American Journal of Epidemiology
Copyright © 1998 by The Johns Hopkins University School of Hygiene and Public Health
All rights reserved
Vol. 147, No. 1
Printed in U.S.A.

# Response Rate to Mailed Epidemiologic Questionnaires: A Population-based Randomized Trial of Variations in Design and Mailing Routines

Sonja Eaker,[1] Reinhold Bergström,[1,2] Anna Bergström,[1] Hans-Olov Adami,[1,3] and Olof Nyren[1]

Although self-administered questionnaires are major sources of information in epidemiology, comparatively little has been done to study practical aspects of design and mailing. The objective of this study was to evaluate various measures taken to increase the response rate. A questionnaire was mailed in July 1995 to a random sample ($n$ = 2,000) of the Swedish population aged 20–79 years. Using a randomized factorial study design, the questionnaire and mailing procedures were changed in three ways: preliminary notification, length of the questionnaire, and mention of telephone contact. The overall questionnaire retrieval rate was 49%. Preliminary notification (adjusted odds ratio of receiving a completed questionnaire = 1.30, 95% confidence interval (CI) 1.08–1.56 relative to the absence of preliminary notification) and short length of the questionnaire (odds ratio = 1.24, 95% CI 1.04–1.48 relative to a long questionnaire) were both independently associated with a higher retrieval rate. Of eight possible combinations, the one comprising preliminary notification, a short questionnaire, and no mention of telephone contact gave the highest retrieval rate, 56%. The lowest retrieval rate, 40%, was observed for the combination of no preliminary notification, a long questionnaire, and mention of telephone contact. Young age, male sex, and urban residence significantly lowered the retrieval rate. Although there was a positive association between the questionnaire retrieval rate and partial nonresponse (missing answers in retrieved questionnaires), the marginal losses due to the latter did not cancel the gains by optimized mailing routines. Old age was the strongest determinant of partial nonresponse. The data provide evidence that design and mailing strategies, as well as demographic characteristics, may greatly influence the response rate of mailed epidemiologic questionnaires. *Am J Epidemiol* 1998;147:74–82.

epidemiologic methods; questionnaires

Mailed questionnaires are standard tools in epidemiologic research. Questionnaire studies are less time-consuming than telephone and personal interviews, free from the influence of interviewers, and relatively cheap. However, mailed questionnaires often have a lower response rate than telephone and personal interviews (1). A high response rate may be of critical importance for validity and cost efficiency; a low rate may cause selection bias in case-control studies and undermine the follow-up in prospective studies.

Several studies in sociology and marketing research have evaluated factors that may affect the response rate, such as preliminary notification (1–5) and length of the questionnaire (1, 3–6). Epidemiologic studies may be perceived by the public as particularly important, but their questions are often complex and may be difficult to answer. Therefore, the response pattern in such studies cannot be inferred from research in other areas. Until now, epidemiologic studies on these matters have generally been performed in highly selected groups of individuals (7–9) and have dealt mainly with the importance of demographic factors, such as age (7), sex, and urban dwelling (10)—factors that cannot be manipulated to improve the response rate. Comparatively little effort has been devoted to studying technical aspects of the administration of mailed epidemiologic questionnaires. Exceptions are studies dealing with the effects of reminders and economic incentives to the participants (9–12) and of various types of stamps on return envelopes (13).

Our aim was to test, with an experimental design, in a population-based sample, several practically relevant modifications of a questionnaire study. A factorial design allowed us to study the effects, both isolated and in combination. Our population-based sample of men and women (20–79 years old) in Sweden was representative of the whole population.

Received for publication October 28, 1996, and accepted for publication July 16, 1997.

Abbreviation: CI, confidence interval.

[1] Department of Medical Epidemiology, Karolinska Institutet, Stockholm, Sweden.
[2] Department of Statistics, Uppsala University, Uppsala, Sweden.
[3] Department of Epidemiology, Harvard School of Public Health, Boston, MA.
Reprint requests to Sonja Eaker, Department of Medical Epidemiology, Karolinska Institutet, Box 281, S-17177 Stockholm, Sweden.



Downloaded from https://academic.oup.com/aje/article/147/1/74/50426 by guest on 20 August 2022

## MATERIALS AND METHODS

### Subjects

The source population comprised all men and women 20–79 years of age living in Sweden in May 1995. By means of the continuously updated computerized population register, a stratified random sample was drawn in two age groups (20–49 and 50–79 years), with 1,000 subjects selected from each stratum. The sample consisted of 49 percent women and 51 percent men. Besides age and sex, each subject was characterized by population size of residence, based on postal area codes, namely, metropolitan areas, cities, towns, and semirural and rural areas.

