SPECIAL ARTICLE

Methodological and strategic insights for online survey studies: an analysis based on the CHERRIES checklist

María-Carmen Torrejón-Guirado\textsuperscript{a,}\textsuperscript{*}, Isabel San Martín-Erice\textsuperscript{b}, Leticia San Martín-Rodríguez\textsuperscript{b}, Marta Lima-Serrano\textsuperscript{a}

\textsuperscript{a} Department of Nursing, School of Nursing, Physiotherapy, and Podiatry, University of Seville, Institute of Biomedicine of Seville (IBIS), Seville, Spain
\textsuperscript{b} Department of Health Sciences, Public University of Navarra, Pamplona, Spain

Received 2 November 2023; accepted 28 February 2024

Abstract The use of online surveys has become a valuable and widely employed tool in health research. However, the use of such instruments necessitates methodological rigor and optimization in their design to achieve the best response rates. Drawing upon relevant literature and the international CHERRIES guidelines for the development of online surveys, this article addresses methodological aspects related to ethical considerations and data protection (with reference to the Association of Internet Research’s online ethics guide), study design and validation, recruitment, data collection processes, and data management and analysis. In conclusion, given the context of overexposure to online surveys, which can influence recruitment and response rates, strategies for their maximization are provided, encompassing both static and dynamic aspects of survey design.

© 2024 The Authors. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Palabras clave

Enuesta online; Metodología; Guía CHERRIES; Reclutamiento

Claves metodológicas y estratégicas para estudios basados en encuestas online: un análisis basado en la iniciativa Checklist for Reporting Results of Internet E-Surveys

Resumen El uso de encuestas online se ha convertido en una herramienta útil y muy utilizada en la investigación sanitaria. Sin embargo, el uso de este tipo de instrumentos necesita de una rigurosidad metodológica y de una optimización en su diseño para obtener las mejores tasas...
Introduction

A large proportion of nursing research questions are resolved through the development of observational studies by means of surveys. Since the use of the internet burst onto the scene years ago and, especially, since the COVID-19 pandemic, the use of surveys online (through emails, platforms, and even social networks) has become common among researchers. This could be due to the relative ease with which data is collected online, compared to traditional face-to-face interviews, positively influencing their response rate, the affordability of large sample sizes, as well as reaching otherwise inaccessible populations. Thus, this type of data retrieval in research has been considered cost-effective for recruitment, data collection and analysis.

Despite the potential advantages of using online surveys, there is controversy about the validity of their findings, especially if they are not conducted according to set standards. Although these standards are shared with other types of designs, in online surveys it is essential to emphasize representativeness, with the most frequent biases being the use of convenience sampling, and information bias.

In addition, among other general aspects such as data protection and ethical considerations, the specific characteristics of the online survey, the prevention of multiple responses by participants or statistical correctness must also be taken into account. In particular, the privacy of participants is often the most questioned topic in online surveys. Thus, for example, although social networks (SNs), such as Twitter, Instagram or WhatsApp, have become a fairly rapid means of disseminating online surveys, very often, researchers do not guarantee achieving reliability and security in accordance with current data protection regulations. Research data management plays an increasingly important role in the scientific process; This data should be managed carefully and stored in an ethically responsible manner.

Therefore, despite the potential usefulness of online surveys, the methodological aspects related to the use and development of online surveys should be discussed, in order to guarantee minimum quality standards. This paper reflects on this point, claiming the need to follow a series of guidelines for effective conducting and communication of online surveys. Finally, keys to maximising recruitment in this type of survey are provided.

Methodological aspects related to the use of online questionnaires and the CHERRIES guide

In a similar way as already existed with other types of design, for which the CONSORT statements (for clinical trials) or STROBE (for observational studies) were developed, given the specificity of studies based on online surveys, since 2004, and with correction of this in 2012, there is a checklist for reporting results from Internet e-surveys (CHERRIES). This guideline is aimed at maximising the validity of the survey and minimising bias in the findings. It can be found on the EQUATOR Network (www.equator-network.org), and continues to be recommended for scientific communication using this data collection technique.

The CHERRIES guide provides a structured framework for the evaluation and submission of studies using online surveys, with the aim of improving the transparency, quality and reproducibility of research in this field. It was adapted to Spanish in 2019 and consists of a series of items related to (A) ethical and legal aspects; (B) design, development and pretest; (C) recruitment process and description of the sample that has access to the survey; (D) running of the survey, the response rate; and (E) prevention of repeated entries by the same user and, finally, data analysis.

