National health surveys by mail or home interview: effects on response

H S J Picavet

Abstract
Study objective—To study the effect of using a mail questionnaire or home interviews on the size and the selectivity of response to national health surveys.

Design—The interview survey and the mail survey were both carried out in the same country (the Netherlands) using the same sample frame, the same study period (1998) and collected partly the same data on demographic, socioeconomic and health characteristics.

Setting—The Netherlands.

Participants—Dutch non-institutionalised inhabitants aged 25 years and over.

Main results—Response to the mail survey was lower (46.9%, n=3664) than to the interview survey (58.5%, n=6061). The mail survey gave higher response rates for women and lower response rates for persons with lower levels of education. Respondents to the mail survey reported lower rates of smoking but a slightly worse health status and higher figures on the use of health care services. No differences by method of data collection were found for age, marital status, region, household composition, work status and categories of body mass index.

Conclusion—Although the response of the mail survey was lower than the home interview survey, respondents showed generally small differences, with exception of level of education.

(J Epidemiol Community Health 2001;55:408–413)

National health surveys are the main sources for data on many (public) health indicators. Data collection can be carried out by face to face interviews, telephone interviews, mail questionnaires or a combination of methods. In the Netherlands we often use face to face interviews conducted at home or mail questionnaires. The choice of the mode of data collection is determined by several factors, including available resources and response expectations. Face to face interviews surveys are much more expensive than mail surveys. Several effects of mode of data collection on response are known.

Firstly, it is generally thought that response rates are better for interview surveys than mail surveys, some evidence for elderly people is available. Secondly, those who respond to interview surveys can be different from those who respond to mail surveys. There is some evidence that lower socioeconomic classes are under-represented in mail surveys compared with interview surveys. Thirdly, people can respond differently to questions on paper than to questions asked by an interviewer. For instance, for some disability indicators it is known that systematically higher prevalences are found using self administered questionnaires compared with interviews. The same is found for other health indicators. In addition, questions that can be affected by social desirability, for example, alcohol consumption, using of car belts, are suspected to do better in self administered questionnaires than in face to face interviews. However, information on height, household composition, work status is considered not to be affected by mode of data collection.

For this paper we were able to study the response to a health mail survey and to a health interview survey, both using the same sample frame (population register), the same target population (the Dutch non-institutionalised population) and mainly the same topics and questions. The question of our study is: does using a mailed survey or an interview survey lead to different response groups? According to the (limited) data in the literature we expect (1) the response on the mail survey to be lower than on the interview survey, (2) that respondents to the mail survey with lower educational levels are under-represented and (3) that population estimates on non-mode dependent questions such as work status, number of persons in the household, height and weight, are not affected by mode of data collection. (4) Our last hypothesis is that population-based estimates of health indicators based on mail survey will represent a less favourable health status compared with the interview survey.

Methods
Two health surveys carried out in the Netherlands were analysed (1) the Netherlands Health Interview Survey (NetHIS) of 1998, which uses face to face interviews at home carried out by trained interviewers combined with a paper questionnaire, and (2) the baseline of the Dutch Musculoskeletal Complaints and Consequences Cohort study (DMC3-study), which uses mail questionnaires. General characteristics of the two studies are presented in table 1. The NetHIS is a continuous survey started in 1981 and carried out by Statistics Netherlands. From 1997 the NetHIS is one module of the integrated system of face to face interviews of Statistics Netherlands. The DMC3-study is carried out by the National Institute of Public Health and the Environment in collaboration with Statistics Netherlands. For the two surveys the same sample frame (population register) and sample method were...
used, although for the DMC3-study this was a stratified sample. The sample frame provides us with data on date of birth, sex, marital status and address details of the persons in the sample. The surveys collected identical information for sociodemographic characteristics, health indicators, risk indicators and the use of health services. We have, however, two exceptions: the questions on household composition and on educational level were much more detailed in the NetHIS than in the DMC3-study.

The net response was calculated by dividing the number of respondents by the number of those actually approached excluding those who were known to be deceased or those whose address was unknown. The response figures are presented for both surveys: for the total group and by sample frame characteristics—that is, age group, sex, marital status and region. For NetHIS two response figures are given: one for the interview and one for questionnaire. After the interview, which is carried out by trained interviewers with laptop computers the interviewees were handed the paper questionnaire and asked to fill it in and send it back by free post return envelope.

To compare the results of the surveys, both surveys were weighted. Weighting factors were constructed in such a way that the distribution of both surveys by age, sex, region and marital status was equal to that of the Dutch population of 1998. The surveys were then compared for (1) sociodemographic characteristics (household composition, education, work status), (2) health indicators (perceived health, limitations in daily life, chronic conditions), (3) risk factors (smoking, body mass index) and (4) the use of health care services.