### Questionnaire

The questionnaire concerned height, weight, diet, medical history and medication, physical activity, eating and drinking habits, tobacco use, and certain background factors such as education and year and country of birth. Women were also asked about their reproduction history (pregnancies, menarche, and menopause), and men were asked about their age at voice change. All participants received the same type of questionnaire, a letter that explained the study, and a prepaid return envelope. As this study was done as part of the preparations for a case-control study of kidney cancer, the alleged purpose given in the accompanying letter was to study risk factors for this cancer. The questionnaires were sent out as first-class mail delivered within 1 day, on two different weekdays in the beginning of July 1995. By doing so, we investigated whether the response rate differed between those who received the questionnaire at the beginning (Tuesday) and at the end (Friday) of the week. A reminder was mailed to all 2,000 subjects in the sample 1 week after mailing the questionnaire. No other reminders were sent out.

### Study design

Three factors were varied according to a randomized $2^3$ factorial design. The first factor tested was sending or not sending a preliminary notification, an "alert" letter that preceded the mailing of the questionnaire by about 1 week and in which the study was presented and the questionnaire announced.

The second factor tested was the length of the questionnaire. A long version included detailed questions about physical activity (at work, at home, and during leisure time) and food consumption. A short version had less complex and shorter questions about physical activity and excluded all questions about food consumption. The long questionnaire consisted of 15 pages (comprising 66 items) for men and 18 pages (79 items) for women, and the shorter version consisted of 11 and 14 pages (50 and 66 items), respectively.

The third factor was the inclusion or not of a clause that mentioned the possibility of a future telephone contact requesting supplementary information if the questionnaire was incompletely answered. Those who received a questionnaire in which the telephone contact was mentioned were asked to provide their telephone number at the end of the questionnaire.

The three factors produced $2^3 = 8$ experimental conditions (treatments) (table 1). Two hundred fifty subjects were randomly allocated to each condition, resulting in 1,000 subjects allocated to each factor modification.

### End of study

We noted the date when each questionnaire was returned. The proportion of all questionnaires retrieved was our main outcome variable. To be defined as "retrieved," the questionnaire had to be returned and had to contain some information, which means that at least one question had to be answered. Subjects who had not answered within 75 days of the initial mailing were considered nonresponders.

### Partial nonresponse

We concentrated our analysis of partial nonresponse (missing answers in returned questionnaires) on questions that were common to all questionnaire types (31 questions for women and 28 for men). Nine questionnaires returned could not be analyzed in this phase: two because of problems with identification (one with incorrect national registration number and one with the first page missing), five because women were erroneously given questionnaires for men, and two because men were erroneously given questionnaires for women.

### Statistical methods

In the analyses of the overall questionnaire retrieval rate, we used logistic regression, assuming that the logarithm of the odds of returning the questionnaire is a linear function of the explanatory variables. The model was estimated by the maximum likelihood method; odds ratios with 95 percent confidence intervals were computed from the estimated parameters and standard errors. Univariate and multivariate models were estimated. The latter included the following variables: age (in 10-year classes), sex, population size (five categories), weekday of mailing, and experimen-



Downloaded from https://academic.oup.com/aje/article/147/1/74/50426 by guest on 20 August 2022

tal condition. The latter variable was included in the form of a categorized variable with eight categories but also as a separate term for each of the three factors, preliminary/no preliminary notification, short/long questionnaire, and mention/no mention of telephone contact, reflecting the factorial design of the study. Not only models with main effects but also those with interaction effects were estimated. Likelihood ratio tests were used to test different nested models against one another. On the basis of univariate analyses, we show confidence intervals for the difference between proportions in certain cases. These intervals were computed using a normal approximation.

In the modeling of partial nonresponse, we used standard linear regression models, estimated by ordinary least squares. The proportion of questions with nonresponse for each respondent was used as the dependent variable. As the basic dependent variable (number of questions not answered) is a variable of count data type, an analysis based on Poisson models would in principle have been preferrable. However, as closer inspection of the data indicated that the Poisson assumption was doubtful, we preferred the ordinary least-squares estimates, which require no distributional assumption for a sample of the present size. Estimation using Poisson models produced qualitatively similar results.

