ONLINE VERSUS POSTAL QUESTIONNAIRES: A COMPARISON BETWEEN TWO DATA COLLECTION METHODS

Views of the choice of data collection methods on the basis of the Physician 2011 study

Alar Sepp
Lecturer of Chair of Pharmacy; Tallinn Health Care College, Estonia

Abstract

Objectives: The purpose of this article is to find out differences between surveys using paper and online questionnaires. The author has deep knowledge in the case of questions concerning opinions in the development of survey based research, e.g. the limits of postal and online questionnaires.

Methods: In the physician studies carried out in 1995 (doctors graduated in 1982-1991), 2000 (doctors graduated in 1982-1996), 2005 (doctors graduated in 1982-2001), 2011 (doctors graduated in 1977-2006) and 457 family doctors in 2000, were used paper and online questionnaires. The response rates were 64%, 68%, 64%, 49% and 73%, respectively.

Results: The results of the physician studies showed that there were differences between methods. These differences were connected with using paper-based questionnaire and online questionnaire and response rate. The online-based survey gave a lower response rate than the postal survey. The major advantages of online survey were short response time; very low financial resource needs and data were directly loaded in the data analysis software, thus saved time and resources associated with the data entry process.

Conclusions: The current article helps researchers with planning the study design and choosing of the right data collection method.

Key terms: online; questionnaire; survey; data collection; research methodology.

1. Introduction

Carrying out a survey involves a wide range of issues which must be resolved before the study. Who are the study group? How to design the sample? What sort of questions should be asked and how to ask them from a target population? (De Leeuw et al. 2008a; Ruskoaho et al. 2010). The aim of using the questionnaire survey is to collect as reliable and representative material about the study group as possible (Tekkel 1997; Ruskoaho et al. 2010). In recent years, the response rate to the questionnaire surveys has fallen significantly which has affected the results, conclusions and trust towards the collected material (De Leeuw 2008; Ruskoaho et al. 2010). It is also connected with a major change in the way of conducting questionnaire surveys in recent years, when the traditional postal surveys have been replaced ever more often by electronic data collection methods. These methods involve a number of benefits which facilitate and speed up the process of studies. However, collecting data electronically implies considerable methodological challenges and problems, which differ from those connected with
traditional paper questionnaires and which are to be taken into account when preparing a study and carrying out data collection. There is a multitude of daily Gallup polls done electronically which do not meet the criteria of a scientific study. In the case of these polls there is no defined study group from which a sample is drawn based on probability. Instead of that, they depend on the personal participation activity of respondents (Couper 2000; Ruskoaho et al. 2010). In the present report we discuss on the one hand the opportunities, challenges and problems of collecting data electronically, and on the other hand, the possibilities of combining paper and electronic questionnaire surveys based on the experience of the Physician 2011 study. The data collection for the Physician 2011 survey is assessed in relation to the sources of error, which is generally found in the regular questionnaire surveys and which comprises the problem of representativeness, sampling errors, nonresponse and measurement (Groves et al. 2004; Ruskoaho et al. 2010). The aim of the report is to help investigators and research teams in the planning of their surveys and in the selection of the correct data gathering method.

The PHYSICIAN 2011 study was conducted by the Estonian Medical Association, Tallinn Health Care College and Finnish research group LÄÄKÄRI 2008 (Heikkilä et al. 2009, Heikkilä et al. 2010; Ruskoaho et al. 2010). For that purpose an extensive survey was carried out in the summer (June, July, August) of 2011. The study population consisted of physicians who had graduated in years 1977 to 2006 (N = 4702) and the study examined their psychosocial background, opinions on working conditions and medical education (undergraduate and postgraduate training), as well as values and professional and communal identity of doctors. In carrying out the survey extra challenges were presented by following: the study population was divided into two cohorts by the graduation year and their questionnaires differed partially from each other. The younger cohort consisted of the physicians who had graduated in 1997 to 2006 (n = 891) and who were called "the junior cohort". The older cohort of doctors ("the seniors") had graduated in the years 1977 to 1996 (n = 3811). The sample for the study of doctors comprised ultimately half of the whole study population (445+1905=2350). The physician surveys were started 17 years ago with the Young Physician 95 study and the Physician 2011 survey carried on the previous research (Saava, Sepp 2003, Sepp, Virjo, Mattila 2007, and Sepp 2012a). Since then, the studies have been repeated in every five years (in 2011, after six years) (Sepp 2012b). In the publications, separate attention has been given to the results of the junior cohorts (Sepp 2009b). In the first study the response rate of young doctors was 64 per cent, which became higher in 2000. Later the response rate started to fall, but in 2005 it was 64 % (Sepp 2009a). In order to achieve a higher response rates, the online questionnaire was used introduced for data collection in the last (2011) study. The program SurveyMonkey® was used for electronic data collection (SurveyMonkey® 2012). Differently from the postal questionnaire method, the online survey produced stratified seniors' and juniors' data and in the end, these data were merged. When the survey was designed, the methodological pretesting was not planned. The objective was to collect representative data cost-effectively. In total, the number of respondents to the electronic survey was 218 (response rate 66%), of which around 48% answered in the first survey round. The response rate stayed approximately the same in spite of the new method of data collection.
collection, as shown by Table 1.