Next, we reflect on some of the issues covered in the CHERRIES standards, as well as other methodological aspects that must be considered in the design and reporting of studies that use online surveys.

Ethical and data protection aspects of online surveys

Both the CHERRIES standards and other studies have identified aspects related to the following needs: to obtain permission from a research ethics committee; to properly use informed consent (disclosure of information and fair treatment); and to adequately protect personal data (anonymity and confidentiality).

The existence of permission from a research ethics committee and informed consent are requirements for any research. In online surveys, it is recommended that an introductory home page be included, explaining the details of the study, such as the sponsor and objectives of the study;
the characteristics and duration of the surveys, as well as aspects related to the protection of participants’ data and respect for the ethical aspects of the research. The participant must be able to accept or not to be part of the study, by clicking on his/her consent.

According to Mondragón-Barrios, informed consent is “a process that consists of the express manifestation of a competent person (whose physical, mental and ethical capabilities enable him or her to make a decision) to participate in research, under conditions that enable him or her to understand the risks, benefits, consequences or problems that may arise during the whole process of the research which he/she is going to participate in.” This consent, therefore, must be preceded by the provision of information on the objective of the research, as well as its potential risks and benefits, in a relationship between the researcher and the person participating in the research, in which the latter is able to clarify, with sufficient or necessary time, the doubts that arise as to the process and thus make a free and uncoerced decision. This condition could be affected in the context of the online survey.

Regarding data protection, anonymity, confidentiality and security, it should be considered that when a user accesses a website, there is a cookie system that retains personal information. This should be disabled or the user should be encouraged to delete their data collection history. More and more institutions are requiring research projects to be subject to personal data protection protocols and regulated by data protection regulations, such as the European Personal Data Protection Act. Management protocols guarantee protection against theft, misuse, damage or loss, and data management in research projects undertaken through online surveys require the use of secure servers, data coding and encryption systems, which prevent the transfer of sensitive information through the use of secure servers, access protection and storage, only for the amount of time strictly necessary. In this regard, it should be considered that certain media companies, especially if they are free, are not completely safe for sharing personal and confidential information and are prone to hacking or espionage. Thus, data must be stored and processed on servers of official bodies such as those of universities or health centres, and data encryption must be used when working outside these networks. In addition, provided that no legal/contractual rule or guideline of the relevant entity prescribes a longer period, the stored research data should be retained for the shortest possible time.

It is particularly important to safeguard the protection of minors if they should not participate in the research. Confirmation that the participant is of legal age can be requested with one click. In cases where minors are involved in the research, informed consent must also be requested from the parents or legal guardians, as in any type of research. National or international data protection regulations must be complied with in the event that the research involves different countries, which can be frequent in this type of survey. Data protection is also particularly sensitive to other vulnerable groups, such as the elderly or people with disabilities, who, due to difficulties in accessing technology, may not understand or appreciate the importance of maintaining data privacy while sharing personal and confidential information online.

As in any survey, certain questions can have a negative impact on participants, causing harm or psychological damage to them. However, in online surveys, the same support from the researcher cannot be provided as in the case of a face-to-face survey, unless some rule is established for this (telephone contact or video call during the survey). It should also be confirmed that virtual space is a safe and private space to ask certain questions.

Finally, the Association of Internet Research (AOR) has developed an ethics guide for online research, which can be considered when conducting this type of research. This paper presents a number of questions that researchers should consider when conducting internet-based research. In addition, on their website, you can find examples and other resources that can be used in research based on online surveys.

Aspects related to the design and validation of the online survey

The CHERRIES standards highlight the importance of describing how the survey was conducted, its usability and functionality. For their part, Crawford, McCabe & Pope allude to the importance of an online survey being friendly and functional during its completion. They identify a series of standards classified into five categories, as follows: general screen design, text, presentation of questions, format of participant input or response, and navigation through - and interaction with - the survey.

It is highly recommended that usability tests be run on the design and content of the survey with small groups of participants and other stakeholders, before conducting the survey on a large scale. All aspects related to the survey, the interface; the interaction between the software and the user; the navigation through the questionnaire; the way in which the questions are arranged; the answer alternatives; and the operation of algorithms for key questions, etc., must all be evaluated.