## RESULTS

We received completed questionnaires from 975 of the 2,000 selected subjects, corresponding to an overall retrieval rate of 49 percent. Of the three randomized factors, preliminary notification had the greatest impact (table 1). This increased the retrieval rate significantly, that is, by 7 percentage units (95 percent confidence interval (CI) 3 to 11), from 45 percent with no preliminary notification to 52 percent. A short questionnaire increased the retrieval rate by 5 percentage units (95 percent CI −0 to 9), 51 percent versus 46 percent with the long questionnaire (table 1). The effect was of borderline significance; the odds ratio for receiving a completed questionnaire was 1.21 (95 percent CI 1.01 to 1.44). Mention of the possibility of a telephone contact caused a 4 percentage-unit (95 percent CI −7 to 8) lower retrieval rate, 47 percent versus 51 percent, than no mention of a telephone contact (not significant). The highest retrieval rate, with the three factors combined, was observed in category 4 (preliminary notification, the shorter version of the questionnaire, and no mention of telephone contact), and the lowest in its opposite category 5 (no preliminary notification, the long questionnaire, and mention of telephone contact) (table 2). The response rate in category 4 was 16 percentage units higher than in category 5, 56 percent versus 40 percent (i.e., almost a 40 percent increase) (table 2).

A test for interaction among the three random assignments was not significant; using a likelihood ratio test, we obtained $\chi^2_{(3)} = 4.17$ ($p > 0.05$), which means that we can largely rely on the main-effects model. Factorial design categories 1 and 8 tended to have higher response rates than expected on the basis of the main-effects model (table 2).

As expected in a randomized trial of sufficient size, the results of multivariate modeling (which included the additional variables of age categorized in 10-year classes, sex, population size, and weekday of mailing) were similar to those of the univariate modeling. Thus, our findings were not confounded by these factors (tables 1 and 3).

**TABLE 1. Odds ratios and 95% confidence intervals of receiving a completed questionnaire with experimental modifications in three factors, Sweden, 1995**

| | Sample size (no.) | Respondents | | Univariate modeling | | Multivariate modeling | |
|---|---|---|---|---|---|---|---|
| | | No. | % | OR* | 95% CI* | OR | 95% CI† |
| Preliminary notification | | | | | | | |
| Yes | 1,000 | 522 | 52.2 | 1.32 | 1.11–1.57 | 1.30 | 1.08–1.56 |
| No | 1,000 | 453 | 45.3 | 1.00 | (ref.) | 1.00 | (ref.) |
| Short questionnaire | | | | | | | |
| Yes | 1,000 | 511 | 51.1 | 1.21 | 1.01–1.44 | 1.24 | 1.04–1.48 |
| No | 1,000 | 464 | 46.4 | 1.00 | (ref.) | 1.00 | (ref.) |
| Telephone | | | | | | | |
| Yes | 1,000 | 469 | 46.9 | 1.00 | (ref.) | 1.00 | (ref.) |
| No | 1,000 | 506 | 50.6 | 1.16 | 0.97–1.38 | 1.14 | 0.95–1.37 |

* OR, odds ratio; CI, confidence interval.

† This model included the variables: allocation group of each factor, age categories in 10-year classes, sex, population size, and weekday of mailing.



Downloaded from https://academic.oup.com/aje/article/147/1/74/50426 by guest on 20 August 2022

**TABLE 2. Odds ratios and 95% confidence intervals of receiving a completed questionnaire by experimental condition according to the factorial design, Sweden, 1995**

| Category | Sample size (no.) | Respondents | | Comparison with reference category | | Univariate modeling | |
|---|---|---|---|---|---|---|---|
| | | No. | % | Difference* | 95% CI† | OR† | 95% CI |
| 1. Preliminary notification<br>Long questionnaire<br>Mention of telephone contact | 250 | 130 | 52.0 | 11.6 | 2.9 to 20.3 | 1.60 | 1.12 to 2.28 |
| 2. Preliminary notification<br>Short questionnaire<br>Mention of telephone contact | 250 | 128 | 51.2 | 10.8 | 2.1 to 19.5 | 1.55 | 1.09 to 2.21 |
| 3. Preliminary notification<br>Long questionnaire<br>No mention of telephone contact | 250 | 124 | 49.6 | 9.2 | 0.5 to 17.9 | 1.45 | 1.02 to 2.07 |
| 4. Preliminary notification<br>Short questionnaire<br>No mention of telephone contact | 250 | 140 | 56.0 | 15.6 | 6.9 to 24.3 | 1.88 | 1.32 to 2.68 |
| 5. No preliminary notification<br>Long questionnaire<br>Mention of telephone contact | 250 | 101 | 40.4 | 1.00 | (ref.) | 1.00 | (ref.) |
| 6. No preliminary notification<br>Short questionnaire<br>Mention of telephone contact | 250 | 110 | 44.0 | 3.6 | −5.1 to 12.3 | 1.16 | 0.81 to 1.65 |
| 7. No preliminary notification<br>Long questionnaire<br>No mention of telephone contact | 250 | 109 | 43.6 | 3.2 | −5.4 to 11.8 | 1.14 | 0.80 to 1.63 |
| 8. No preliminary notification<br>Short questionnaire<br>No mention of telephone contact | 250 | 133 | 53.2 | 12.8 | 4.1 to 21.5 | 1.68 | 1.18 to 2.39 |

* Difference in percentage of responders in comparison with reference category, number 5.
† CI, confidence interval; OR, odds ratio.