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Junior</th>
<th>Senior</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>The survey population</td>
<td>891</td>
<td>3811</td>
<td>4702</td>
<td></td>
</tr>
<tr>
<td>One half of population</td>
<td>445</td>
<td>1905</td>
<td>2350</td>
<td></td>
</tr>
<tr>
<td>Incomplete addresses</td>
<td>115</td>
<td>495</td>
<td>610</td>
<td></td>
</tr>
<tr>
<td>Sample</td>
<td>330</td>
<td>1410</td>
<td>1740</td>
<td></td>
</tr>
<tr>
<td>Responding to electronic questionnaires in the 1st round</td>
<td>104</td>
<td>257</td>
<td>361</td>
<td></td>
</tr>
<tr>
<td>Responding to electronic questionnaires in the 2nd round</td>
<td>77</td>
<td>201</td>
<td>278</td>
<td></td>
</tr>
<tr>
<td>Responding to electronic questionnaires in the 3rd round</td>
<td>37</td>
<td>105</td>
<td>142</td>
<td></td>
</tr>
<tr>
<td>All respondents</td>
<td>218</td>
<td>563</td>
<td>781</td>
<td></td>
</tr>
<tr>
<td>Total complete (%)</td>
<td>137 (42%)</td>
<td>390 (28%)</td>
<td>527 (30%)</td>
<td></td>
</tr>
<tr>
<td>Response rate</td>
<td>66%</td>
<td>40%</td>
<td>49%</td>
<td></td>
</tr>
</tbody>
</table>

2. The planning of questionnaires and data collection

Different data collection methods have different advantages and disadvantages. Different methods may bring about new errors. For starting the collection of data, the most efficient and cost-effective way is to choose the cheapest method (De Leeuw et al. 2008b; Ruskoaho et al. 2010). Therefore, in the Physician 2011 study electronic questionnaire was applied for data collection (SurveyMonkey®, Estonian Medical Association 2011), which was considerably cheaper than the surveys carried out in the years 1995, 2000 and 2005. In 2011 the total cost was 1,432 euros (1995 - 10,794 euros, 2000 - € 11,565; 2005 - € 5,782) (see Table 2). Invitations to participate in the study were sent to those doctors whose email addresses were found in the register of the Health Board (agency of Ministry of Social Affairs) and the Estonian Medical Association. Compared to the use of the postal questionnaire, the electronic data collection brings about an obvious reduction in expenditure through saving on account of postal charges, consumption of materials or the costs of data entry. So, the online-based survey was very much cheaper than the postal surveys (see Table 2). At the same time, these costs do not depend on the sample size, length of the questionnaire or the number of survey rounds (Manfreda, Vehovar 2008; Ruskoaho et al. 2010). Those who previously would not respond to postal surveys may readily accept the online survey, if the members of the study group find the new mode of being surveyed more pleasant. At the same time the identification of differences in the collected responses takes some extra work and adds costs. Designing the survey questionnaire was an intricate process in the Physician 2011 study. Firstly, we wanted to maintain comparability of the survey with the postal surveys of earlier series. Secondly, the electronic questionnaire had to be similar to the previous paper-based questionnaires. Thirdly, the differences in junior and senior doctors' questionnaires had to be taken into account. The use of the electronic data collection method produced separate data for the
cohorts of senior and junior doctors which were then merged into one data. Quite certainly, in electronic data collection, one might use dynamic interval responses for hiding the questions on undergraduate education from the view of the senior cohorts and likewise for hiding management-related questions from the juniors, so that only the seniors would answer to them. It was not the practice in year 2011. By using the electronic questionnaire, we can lead the respondent past these questions which do not concern him/her. It is unnecessary to ask questions about work or residency from doctors, who are not working or are not in postgraduate training (residency). The respondent saves time if he or she does not need to scroll down over irrelevant questions. At the same time, the advisory questions in electronic questionnaires have to be answered. The process of providing answers may stop if the respondent is unwilling to answer a particular question or fails to understand that moving on to the next page of the questionnaire is impossible if one skips a question. The use of these characteristics allows interrupting responding when necessary and resuming it sometime later.