As with face-to-face surveys and tests, it is important to confirm the validity and reliability of the survey beforehand, through a pilot study. In this regard, it must be borne in mind that a survey or test initially produced to be run in a face-to-face format may not be as valid and functional when run online. Another important aspect is cultural adaptation, especially when the survey is conducted in different countries.

Methodological aspects related to recruitment

With regard to recruitment, there are certain aspects that must be taken into account since, if this is not adequate, it may prejudice ethical aspects as regards the representativeness and extrapolation of the results or response rate. The CHERRIES standards refer to the importance of communicating whether or not it is an open survey, i.e. available to any visitor to a certain website, or closed: only available to participants that the researcher decides on (for example, using a password to access it). Another relevant aspect is the method of contact with the participants: it will be necessary to take into account whether this contact has been made exclusively through the internet, as well as the public-
Conducting simulations, Williams notes that some vulnerable populations, such as certain ethnic groups or people with lower socio-economic status, may be difficult to reach because of their social isolation or discrimination. Stratifying sample selection to represent the variability of the study population may be a useful strategy to address these biases.

It is also necessary to take into account the difficulties that certain populations may have in accessing the electronic resources used for data collection or other aspects related to the digital divide. All of this could influence the existence of selection biases. This is especially relevant when recruitment, and also data collection, is undertaken through specific platforms or social networks, since these can enhance the bias of self-selection, overrepresentation and/or exclusion of certain population groups. This aspect has been highlighted by previous authors as an ethical challenge in this type of research, especially when participation is remunerated. Advantages and disadvantages of using these types of tools have been discussed by Newman et al.

On the other hand, Williams also reflects on the bias that could occur in longitudinal studies, since people change their e-mail more frequently than they change their home address, for example, so different methods of access to participants must be ensured.

Another issue that Im and Chee refer to in their study is the time when the survey is launched, as this can influence recruitment, response rate, and representativeness. On this point, it is necessary to take into account aspects such as the time of day the survey is launched, the season of the year, or the time given for its completion.

In summary, researchers should consider how the population will be recruited before starting the study, to avoid selection biases, as well as providing information related to the existence of such biases. Previous authors also allude to the use of large samples to mitigate possible selection biases. It should also be indicated whether or not any method of randomisation of the sample has been used. The previously randomised population may be invited to take the survey. It should be noted that the response rate in cases of randomisation is usually decreased, compared to samples obtained for convenience.

In terms of extrapolating the results, as we said, online surveys can be run on a large number of people very easily, through media such as social networks. This fact could bias the results of the studies since, if we analyse this from a statistical perspective, we come up against the following controversy: on the one hand, the traditional trend encourages obtaining a large sample size in order to achieve greater predictive power. But, on the other hand, by obtaining this large sample size, the results could be statistically significant in groups where, in reality, the differences between them were insignificant, thus indicating a lack of representativeness of the population.

In addition, we must remember that there is a population group to which, by its very nature, it is difficult to extrapolate the results of online surveys: the population with low socio-economic status. This population group may have difficulties accessing the internet or possessing the necessary resources to complete the online survey, so the results of the studies could not be extrapolated to them.

Depending on the means we use to share the survey with participants, we must take into account who the link can reach. It will not be the same to contact the participants through a personal email, where we are ensuring the arrival of the survey to that exact person, as to publish the survey on a social network and lose control of that link (pre-selection bias of the sample). In the latter case, it would be important to pause, in order to design the study correctly, adding, for example, questions that enable us to select the sample so that the survey software has timestamps for the control of fraudulent survey responses.

Other aspects to highlight would be the level of commitment of the participants when completing the survey, the comprehension of the questions, and the social desirability bias (i.e., when the participant responds according to what society expects or wants them to answer). It is also important to select survey software that avoids multiple entries by the same participant. Conducting validation studies to establish the psychometric properties of online surveys could be a fairly effective solution to apply.

Finally, since one of the problems with online surveys is the low response rate, in this article we will dedicate a specific section to addressing how to maximise recruitment in online surveys later on.

Aspects related to the running and data collection process

Regarding data collection, the question arises: is the person who completed the survey a real person or the person who initiated it? Is the context in which the survey is conducted suitable for completing it?