We also analyzed the effect of the design variables in different strata of the background variables to see if there was any effect modification (table 4). The only statistically significant effect modification, albeit barely significant at the 5 percent level, was with regard to sex. Men seemed to be particularly attracted by a short questionnaire. Nonsignificant but quite large differences were also found for age, where people older than 40 years of age seemed to be more sensitive to preliminary notification, and for weekday of mailing, where preliminary notification appeared to be more useful when the questionnaire reached its addressee on Fridays.

The effects of the design modifications on partial nonresponse were moderate. Preliminary notification was associated with a slightly higher and statistically nonsignificant mean proportion (0.9 percentage units) of missing answers (table 5) compared with no preliminary notification. The longer questionnaire entailed no significantly higher proportion of missing answers than did the short questionnaire nor did mention of a telephone contact seriously affect the partial nonresponse. When the effects were combined, category 5 (no preliminary notification, a long questionnaire, and mention of telephone contact), which showed the lowest questionnaire retrieval rate, had the lowest proportion of missing answers (table 6). Although the differences between the means were, on the whole, not large (up to 2 percentage units, corresponding to a 59 percent increase of the proportion of missing answers), multivariate analyses showed that categories 1 and 4 had a significantly greater number of missing answers relative to category 5.

We related the questionnaire retrieval rates to the mean proportions of missing answers for the eight observation points constituted by the various experimental conditions and found a strong positive correlation (Pearson's correlation coefficient = 0.82, $p$ = 0.01). Thus, the gains from a higher questionnaire retrieval rate were to some extent counteracted by a



Downloaded from https://academic.oup.com/aje/article/147/1/74/50426 by guest on 20 August 2022

TABLE 3. Odds ratios and 95% confidence intervals of receiving a completed questionnaire by demographic factors and weekday of mailing, Sweden, 1995

| | Sample size (no.) | Respondents | | Univariate modeling | | Univariate modeling | |
|---|---|---|---|---|---|---|---|
| | | No. | % | OR* | 95% CI* | OR | 95% CI† |
| Age (years) | | | | | | | |
| 20–24 | 160 | 59 | 36.9 | 1.00 | (ref.) | 1.00 | (ref.) |
| 25–29 | 171 | 64 | 37.4 | | | | |
| 30–34 | 171 | 67 | 39.2 | 1.21 | 0.89–1.65 | 1.23 | 0.90–1.69 |
| 35–39 | 160 | 71 | 44.4 | | | | |
| 40–44 | 160 | 85 | 53.1 | 2.09 | 1.54–2.85 | 2.06 | 1.50–2.81 |
| 45–49 | 178 | 102 | 57.3 | | | | |
| 50–54 | 231 | 132 | 57.1 | 1.89 | 1.40–2.53 | 1.91 | 1.41–2.58 |
| 55–59 | 175 | 82 | 46.9 | | | | |
| 60–64 | 155 | 81 | 52.3 | 1.99 | 1.45–2.73 | 1.97 | 1.43–2.71 |
| 65–69 | 154 | 86 | 55.8 | | | | |
| 70–74 | 155 | 89 | 57.4 | 1.78 | 1.29–2.45 | 1.70 | 1.22–2.35 |
| 75–79 | 130 | 57 | 43.8 | | | | |
| Sex | | | | | | | |
| Men | 1,025 | 457 | 44.6 | 1.00 | (ref.) | 1.00 | (ref.) |
| Women | 975 | 518 | 53.1 | 1.41 | 1.18–1.68 | 1.39 | 1.16–1.66 |
| Population size | | | | | | | |
| Metropolitan areas | 595 | 268 | 45.0 | 1.00 | (ref.) | 1.00 | (ref.) |
| Cities | 273 | 133 | 48.7 | 1.16 | 0.97–1.55 | 1.17 | 0.87–1.56 |
| Towns | 384 | 195 | 50.8 | 1.26 | 0.97–1.63 | 1.28 | 0.99–1.67 |
| Semirural areas | 323 | 159 | 49.2 | 1.18 | 0.90–1.55 | 1.16 | 0.88–1.53 |
| Rural areas | 425 | 220 | 51.8 | 1.31 | 1.02–1.68 | 1.26 | 0.97–1.62 |
| Weekday | | | | | | | |
| Tuesday | 1,000 | 505 | 50.5 | 1.00 | (ref.) | 1.00 | (ref.) |
| Friday | 1,000 | 470 | 47.0 | 0.87 | 0.73–1.04 | 0.88 | 0.74–1.06 |
| Total | 2,000 | 975 | 48.8 | | | | |