Household composition presents the number of persons living in the household. Level of education was measured as the highest level reached and then summarised in four groups: primary school, junior (vocational) education, secondary (vocational) education, vocational colleges/university. Work status is defined according to four categories: have paid work for more than 12 hours a week, does not have paid work for more than 12 hours per week but wants to have work for more than 12 hours a week, does not have nor wants work for more than 12 hours a week, and those who are work disabled or have a pension (at least everybody above 65 years of age). Perceived health was measured with the question, “How do you rate your health using a mark for a school report?” In the Netherlands these marks are between 1 and 10 with 10 as the best mark. The following groups were made according to other descriptions of health: mark 1 to 5 (bad health), 6 to 7 (average health), 7.5 to 8.5 (good health) and higher than 8.5 (excellent health). Limitations in daily life were measured by a simple question “Do you have any limitation in your daily activities due to a health problem?” For the assessment of chronic conditions a list of chronic health problems was used. The majority of the descriptions of chronic conditions were identical for the DMC3-study and the NetHIS. These were: COPD (chronic obstructive pulmonary disease), sinusitis, coronary heart disease or other severe heart disease, hypertension, (consequences of) stroke, peptic ulcer, severe intestinal disorder, diabetes, thyroid disorders, epilepsy, dizziness and falling, migraine, severe skin disease and cancer.

For smoking five fixed response categories were used: every day smokers, occasional smokers, former every day smokers, former occasional smokers and never smokers. “Occasional” smokers are those who smoke less than one cigarette a day. Body mass index (BMI) is calculated by dividing weight (kg) by squared length (m) and was categorised as < 18.5 (underweight), 18.5–25 (normal weight), 25–30 (moderate overweight) and > 30 (severe obesity).
The use of health care services was measured identically in both surveys for contact with general practitioner (GP), medical specialist and physiotherapist. For the GP and the medical specialist two indicators were calculated: contact in the past two months and contact in the past year. For the physiotherapist only contact in the past year was measured. Differences in results between the surveys were calculated using the weighted percentages and the 99% confidence limits were calculated using the standard errors of the unweighted prevalences.

Results
The response to the interview of the NetHIS was higher (58.4%) than the response of the mail questionnaire of the DMC3-study (46.9%) (table 2). Using a questionnaire as a second step after the interview survey gave some additional non-response, resulting in a net response to the NetHIS questionnaire of 47.9%. The mail survey gave higher response rates for women than for men whereas the NetHIS shows no differences by sex. The response patterns of both surveys for marital status and region were similar. The response of persons of 65 years and over was slightly lower than of the other age groups. Those who were married showed the highest response and those who were not married the lowest. By region of living we found a slightly lower response in the west, which is the most urbanised region of the Netherlands, including Amsterdam, Rotterdam and The Hague.

Table 3 presents the estimations of different characteristics of the Dutch population aged 25 years and over according to the two surveys. These estimations were almost identical for household composition and work status, but for level of education we found that those with only primary school were underrepresented in the DMC3-study with 14.5% compared with 21.1% in the NetHIS. If this figure is estimated...
on the basis of those who participated in the NetHIS and also returned the questionnaire, we get a percentage of those with only primary school of 19.5%. Of the NetHIS respondents who did not return the questionnaire 28.4% (p<0.01) had only primary school. For the other characteristics there was no difference between the respondents of NetHIS who did and who did not return the questionnaire (not shown).

For the health indicators we see that estimated proportions of the extreme values of subjective/perceived health (bad and excellent) were the same for both surveys but that there was a small shift to better health in the NetHIS.

In contrast with that the NetHIS reported a higher prevalence of persons with limitations in daily life (16.6% versus 12.7%), suggesting a worse health for the respondents of NetHIS. The estimated prevalences of 14 chronic conditions were the same in both surveys or slightly higher in the DMC3-study. The latter was true for: sinusitis, severe intestinal disorder, dizziness with falling, and migraine.

The DMC3-study gave a lower figure for every day smokers than the NetHIS, 24.2% versus 29.4% but the estimations of BMI categories were the same for both surveys. The use of health care services (GP or specialist) during the past two months and contact with the physiotherapist during the past year was more often reported by the respondents of the DMC3-study. The percentage of persons with contact with the GP during the past year was the same for both studies and the percentage with contact with medical specialist during the past year was relative lower in the DMC3-study.

Discussion
In this study we found that a health mail survey had a lower response rate than a health interview survey but that the differences in the respondents by mode of data collection were small with exception of educational level.

Other studies also reported higher response rates for interview surveys than mail surveys. However, it is very difficult to make absolute statements about this because many other factors could have affected response such as study design differences. In our study one important drawback of the mail survey was the length of the questionnaire. The questionnaire was rather long, with 28 pages. However, the questionnaire was divided in a few parts indicated by coloured paper and if a screening question was negative many pages could be skipped.

There is some evidence that long questionnaires affects response negatively, although one study did not found a difference in response by using a 4 page or 16 page questionnaire. We think that the response of the mail survey would have been higher if we had used a much shorter questionnaire but then the amount of information gathered would have been much less.

In both surveys analysed in this paper, high risk groups for low response were: higher age groups (65 years and over), and those who were not married (anymore), and this was found before. Usually men were also underrepresented but we did not find this in the NetHIS. We found no differences between the mail survey and the interview survey with regard to age, marital status and region of living. Other studies also reported no difference by region although rural areas had a slightly higher response than urbanised areas. For the demographic characteristics, we conclude that our national mail and interview surveys with respect to health related topics have similar, slightly selective, response.