Table 2. Characteristics of the physician studies by method, respondents and costs.

<table>
<thead>
<tr>
<th>Years of the physician studies</th>
<th>1995</th>
<th>2000</th>
<th>2005</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method of study</td>
<td>Postal</td>
<td>Postal</td>
<td>Postal</td>
<td>Online</td>
</tr>
<tr>
<td>Costs of the study</td>
<td>10,794 €</td>
<td>11,565 €</td>
<td>5,782 €</td>
<td>1,432 €</td>
</tr>
<tr>
<td>Number of physicians who responded</td>
<td>483</td>
<td>482</td>
<td>635</td>
<td>781</td>
</tr>
<tr>
<td>Costs per one respondent</td>
<td>22.5 €</td>
<td>24.00 €</td>
<td>9.11 €</td>
<td>1.84 €</td>
</tr>
<tr>
<td>Expenses of the earlier studies compared to the Physician 2011 study</td>
<td>−12 X</td>
<td>−13 X</td>
<td>−5 X</td>
<td></td>
</tr>
</tbody>
</table>

In the earlier studies we used paper questionnaires comprising 20 pages (in 1995), 27 pages (in 2000) and 24 pages (in 2005) (Estonian Science Foundation, University of Tartu 1995; University of Tartu, Estonian Science Foundation 2000; Tallinn University of Technology, Ministry of Social Affairs 2005). Electronic questionnaires consisted of 29 pages (SurveyMonkey® Estonian Medical Association 2011), with time expenditure (~30 minutes by the estimate of the test survey manager) which can pose a problem, as already 15 minutes spent for responding, excl. the special groups, is considered to be lengthy in the case of electronic questionnaires (Czaja, Blair 2005; Ruskoaho et al. 2010). Despite spreading the questions onto numerous pages, the process of answering still occasionally required scrolling down the page, which is inconvenient in online questionnaires (De Leeuw, Hox 2008; Ruskoaho et al. 2010). The text visibility of different computer monitors varies and due to long sets of questions in the questionnaire, scrolling the screen pages could not be completely avoided. Test-questionnaires were drawn up for establishing proper balance between the page dimensions and the scrolling of pages. Electronic data collection provides technical capabilities which can improve the quality of the collected material. The use of compulsory questions reduces partial non-response and the intentional bypassing of certain questions shortens the particular questionnaire and saves time for the respondent. For paper questionnaires, advisory section can also be added, but detailed guidelines lengthen the questionnaires.
Additionally, one can use drop-down menus which enable to compress long variable questions into shorter form. Examples of other options available in electronic questionnaires include graphical tools, new question types and instructions for respondents placed in pop-up-windows. Using these options allows introducing special features of computer-assisted surveys cost-effectively into questionnaires (Dillman 2008; Ruskoaho et al. 2010). The *Physician 2011* study did not use all these possibilities of electronic inquiry, but rather tried to maintain the similarity to the earlier postal questionnaires. At the choice of various data collection methods, one must be aware that using different methods, it is worthwhile to retain the completeness and make compromises if needed.