* OR, odds ratio; CI, confidence interval.
† This model included the variables: experimental condition, age categorized in 10-year classes, sex, population size, and weekday of mailing.

higher mean proportion of missing answers. However, since the gains in questionnaire retrieval were much larger than the losses due to partial nonresponse, the net effects of the studied factors on total response (i.e., the total number of answered questions that we received from the 2,000 subjects) were essentially the same as the effects on questionnaire retrieval (data not shown).

## Demographic characteristics

Age influenced the retrieval rate markedly (table 3). Less than 40 percent of the subjects under the age of 35 returned their questionnaire. For the 40- to 74-year age group, the retrieval rate varied between 47 and 57 percent, that is, up to 20 percentage units higher than in the youngest age group. In the oldest age group, the retrieval rate was again slightly lower (44 percent). A model with age in continuous form including a second-order term revealed a significant effect of the second-order term, confirming the nonlinear relation. Overall, the retrieval rate from women was 8 percentage units (95 percent CI 4 to 13) higher than among men, 53 percent compared with 45 percent (table 3). The population size also affected the retrieval rate significantly. In rural areas, the retrieval rate was 7 percentage units (95 percent CI 1 to 13) higher than in metropolitan areas, 52 percent compared with 45 percent (table 3). The retrieval rate among those who received the questionnaire on a Tuesday was 4 percentage units (95 percent CI −1 to 8) higher than for those who received it on a Friday, 51 percent versus 47 percent (table 3). This difference was not significant.

Downloaded from https://academic.oup.com/aje/article/147/1/74/50426 by guest on 20 August 2022

**TABLE 4. Stratified analyses of the effects of the experimental treatments on the odds ratios of receiving completed questionnaires, Sweden, 1995*,†**

| Variable | Sample size (no.) | Preliminary notification‡ | | Short questionnaire§ | | Mention of telephone contact¶ | | Test for interaction | |
|---|---|---|---|---|---|---|---|---|---|
| | | OR# | 95% CI# | OR | 95% CI | OR | 95% CI | $\chi^{2**}$ | *p* value |
| Age (years) | | | | | | | | | |
| ≤39 | 662 | 1.04 | 0.76–1.42 | 1.43 | 1.04–1.95 | 1.19 | 0.87–1.62 | 4.34 | NS# |
| ≥40 | 1,338 | 1.46 | 1.17–1.81 | 1.15 | 0.93–1.43 | 1.14 | 0.91–1.41 | | |
| Sex | | | | | | | | | |
| Men | 1,025 | 1.37 | 1.07–1.76 | 1.54 | 1.20–1.98 | 1.07 | 0.83–1.37 | 7.96 | 0.048 |
| Women | 975 | 1.27 | 0.98–1.63 | 0.96 | 0.74–1.22 | 1.23 | 0.95–1.58 | | |
| Population size | | | | | | | | | |
| Metropolitan areas | 595 | 1.33 | 0.96–1.84 | 1.24 | 0.89–1.71 | 1.16 | 0.84–1.60 | 0.04 | NS |
| All other areas | 1,405 | 1.31 | 1.06–1.62 | 1.19 | 0.97–1.47 | 1.16 | 0.94–1.44 | | |
| Weekday | | | | | | | | | |
| Tuesday | 1,000 | 1.17 | 0.91–1.49 | 1.30 | 1.02–1.67 | 1.09 | 0.85–1.40 | 3.13 | NS |
| Friday | 1,000 | 1.50 | 1.17–1.92 | 1.12 | 0.87–1.44 | 1.24 | 0.96–1.59 | | |

* Stratification by age, sex, population size, and weekday.
† Each odds ratio given in the table is relative to the reference category for the respective experimental treatment (see table 1), and mutually adjusted results are given.
‡ Reference category = no preliminary notification.
§ Reference category = short questionnaire.
¶ Reference category = mention of telephone contact.
# OR, odds ratio; CI, confidence interval; NS, not significant.
** Test for interaction between experimental conditions and the various background variables.

