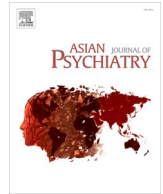




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Short communication

A critical look at online survey or questionnaire-based research studies during COVID-19

Swarndeeep Singh, Rajesh Sagar*

Department of Psychiatry, All India Institute of Medical Sciences, Delhi, India

ARTICLE INFO

Keywords:

Internet survey
 Limitations
 Survey fraud
 Methodology

ABSTRACT

In view of restrictions imposed to control COVID-19 pandemic, there has been a surge in online survey-based studies because of its ability to collect data with greater ease and faster speed compared to traditional methods. However, there are important concerns about the validity and generalizability of findings obtained using the online survey methodology. Further, there are data privacy concerns and ethical issues unique to these studies due to the electronic and online nature of survey data. Here, we describe some of the important issues associated with poor scientific quality of online survey findings, and provide suggestions to address them in future studies going ahead.

1. Introduction

Online survey or questionnaire-based studies collect information from participants responding to the study link using internet-based communication technology (e.g. E-mail, online survey platform). There has been a growing interest among researchers for using internet-based data collection methods during the COVID-19 pandemic, also reflected in the rising number studies employing online survey to collect data since the beginning of COVID-19 pandemic (Akintunde et al., 2021). This could be due to the relative ease of online data collection over traditional face-to-face interviews while following the travel restrictions and distancing guidelines for controlling the spread of COVID-19 pandemic. Further, it offers a cost-effective and faster way of data collection (with no interviewer requirement and automatic data entry) as compared to other means of remote data collection (e.g. telephonic interview) (Hlatshwako et al., 2021), both of which are important for getting rapid results to guide development and implementation public-health interventions for preventing and/or mitigating the harms related to COVID-19 pandemic (e.g. mental health effects of COVID-19, misconceptions related to spread of COVID-19, factors affecting vaccine hesitancy etc.). However, there have been several concerns raised about the validity and generalizability of findings obtained from online survey studies (Andrade et al., 2020; Sagar et al., 2020). Here, we describe some of the important issues associated with scientific quality of online survey findings, and provide suggestions to address them in future studies going ahead. The data privacy concerns and ethical issues unique to these

studies due to the electronic and online nature survey data have also briefly discussed.

2. Limited generalizability of online survey sample to the target general population

The findings obtained from online surveys need to be generalized to the target population in the real world. For this, the online survey population needs to be clearly defined and should be representative of the target population as much as possible. This would be possible when there is reliable sampling frame for online surveys, and participants could be selected using randomized or probability sampling method. However, online surveys are often conducted via email or online survey platform, with survey link shared on social media platforms or websites or directory of email ids accessed by researchers. Also, participants might be asked to share the survey link further with their eligible contacts. In turn, the population from which the study sample is selected often not clearly defined, and information about response rates (i.e. out of the total number people who viewed the survey link, how many of them did actually respond) are seldom available with the researcher. This makes generalization of study findings unreliable.

This problem may be addressed by sending survey link individually to all the people comprising the study population via email and/ or telephonic message (e.g. all the members of a professional society through membership directory, people residing in a society through official records etc.), with a request not to share the survey link with

* Corresponding author.

E-mail address: rsagar29@gmail.com (R. Sagar).

anyone else. Alternatively, required number of people could be randomly selected from the entire list of potential subjects and approached telephonically for taking consent. Basic socio-demographic details could be obtained from those who refused to participate and share the survey link with those agreeing to participate. Although, if the response rates are low or the socio-demographic details of non-responders significantly differ from that of responders, then the online survey sample is unlikely to be representative of the target study population. Further, this is a more resource intensive strategy and might not be always feasible (as it requires a list of contact details for the entire study population prior to beginning of data collection). In certain situations, when the area of research is relatively new and/or needs urgent exploration for hypothesis generation or guiding immediate response; the online survey study should list all possible attempts made to achieve a representative sample and clearly acknowledge it as a limitation while discussing their study findings (Zhou et al., 2021).