3. Checking the Bias in *Physician 2011* study

*The problem of coverage*

Coverage refers to the extent to which the studied population is represented among those who serve as the basis for forming the sample. The sample cannot include persons who do not belong to the target group of the study, nor can one person respond several times (De Leeuw et al. 2008b; Ruskoaho et al. 2010). For electronically performed surveys there is often a significant coverage problem, because email addresses are changed more often than the postal ones and the incoming mail is checked with less regularity than traditional letter boxes. Also, spam-filter programs used by organizations for inspecting incoming emails may filter off messages about the survey (Lohr 2008; Ruskoaho et al. 2010). In order to improve coverage, the same survey will combine the use of the postal and electronic questionnaire (De Leeuw et al. 2008b; Ruskoaho et al. 2010). The studies show that the coverage of email addresses varies and depends e.g. on the level of education, income and age of the respondent. The selection of the sample becomes more difficult if certain respondent groups are over- or underrepresented in the sample (Couper 2000; Ruskoaho et al. 2010). Highly educated medical doctors can be seen as a suitable group for conducting questionnaire surveys. The overall introduction of e-Health (electronic health record system) and e-Patient Case Records has made the nature of doctor’s work more computer-centred and using the computer has become a professional requirement (Winblad et al. 2009; Ruskoaho et al. 2010). In the case of the doctors sampled for the *Physician 2011* survey, about 13 per cent had not reported any email address for mailing the survey. Email addresses were known in the case of 87 per cent of the sample. Part of the email addresses appeared to be false and it was possible to send survey messages to about 80% of the sample. However, there is no certainty in the case of non-responses if the existing email addresses were actively used or not. The Estonian Medical Association verified data on doctors included in their register, incl. their email addresses. The vast majority of email addresses were collected by the Internet search on hospitals, medical specialists and general practitioners. Typically, doctors have an email address at place of work. Younger doctors are prone to change their working organization and their precise email addresses may become a problem. In some cases more than one doctor was placed under one email addresses. In these cases we called and agreed about the person who would respond to
the questionnaire. The registered email addresses did not appear to be always personal. Problems emerged with email addresses of the target group doctors. For example, doctors who had left to work abroad may have had outdated email addresses and they could not respond to questionnaires. The addresses of the members of the Estonian Medical Association are updated, as they receive the "Estonian Physician" magazine, but the email addresses of non-organized Estonian doctors tend to be incorrect more frequently. After receiving the electronic questionnaire, some persons asked to forward it to another address, and hence, they were sent survey invitations with links to the survey. In the Physician 2011 study the survey manager did not consider multiple responses to constitute a reasonable risk, as filling in the online questionnaire could be done from one email address only once. The possibility of multiple responses would be conceivable if the respondent had two email addresses. That possibility was eliminated with the exact selection of doctors for the sample and with careful recording of email addresses of the selected sample members. For each person only one email address was found and a message about the Physician 2011 study was dispatched to a single address. In order to improve the survey's coverage one might consider sending survey messages to all known addresses, if one can be certain that the sampled persons would provide answers only once.