**TABLE 5. Partial nonresponse: multivariate linear regression analysis of factors related to the percentage of missing answers in returned questionnaires, Sweden, 1995**

| Variable | Respondents | | Multivariate modeling* | | |
|---|---|---|---|---|---|
| | No. | Mean† | $\hat{\beta}$ | SE‡ ($\hat{\beta}$) | *p* value |
| Preliminary notification | | | | | |
| Yes | 518 | 5.5 | 0.95 | 0.57 | NS‡ |
| No | 448 | 4.6 | 0.00 | (ref.) | |
| Short questionnaire | | | | | |
| Yes | 503 | 5.1 | −0.13 | 0.58 | NS |
| No | 463 | 5.0 | 0.00 | (ref.) | |
| Telephone | | | | | |
| Yes | 461 | 4.7 | 0.00 | (ref.) | |
| No | 505 | 5.4 | 0.87 | 0.57 | NS |

* This model included the variables: allocation group of each factor, age categorized in 10-year classes, and sex.
†Mean percentage of answers missing.
‡SE, standard error; NS, not significant.

The greatest differences in partial nonresponse were found among the various age groups (table 6). The mean percentage of missing answers in the oldest group, 70–79 years, was more than 5 percentage units higher than in the groups aged 30–49 years, who showed the lowest proportion of missing answers. No large differences in partial nonresponse were found between men and women (table 6).

Those who answered only after a reminder had a significantly (1 percentage unit) higher partial nonresponse than did the others. This effect disappeared after adjustment for age, since older people were overrepresented among those who required a reminder to answer. In a model containing both the reminder variable and the three manipulated factors, only the reminder variable turned out to be significant. Further adjustment for age eliminated the effect of this variable (data not shown).

To assess the robustness of our results, we performed a sensitivity analysis. The replies of individu-



Downloaded from https://academic.oup.com/aje/article/147/1/74/50426 by guest on 20 August 2022

**TABLE 6. Partial nonresponse: multivariate linear regression analysis of factors related to the percentage of missing answers in returned questionnaires, Sweden, 1995**

| Variable | Respondents | | Multivariate analysis* | | |
|---|---|---|---|---|---|
| | No. | Mean† | $\hat{\beta}$ | SE‡ $(\hat{\beta})$ | *p* value |
| Category | | | | | |
| 1. Preliminary notification<br>Long questionnaire<br>Mention of telephone contact | 130 | 5.9 | 2.49 | 1.17 | 0.034 |
| 2. Preliminary notification<br>Short questionnaire<br>Mention of telephone contact | 124 | 5.0 | 1.21 | 1.19 | NS‡ |
| 3. Preliminary notification<br>Long questionnaire<br>No mention of telephone contact | 124 | 5.0 | 1.19 | 1.18 | NS |
| 4. Preliminary notification<br>Short questionnaire<br>No mention of telephone contact | 140 | 5.9 | 2.32 | 1.15 | 0.045 |
| 5. No preliminary notification<br>Long questionnaire<br>Mention of telephone contact | 101 | 3.7 | 0.00 | (ref.) | |
| 6. No preliminary notification<br>Short questionnaire<br>Mention of telephone contact | 106 | 3.7 | −0.32 | 1.23 | NS |
| 7. No preliminary notification<br>Long questionnaire<br>No mention of telephone contact | 108 | 5.2 | 1.87 | 1.23 | NS |
| 8. No preliminary notification<br>Short questionnaire<br>No mention of telephone contact | 133 | 5.3 | 1.74 | 1.17 | NS |
| Age (years) | | | | | |
| 20–29 | 120 | 4.2 | 0.00 | (ref.) | |
| 30–39 | 136 | 3.4 | −0.86 | 1.11 | NS |
| 40–49 | 186 | 3.3 | −1.16 | 1.04 | NS |
| 50–59 | 213 | 4.4 | 0.00 | 1.01 | NS |
| 60–69 | 166 | 6.6 | 2.41 | 1.06 | 0.0232 |
| 70–79 | 145 | 8.6 | 4.42 | 1.09 | <0.0001 |
| Sex | | | | | |
| Men | 453 | 5.1 | 0.00 | (ref.) | |
| Women | 513 | 5.0 | −0.01 | 0.57 | NS |

* This model included the variables: experimental condition, age categorized in 10-year classes, and sex.
†Mean percentage of answers missing.
‡SE, standard error; NS, not significant.

als with much partial nonresponse are sometimes considered so unreliable that they are not used. Among the 975 individuals classified as responders, 34 had more than 25 percent partial nonresponse. If we move them to the nonresponse category, the overall response rate diminishes to 47.1 percent. A reanalysis of table 1 after such reclassification produced odds ratios in the multivariate modeling of 1.26, 1.25, and 1.14, results that are very close to those obtained in the original analysis of 1.30, 1.24, and 1.14. The analysis in table 6 was also repeated after exclusion of the same 34 individuals, who could be considered outliers with respect to partial nonresponse. As expected, the standard errors of the estimates were reduced. The age effects were still significant, but they were reduced in magnitude. The differences between the design modifications were of about the same magnitude as before, but the maximum difference was now somewhat less than 2 percentage units.

