confounding	7-8
Statistical methods	#12b	Describe any methods used to examine subgroups and interactions	6-8
Statistical methods	#12c	Explain how missing data were addressed	7,8,19
Statistical methods	#12d	If applicable, describe analytical methods taking account of sampling strategy	6
Statistical methods	#12e	Describe any sensitivity analyses	n/a
Results			
Participants	#13a	Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed. Give information separately for for exposed and unexposed groups if applicable.	6,7,17
Participants	#13b	Give reasons for non-participation at each stage	17
Participants	#13c	Consider use of a flow diagram	17
Descriptive data	#14a	Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders. Give information separately for exposed and unexposed groups if applicable.	9,10,20
Descriptive data	#14b	Indicate number of participants with missing data for each variable of interest	19
Outcome data	#15	Report numbers of outcome events or summary measures. Give information separately for exposed and unexposed groups if applicable.	18,19,21
Main results	#16a	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	n/a

Main results	#16b	Report category boundaries when continuous variables were categorized	n/a
Main results	#16c	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	#17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	n/a
Discussion			
Key results	#18	Summarise key results with reference to study objectives	10,11
Limitations	#19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.	12,13
Interpretation	#20	Give a cautious overall interpretation considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence.	10-13
Generalisability	#21	Discuss the generalisability (external validity) of the study results	12,13
Other Information			
Funding	#22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	13