It is the running of online surveys that the CHERRIES standards place the greatest emphasis on. Thus, the tool used must be taken into account, whether it is a web platform for collecting surveys or this is done by email (and in the latter case, how the data was collected); secondly, the web/social media context in which the survey was published or run and how this could influence the results; thirdly, the voluntary/obligatory nature of the survey and whether incentives were offered; fourthly the time at which the survey was run and, fifthly, mechanisms, if they are any, for randomising questions or adapting the questionnaire to the participants’ answers. In addition, it is important to report on the number of items per page; the number of screens; whether or not there is a possibility for the participants to review the answers; or mechanisms for checking complete tests. All these aspects influence the validity and quality of the data collection process.

Other issues that influence the response rate are the medium used for its publication, which must be attractive enough, and the usability of the platform. As well as other technical issues such as the capacity and speed of the server or network used, or the existence of antivirus software that
blocks access to the website or restrictions on the internet service provider. When surveys are sent via email, you should also consider the possibility that they may arrive as spam in the participant’s email. The trust in the site by the user, as well as in the use of the internet is also an aspect to be highlighted.\textsuperscript{13} Strategies that can improve the response rate can act as support or follow-up during the running of the survey, plus prior training on the use of the survey or the use of incentives (the latter also considering the risk this may pose for the representativeness of the sample).\textsuperscript{27}

In addition, it is necessary to take into account issues that may affect the validity of the data, such as the participant creating a false identity, paying little attention when completing the questionnaire, false answers or incorrectly completed questionnaires. These issues also affect face-to-face surveys but have been more closely linked to online surveys.\textsuperscript{28}

A strategy to improve representativeness and reduce self-selection bias is to publish the survey in different media, such as websites, social networks, etc., to improve attraction to the study, always maintaining control mechanisms over repeated entries to the questionnaire by the same user.\textsuperscript{5,28}

**Aspects related to data management and analysis**

The CHERIES standards emphasize the analysis of the response rate, the prevention of repeated user inputs, and the performance of proper data debugging and analysis.\textsuperscript{1,3,4}

One of the most common aspects of online surveys is the response rate, especially when compared to the number of visits.\textsuperscript{25} It is therefore important to analyse and report the response rate, which includes different aspects such as identifying the number of unique users/visitors based on aspects such as IP addresses, cookies or both. Another issue is the viewing rate (number of views of the first page of the survey versus the number of single users). The participation or recruitment rate refers to the ratio of single users who agree to participate to the number of visitors to the first page of the survey. Finally, the completion rate refers to the ratio of the number of users who complete the survey to the number of users who agreed to participate.\textsuperscript{1,23}

For the prevention of repeated entries, techniques such as the use of cookies, IP checking, analysis of the log file, or user registration are recommended. The organisation of the database is essential. Regarding the analysis, it is important to refer to how incomplete questionnaires are handled - questionnaires with atypical timestamps: for example, excessively short completion times, and whether statistical corrections are made to adjust the representativeness of the sample, such as the use of propensity scoring.\textsuperscript{1,14}

**How to maximize recruitment in the use of online questionnaires**

As already mentioned, in recent years we have seen a great proliferation of online surveys, which can exhaust potential respondents. How can the participation rate be increased in this context? This is the question that many researchers in all areas face and which we will try to answer by providing the existing evidence in this regard.

**Invitation to take part**

Studies have shown that the engagement rate improves when the user is invited to participate via a text message on the mobile phone.\textsuperscript{24} This form of invitation has been shown to be better than e-mail, especially in cases where the user will use their phone to respond to the survey.

In addition, the review undertaken by Sammut et al.\textsuperscript{24} argues that when you invite by e-mail, you get a higher response rate when you leave the subject empty than when you use a subject related to the topic. In the case of using a subject line, studies have shown that a subject that conveys that the person has been chosen to respond to that survey is better than any other type of issue.\textsuperscript{24,29}

Personalising messages with the person’s name and/or surname has also been shown to maximize the response rate, as opposed to generic messages.\textsuperscript{24} In this sense, paradoxically, a study showed that lower response rates are obtained when the invitation to participate in the survey is personalised in a more formal way, with their name, surname, academic title, etc.\textsuperscript{30} These findings suggest that a very high degree of personalisation may be perceived as worse by the potential participant, perhaps, as the authors point out, due to confidentiality concerns. The use of a lot of personal data can make it easier to identify the respondent in the offline world.