## DISCUSSION

Our principal finding was that simple modifications of mailing strategies and of the questionnaire itself



Downloaded from https://academic.oup.com/aje/article/147/1/74/50426 by guest on 20 August 2022

may strongly influence the subjects' willingness to respond. In our study, the best combination of modifications yielded a response rate that was almost 40 percent higher than the rate obtained with the worst combination. Preliminary notification substantially increased the questionnaire retrieval rate, but the length of the questionnaire had less effect. Demographic factors such as age, sex, and population size were significant determinants of response propensity, but they did not confound our findings. The effect of questionnaire length was, however, modified by sex, and there was a tendency toward a stronger effect of preliminary notification among older subjects and among those who received the questionnaire on a Friday.

Our findings are probably internally valid because of the randomized design and externally valid because the study was population based. Since all selected subjects contributed information, whether they responded or not, selection bias due to drop-outs after randomization was not a problem in our study. Our conclusions presuppose that the questionnaire reached all of the subjects, but the efficient postal services and the continuously updated population registers in Sweden ensured that this was accomplished. Even if some questionnaires did not reach their addressees, such a failure should be nondifferential, which would cause a slight underestimation of the effects that would not greatly affect our conclusions.

Our overall questionnaire retrieval rate of 49 percent may seem low. Even in category 4, the retrieval rate was only 56 percent. We have no information about the individual reasons for not returning the questionnaire. Our questionnaire was of average size and complexity for an epidemiologic study. In the cover letter, the purpose of finding risk factors for kidney cancer was explained. Although the questionnaire was sent out during the summer, when a high proportion of the addressees may have been away from home, it is common to have the mail forwarded to the summer address. It was probably more important that only one reminder was sent out and that no other follow-up measures were taken. In a study by Marrett et al. (10), a response rate of about 50 percent with one reminder was increased to 70 percent with a second reminder. This is approximately the same response rate as in recent case-control studies in the Swedish population (14, 15). A second reminder would probably have increased the overall response rate, but it might also have changed the relative importance of each of the three design modifications. The importance of the preliminary notification would perhaps have been less if a second reminder had been used as well. Our finding of a positive effect of preliminary notification is, however, in good agreement with the results of several previous studies (1–4, 9). Although it may be preferable to use both preliminary notifications and two reminders, economic restraints often force the investigator to choose to use either a preliminary notification or a second reminder. We could not determine whether preliminary notifications are more cost effective than reminders after the main mailing, and this remains an open question.

We found a higher questionnaire retrieval rate with the shorter version of the questionnaire, although of borderline significance. Our findings agree with those of several other investigators, showing marginal effects of changes in the size of the questionnaire (1, 3–5, 8, 9). However, the results of previous studies are not consistent. There may be several reasons for the ambiguous results in the literature. One is that the effect may require a large relative difference in the number of pages. Neither our data nor the available literature gives an unequivocal answer as to whether a linear dose-response relation or a threshold effect exists between the length of the questionnaire and the response rate. Moreover, effects of the length are likely to be confounded by the content. The difference in the response rate may be due to the presence or absence of particular questions rather than to the length. (This can be seen clearly in our study, where women who generally had a higher response rate than men also received a longer questionnaire.)

Despite the weak and inconsistent empiric support, short questionnaires are often recommended in epidemiologic studies because of concern about the willingness of people to respond. Our findings belie some of these concerns. It may be more rewarding to pay attention to details in mailing procedures and to increase people's motivation than to cut out pages in the questionnaire.

Mention of a telephone contact reduced the questionnaire retrieval rate, but only marginally. It did not affect the partial nonresponse. This, to our knowledge, is the first study to investigate the effects of such information that is commonly found in mailed questionnaires. It is conceivable that some perceived the phone contact as a greater encroachment on their privacy and were therefore less willing to participate.