Non-response

The response rate is vital from the viewpoint of analysing and generalising the results of the questionnaire survey. The randomly distributed response activity is not very detrimental, because it reduces only the accuracy of calculations and evaluations (De Leeuw et al. 2008a; Ruskoaho et al. 2010). Out-drop is analysed by statistical methods, such as data weighting (weighted average). The failure to weight the data may bring along incorrect conclusions if the drop-out is connected with the target group of the study or with important background variables (Biemer, Christ 2008; Ruskoaho et al. 2010). In addition to the non-response, there is usually also partial non-response which provides a general reason to use statistical methods intended for correcting the incomplete data (Rässler 2008; Ruskoaho et al. 2010). According to the meta-analysis of 35 questionnaire surveys conducted in 1992-2005, the response rate with the electronic data collection method was below that of the paper surveys (Shih, Fan 2009; Ruskoaho et al. 2010). The lower response activity can be explained by a number of IT-related problems, such as incompatibility of different monitors, questionnaires and programs, slow or interrupted internet connections and the technical problems of computers. In comparison with the paper-based questionnaires, answering to the electronic questionnaires requires scrolling down the screen to gain full view, which may seem difficult to some respondents (Dillman, Bowker 2001; Ruskoaho et al. 2010). One way to increase the participation in the survey is to motivate the respondents by clarifying the importance and objectives of the survey. Additionally, one can offer incentive rewards by drawing lottery to the participants of the study (Lynn 2008; Ruskoaho et al. 2010). In the case of the Physician 2000 and Physician 2005 studies the doctors were notified in advance by articles published in the journal "Eesti Arst" (Estonian Physician) and the magazine "Meditsiiniuudised" (Medical News).
accompanying letter explained the importance of the survey from the viewpoint of improving the working conditions of doctors. The survey was sent on behalf of the Estonian Medical Association (EMA), accompanied by a letter from the president of the EMA, which stressed the high priority of the survey and distinguished it from other surveys sent to doctors for business purposes. The fact that the Physician 2011 study series have a history of being conducted since 1995 may also serve as a favourable reference for respondents. Differently from an electronic questionnaire, filling in the paper form can be conveniently stopped and resumed later. However, also the electronic inquiry could be continued after interruption from where it had stopped which was not known to respondents and unfortunately not emphasized by organizers of the study. There is a pressure to make online questionnaires short, as simultaneously with responding to them the persons often use their computers for alternative activities (with several programs in operation at the same time), which contributes to interruptions (De Leeuw 2008; Ruskoaho et al. 2010). The Physician 2011 survey was responded to on approximately 800 occasions, while nearly 250 respondents had discontinued answering the survey. A number of respondents who provided feedback considered the questionnaire excessively long. The progress indicator at the questionnaire measured the progress of answering by the scale of percentage of all questions (every respondent would see only of the two questionnaires – either for the junior or the senior cohort). It was possible to resume the once-stopped process of answering from the point of interruption (the participants of the study were not notified of this option) and the unaware respondents may have stopped answering. The latter is confirmed by the fact that every third doctor (254 persons (32.5%) from 781) who started responding did not complete it. As the Physician 2011 study did not entail testing the effectiveness of the data collection method, one cannot compare the impact of the online inquiry to that of the paper survey as regards the drop-out analysis. There were no differences in the response activity by men and women in the case of the electronic survey, whereas women generally had responded to postal questionnaires more actively than men throughout the survey series and the current lack of difference surprised the researcher. The response activity of different age groups of women did not vary significantly, while the younger age groups of men responded more frequently than the men from older age groups (it was vice versa in the case of the similar Finnish survey Lääkäri 2008) (Heikkilä et al. 2010; Ruskoaho et al. 2010). One factor which possibly influenced the drop-out in the survey is the respondents' mother tongue, as the questionnaires were exclusively in Estonian this time. In previous surveys the postal questionnaires had been in Russian and Estonian. Responding to the questionnaire requires proficiency in language and in the linguistic sense the ability to recognise nuances. There was no question about the respondent's native language in the survey and, hence, the response rates in this regard cannot be determined. Because of the internationalization of the medical community, the surveys should be set up with questionnaires in different languages.

**Sampling error**

Different samples selected from the same population produce more often than not a varying...
result which is known as the sampling error (Kuusela 2009; Ruskoaho et al. 2010). The sampling theory assumes that every member of the target group has the chance to be sampled, after which the statistical analysis would help to evaluate the sampling error. A typical way of selecting the sample is by simple random sampling, where each member of the population has equal probability of inclusion in the sample. The use of the stratified sample gives certainty that also the samples formed from smaller groups will have a sufficient possibility of providing data. If the sample is not based on probability, there must be strong reasons why the selected target group can represent the entire population. For the Physician 2011 study, the sample was chosen randomly by including in it every second doctor who had graduated in years 1977-2006. In the previous studies the sample selection was based on the birth date of doctors (persons born on odd and even days). In 1995 the sample included those born on odd day, in 2000 the persons born on an even day, in 2005 – again those with an odd birthday, etc. This principle of forming the sample was used for reducing the response strain of doctors, involving them in the comprehensive study not in every five years, but only in every ten years. Also, the sample is large, including almost half of the population, which does not put the sample’s representativeness in question.