**Reminders**

Sending reminders also increases the response rate (compared to not using them). Several studies have tried to clarify what is the best way to do this, showing that it is better to give 2 reminders (the first one increases the rate more than the second one), through e-mail or text message on the phone, changing the text from the first to the second reminder (30% increase in the response rate when this is done) and without having an impact on the time that elapses between both messages.\textsuperscript{24}

The use of multiple reminders can be beneficial in terms of monitoring and tracking the results, although the balance between the marginal increase in the response rate and the overload of messages on the recipient must be assessed, so it is advisable to include a stop receiving reminders link.\textsuperscript{30}

**Days of the Week for sending out the survey**

The study by Sauermann and Roach (2013) found no significant differences in response rates between days of the week, although Wednesdays and weekends appeared to be slightly worse than the other days (especially if respondents had children). If the invitation arrived over the weekend, respondents were less inclined to respond to the spot, deferring the decision to a weekday. There were no significant differences in response rates depending on the time of day. The average response delay (time elapsed between invitation and response) for emails sent at night was about 12h,
compared to 3–4 h for emails sent at other times of the day.10

Survey design

Several years ago, De Bruijne and Wijnant studied the best way to present a survey to mobile phone users, which can now be considered commonplace. The evidence from their study pointed to the importance of the questions being presented in a row, one after the other and using scrolling to go up and down from the beginning to the end of the survey, rather than dividing the survey into different pages which the user has to progress through.31

In addition, the data also show that the vertical orientation of the answer scale for each question is slightly better than the horizontal orientation. Along these lines, the study by Weigold et al. showed that even better than the horizontal orientation is the drop-down menu, that produced a higher response rate,12 although it increased the response time.

Incentives

The response rate can be increased by using incentive strategies or motivational text messages. Specifically, in the study by Pedersen and Nielsen it was the combination of lottery draws of small amounts of money, along with messages that indicated the participant as chosen to respond to the survey, that was the strategy shown to maximize the response rate ahead of other similar strategies.29

Seuermann & Roach found a positive effect in the use of incentives such as post-survey sweepstakes, which are easier to implement and more cost-effective than previous incentives, especially in large samples. For a limited budget, they found a small number of large prizes more effective than a large number of small prizes.30

Conclusions

The use of online surveys has become a useful and widely used tool in recent years in healthcare research, partly due to accessibility to large population groups and because the time taken to run and complete them has decreased. However, this type of instrument requires methodological rigour for the results to be valid and reliable, as well as optimisation in its design to obtain the best response rates. Aspects related to methodological design such as ethical and data protection considerations; those related to the design and validation of the online survey; the recruitment of participants; the management and analysis of data; or the running and data collection process are protocolised and well defined thanks to verification guides such as the CHERRIES standards.

Likewise, based on the literature, strategies are presented to maximize recruitment and increase the response rate, in a context of overexposure to online surveys, it being key to produce an effective design to obtain response rates that provide valid and reliable data. Input is provided on the static and dynamic design of the survey. Regarding dynamic design, it is recommended that a maximum of 2 reminders be sent at specific intervals; that sending should occur during the week and preferably during daylight hours; and that the messages are personalised and with small changes in the wording. Regarding the more static aspects of the survey design, it is evident that it is important that the questions are presented in a row, one after the other, using scrolling to go up and down from the beginning to the end of the survey, through vertical orientation of the response scale for each question. The use of incentives has also been shown to be effective in increasing the response rate: these can be motivating, economic or of a mixed nature.

Funding

No funding has been received.

Conflict of interest

There are no conflicts of interest.

References

1. López-Rodríguez JA. Declaración de la iniciativa CHERRIES: adaptación al castellano de directrices para la comunicación de resultados de cuestionarios y encuestas online [Improving the quality of Spanish web surveys: Spanish adaptation of the checklist for reporting results of internet e-surveys (CHERRIES) to the Spanish context]. Aten Primaria. 2019;51(9):586–9, http://dx.doi.org/10.1016/j.aprim.2019.03.005.


17. AoIR ethics working committeeEss C, Available online: Ethical decision-making and Internet research: Recommendations from the aoir ethics working committee; 2002 www.aoir.org/reports/ethics.pdf.