There was a clear positive relation between questionnaire retrieval rates and the mean proportions of missing answers. This suggests that the marginal addition of participants resulting from our efforts to increase the response rate was recruited from a less motivated stratum of the population. The moderate losses due to partial nonresponse did not outweigh the overall gains produced by the mailing modifications. Our findings, however, show a negative side of zeal-



Downloaded from https://academic.oup.com/aje/article/147/1/74/50426 by guest on 20 August 2022

ous attempts to increase the response rate at all costs. We were unable to analyze the validity of the answers in the various subgroups, but a high partial nonresponse rate may indicate a lower validity of the answers. The design of our study did not allow firm conclusions as to whether preliminary notifications are superior to repeated reminders in this respect. A preliminary notification can greatly influence people's attitudes toward the investigation before the main mailing, whereas reminders may be perceived as annoying to the extent that they impair the quality of the responses (3).

Age and sex had strong effects on the questionnaire retrieval rate. Subjects between 40 and 74 years were particularly willing to answer, as was also found in one interview study (7). Possible explanations include a higher prevalence of physical or mental diseases and of visual impairment in old people, as well as a more suspicious attitude (7). The perceived aims of the study may be of importance. Our study explicitly dealt with risk factors for kidney cancer, which may partially explain the low response rate among the youngest persons, who may be less concerned about such diseases that typically occur in old age. The effect of sex on response may be underestimated because women received a somewhat longer questionnaire. More research is needed to determine whether men and women evaluate health issues differently and whether they should be approached and motivated differently. The oldest age group (70–79 years) had not only a lower response rate but also a higher proportion of missing answers. This may be partly related to the type of questions that we asked. Many questions dealt with circumstances far back in time, requiring older people to tax their memories farther back than would be the case for younger people.

We conclude that mailing strategies may be as important for the response rate as the characteristics of the questionnaire itself. The combined effects of simple modifications in the mailing routines and in the appearance of the questionnaire may be of decisive importance for the success of an epidemiologic study. Our study also indicates that special measures may be required to attain high response rates among young and very old people, men, and city dwellers.

## ACKNOWLEDGMENTS

Supported by grants from the Swedish Cancer Society.

The authors thank Rick Kelso, Department of Cancer Epidemiology, for excellent technical assistance.

## REFERENCES

1. Yu J, Cooper H. A quantitative review of research design effects on response rates to questionnaires. J Marketing Res 1983;20:36–44.
2. Fox RJ, Crask MR, Kim J. Mail survey response rate. A meta-analysis of selected techniques for inducing response. Public Opinion Q 1988;52:467–91.
3. Harvey L. Factors affecting response rates to mailed questionnaires: a comprehensive literature review. J Market Res Soc 1987;29:341–53.
4. Linsky AS. Stimulating responses to mailed questionnaires: a review. Public Opinion Q 1975;39:82–101.
5. Yammarino FJ, Skinner SJ, Childers TL. Understanding mail survey response behaviour. A meta-analysis. Public Opinion Q 1991;55:613–39.
6. Dillman DA. Mail and telephone surveys: the total design method. New York: Wiley & Sons, 1978.
7. Kelsey JL, O'Brien LA, Grisso JA, et al. Issues in carrying out epidemiologic research in the elderly. Am J Epidemiol 1989; 130:857–66.
8. Kuskowska-Wolk A, Holte S, Ohlander E-M, et al. Effects of different designs and extension of a food frequency questionnaire on response rate, completeness of data and food frequency responses. Int J Epidemiol 1992;21:1144–50.
9. Spry VM, Hovell MF, Sallis JG, et al. Recruiting survey respondents to mailed surveys: controlled trials of incentives and prompts. Am J Epidemiol 1989;130:166–72.
10. Marrett LD, Kreiger N, Dodds L, et al. The effect on response rates of offering a small incentive with a mailed questionnaire. Ann Epidemiol 1992;2:745–53.
11. Perneger TV, Etter J-F, Rougemont A. Randomized trial of use of a monetary incentive and a reminder card to increase the response rate to a mailed health survey. Am J Epidemiol 1993;138:714–22.
12. Tambor ES, Chase GA, Faden RR, et al. Improving response rates through incentive and follow-up: the effect on a survey of physicians' knowledge of genetics. Am J Public Health 1993;83:1599–603.
13. Choi BCK, Pak AWP, Purdham JT. Effects of mailing strategies on response rate, response time, and cost in a questionnaire study among nurses. Epidemiology 1991;1:72–4.
14. Bergqvist L, Persson I, Adami H-O, et al. Risk factors for breast and endometrial cancer in a cohort of women treated with menopausal oestrogens. Int J Epidemiol 1988;17:732–7.
15. Galanti MR, Hansson L, Lund E, et al. Reproductive history and cigarette smoking as risk factors for thyroid cancer in women: a population-based case-control study. Cancer Epidemiol Biomarkers Prev 1996;5:425–31.

Downloaded from https://academic.oup.com/aje/article/147/1/74/50426 by guest on 20 August 2022