An under-representation of lower socioeconomic groups in the mail survey has been found before, regardless the indicator being income levels or level of education. However, also examples exist with no differences in response by income groups. Because level of education is in general such an important determinant of health and health related behaviour it was important to look for response bias attributable to educational level.

Unfortunately there is no other source than population surveys for the information on the distribution by level of education in the Netherlands. So it is impossible to say whether or not the estimations of the surveys represent an underestimation or overestimation. This study indicates that those with only primary school were less likely to respond to a mail questionnaire than to an interview survey.

For a health survey it was also important to identify response bias attributable to health. What are the health characteristics of respondents and non-respondents? In general it is suggested that respondents to health surveys are the “worried well” healthy people who see their doctor regularly and follow healthy lifestyle practices. Our study showed that the mail survey gave higher rates of health care utilisation and a lower prevalence of smoking compared with the health interview survey.

The overestimation of the utilisation of health services by mail surveys was found before and the same was true for the underestimation of smokers. The mail survey gave a slightly more unhealthy picture of the population than the interview survey, based on subjective health and a few chronic conditions. However, in contrast with what we expected the prevalence of limitations in daily life was higher in the interview survey than in the mail survey. Examples of studies on non-response bias with respect to health characteristics are mail surveys among the very elderly and non-response bias was often found attributable to bad health and

KEY POINTS
- For many health topics carefully designed mailed surveys are probably an equally good alternative for the—much more expensive—interview surveys.
- People with only primary school should receive extra attention in the health survey design.
- Investigating non-response bias should be part of every survey regardless of the response rate.
diminished cognitive functioning,\textsuperscript{28} affecting health related estimations more than factual information.\textsuperscript{1} Elderly non-responders also had higher death rates.\textsuperscript{2,14} In general we can expect that health surveys miss the very unhealthy part of the population and that respondents of mail surveys can be described as the “worried well”.

It was important to note however that although some of the differences between the NetHIS and the DMC\textsubscript{3}-study are statistically significant the differences in terms of percentage points are small. A lack of differential non-response in comparing mail surveys and interview surveys was also found in other studies.\textsuperscript{21,29}

A limitation of this study is that both surveys have a rather low response. Surveys in the Netherlands are in general confronted with lower response rates than health surveys elsewhere in Europe,\textsuperscript{30} or compared with mail surveys published in medical journals\textsuperscript{31} despite intensive efforts to reduce non-response. General guidelines to increase motivation for participation\textsuperscript{41} were used where possible, including an interesting topic (health), confidentiality, a good reputation of the two organisations responsible for the surveys, approval by official institutions, use of advance letters in the home interview survey, repeated contacts (with a maximum of three), use of trained interviewers and attractive questionnaires. Systematic differences in health and health related topics between those participating and not participating in health surveys—whether or not by mail or interview—are possible. If such extreme groups exist we are never able to investigate them but they do not differ in terms of sociodemographic characteristics and they do hardly differ between a mail or interview survey.

An additional limitation is the difference in time period of data collection between the NetHIS (all year in 1998) and DMC\textsubscript{3}-study (September to December 1998). In 1998 the response to the HIS per month varied from 55.7\% (June) to 64.1\% (October). The response to the NetHIS is somewhat lower in the summer months because of holidays. There is however no systematic variation in health prevalences by month or season.\textsuperscript{10}

Differences in sponsorship or themes of the surveys can also contribute to the response. Statistics Netherlands was involved in both studies and in the DMC\textsubscript{3}-study also the National Institute of Public Health. Both organisations are national government associated non-profit organisations. We do not think that differences in the perception exist or should have affected response rates. Because the focus of the DMC\textsubscript{3}-study was on musculoskeletal health problems, and the NetHIS was a general health survey, the response could have been higher for those with musculoskeletal health problems than for those without musculoskeletal health problems. Because musculoskeletal health problems are very common this would not have affected the response rate but an overrepresentation of musculoskeletal health problems can be expected. For limitations in daily life, however, we found a higher prevalence in the NetHIS than in the DMC\textsubscript{3}-study.

Our comparisons showed that the response was lower on the mail survey than the interview survey, that responders with lower educational levels are underrepresented in the mail survey, and that estimates of questions such as work status, number of persons in the household, height and weight, are not affected by mode of data collection. Our fourth hypothesis, stating that population-based estimates of health indicators based on mail survey will represent a less favourable health status compared with the interview survey, is not confirmed for all health indicators. In general, we can conclude that the differences in respondents between interview surveys or mail surveys are no reason for great concern.

The author wishes to thank H W V van Gils from University of Maastricht and Statistics Netherlands for the analyses of NetHIS and Dr J S A G Sonneboom for help in the design of the study and colleagues for useful suggestions for the text.

Funding: the Dutch population-based Musculoskeletal Complaints & Consequences Cohort (DMC\textsubscript{3}-study) was financially supported by the Ministry of Health, Welfare and Sport of The Netherlands and the National Institute of Public Health and the Environment and carried out in collaboration with Statistics Netherlands.

Conflicts of interest: none.