**Measurement error**

In electronic questionnaires the measurement of phenomena may be more accurate due to the possibility of presenting various specifying instructions by way of text, image or sound to the respondent (Dillman 2008; Ruskoaho et al. 2010). When combining the methods of data collection, the same impulses must be given to the receiver to ensure that the way of data collection will not affect answering (De Leeuw et al 2008b; Ruskoaho et al. 2010). The empirical and systematic study of the comparative influence of various data collection methods on survey results regarding the use of online and paper-based questionnaires is still embryonic. There have been some observations of variability of the stratified out-drop due to different methods of data collection and sometimes the qualitative differences between responses can be explained by different background variables, i.e. the online questionnaires have had other respondents than the postal questionnaires. For finding out the differences between inquiry methods, there is a need for further studies (De Leeuw 2005; Ruskoaho et al. 2010). In designing the questionnaire of the Physician 2011 survey, an attempt was made to achieve the greatest possible degree of similarity to the previous questionnaires (questions). The online questionnaire differs from the earlier postal questionnaires by the number of questions on a page. We realized that most computer programs were unable to scroll lengthy pages containing numerous questions. We were forced to change the questionnaire and accommodate only one to seven questions per page. Questions which were longer and had multiple response options were placed on separate pages and shorter questions with few response options were also concentrated on one page. The advisory texts in the web questionnaire were identical with those used in postal questionnaires and did not include graphical elements or annexed advisory texts. In the earlier studies of the survey series we used the VAS (Visual Analogue Scale) type of questions. Taking into use the SurveyMonkey program, that type of questions were not available among
possible multiple choice questions and therefore, these were replaced by Likert scale questions. So, the present survey is not entirely comparable with the previous studies of our series. Not all question types of the previous paper-based questionnaires were retained. In the online questionnaires we used drop-down lists instead of open-ended questions asked in the postal questionnaires. In the case of paper-based questionnaires it had caused extra problems during the entry of data, but due to the unambiguous nature of the questions it hardly affected the quality of data. Although we performed several tests, all problematic parts were not detected and, for example, after the first round of the inquiry, we made the questionnaire pages shorter and placed fewer questions on them. This change improved the response activity considerably and increased the number of fully completed questionnaires. The electronic questionnaire became more extensive than the previous postal questionnaires as it contained more than 400 separate questions. The online questionnaire did not ask about the geographical location of the doctor's workplace anymore and the doctors practising in Estonia (the country as a whole, not divided by counties and major cities) were studied.

4. Discussion: the selection of the proper method of data collection

The possibilities of electronic data collection have improved carrying out the questionnaire surveys. Internet-based data collection instruments are intended for public use, which has caused the explosive growth of the number of different surveys in the 21st century (Bethlehem, Biffignandi 2012). The massive flood of online questionnaires and the consequent response fatigue provide a possible explanation for the fact that the response rates of various surveys have fallen significantly in the recent time. This is clearly reflected also by the Estonian physician’s studies. There is a risk that the quality of the collected survey data may fall, because the representativeness of the data collected by online surveys remains below that of traditional postal questionnaires. That constitutes a major problem from the aspect of the generalizability of results. The coverage of data is impaired also when survey messages are ever more frequently filtered out as spam by work organizations, meaning that the questionnaires do not reach the target population with the same level of coverage as before. The success of electronic surveys requires beforehand the availability of correct email addresses and certainty that the email messages will reach the respondents. However, data can be collected online even if the addresses are unknown. For that purpose the respondents must be sent postal letters telling them the Web site for responding to the survey via the use of their ID number and password. Anyhow, the background (i.e. socio-demographic characteristics) data of the survey population must be fully known to the researchers, which allows performing a careful drop-out analysis for evaluating the representativeness of the collected data. By combining various data collection methods one can improve the representativeness of the collected data and save on material and data entry costs. At the same time the merging of data collection methods requires a larger work contribution which may even raise the cost of the survey. The selection of the particular data collection method and techniques for motivating respondents must be done individually in each case, considering the different target populations and the availability of respondents. If the surveys
are focused repeatedly at the same study population, it would be advisable to explore in this connection also which particular method yields higher response rates in different study groups. The collection of data for the Physician 2011 study proved the need for such testing. It is the quality of collected information that should have the first priority in choosing the data collection method. The method should not be selected merely on the basis of cost and benefit factors.

5. References
Estonian Science Foundation, University of Tartu. Noor arst’95 [Young Physician 1995].Tartu: University of Tartu, Department of Public Health; 1995.


Kirjastus, OÜ Infotrükk.

To cite this article (According to APA)