A more recent innovative solution to this problem involves partnership between academic institutions (Maryland University and Carnegie Mellon University) and the Facebook company for conducting online COVID-19 related research (Barkay et al., 2020). The COVID-19 Symptom Survey (CSS) conducted (in more than 200 countries since April 2020) using this approach involves exchange of information between the researchers and the Facebook without compromising the data privacy of information collected from survey participants. The survey link is shared on the Facebook, and user voluntary choose to participate in the study. The Facebook's active user base is leveraged to provide a reliable sampling frame for the CSS survey. The researchers select random ID numbers for the users who completed the survey, and calculate survey weights for each them on a given day. Survey weights adjust for both non-response errors (helps in making them sample more representative of the Facebook users) and coverage related errors (helps in making generalizing findings obtained using FAUB to the general population) (Barkay et al., 2020). A respondent belonging to a demographic group with a high likelihood of responding to the survey might get a weight of 10, whereas another respondent belonging to a demographic group with less likelihood of responding to survey might get a weight of 50. It also accounts for the proportion or density of Facebook or internet users in a given geographical area. Thus, findings obtained using this approach could be used for drawing inferences about the target general population. The survey weights to be used for weighted analysis of global CSS survey findings for different geographical regions are available to researchers upon request from either of the two above-mentioned academic institutions. For example, spatio-temporal trends in COVID-19 vaccine related hesitancy across different states of India was estimated by a group of Indian researchers using this approach (Chowdhury et al., 2021).

3. Survey fraud and participant disinterest

Survey fraud is when a person takes the online survey more than once with or without any malicious intent (e.g. monetary compensation, helping researchers collect the requisite number of responses). Another related problem is when the participant responds to some or all the survey questions in a casual manner without actually making any attempt at reading and/or understanding them due to reasons like participant disinterest or survey fatigue. This affects the representativeness and validity of online survey findings, and is increasingly being recognized as an important challenge for researchers (Chandler et al., 2020). While providing monetary incentives improves low response rates, it also increases the risk of survey fraud. Similarly, having a shorter survey length with few simple questions decreases the chances of survey fatigue, but limits the ability of researchers to obtain meaningful information about relatively complex issues. A researcher can take different approaches to address these concerns, ranging from relatively simpler ones such as requesting people to not participate more than once, providing different kind of monetary incentives (e.g. donation to a

charity instead of the participant), or manually checking survey responses for inconsistent (e.g. age and date of birth responses not consistent) or implausible response patterns (e.g. average daily smart-phone use of greater than 24 h, "all or none" response pattern) to more complex ones involving use of computer software or online survey platform features to block multiple entries by same person using IP address and/or internet cookies check, analysis of response time, latency or total time taken to complete survey for detecting fraudulent responses. There have been several different ways described in the available literature to detect fraudulent or inattentive survey responses, with a discussion about merits and demerits of each of them (Teitcher et al., 2015). However, no single method is completely fool proof, and it is recommended to use a combination of different methods to ensure adequate data quality in online surveys.

4. Possible bias introduced in results by the online survey administration mode

One of the contributory reasons for surge in online survey studies assessing mental health related aspects during the COVID-19 pandemic stems from the general thought that psychiatry research could be easily accomplished through scales or questionnaires administered through online survey methods, especially with the reliance on physical examination and other investigation findings being much less or non-existent. However, the reliability and validity of the scales or instruments used in online surveys have been traditionally established in studies administering them in face-to-face settings (often in pen/pencil-paper format) rather than online mode. There could be variation introduced in the results with different survey administration modes, which is often described as the measurement effect (Jäckle et al., 2010). This could be due to differences in the participants' level of engagement, understanding of questions, social desirability bias experienced across different survey administration methods. Few studies using the same study sample or sample sampling frame have compared the results obtained with difference in survey administration mode (ie. traditional face-to-face [paper format] vs. online survey), with mixed findings suggesting large significant differences to small significant difference or no significant differences (Determann et al., 2017; Norman et al., 2010; Saloniki et al., 2019). This suggests the need for conducting further studies before arriving at a final conclusion. Hence, we need to be careful while interpreting the results of online survey studies. Ideally, online survey findings should be compared with those obtained using traditional survey administration mode, and validation studies should be conducted to establish the psychometric properties of these scales for online survey mode.

5. Inadequately described online survey methodology

A recent systematic review assessing the quality of 80 online survey based published studies assessing the mental health impact of COVID-19 pandemic, reported that a large majority of them did not adhere to the CHERRIES (Checklist for Reporting Results of Internet E-Surveys) guideline aimed at improving the quality of online surveys (Eysenbach, 2004; Sharma et al., 2021). Information related to parameters such as view rate (Ratio of unique survey visitors/unique site visitors), participation rate (Ratio of unique visitors who agreed to participate/unique first survey page visitors), and completion rate (Ratio of users who finished the survey/users who agreed to participate); which gives an idea about the representativeness of the online study sample as described previously were not mentioned in about two-third studies. Similarly, information about steps taken to prevent multiple entries by same participant or analysis of atypical timestamps to check for fraudulent and inattentive survey responses was provided by less than 5% studies. Thus, it is imperative to popularize and emphasize upon the use of these reporting guidelines for online survey studies to improve the scientific value of findings obtained from internet-based studies.

6. Data privacy and ethics of online survey studies

Lastly, most of the online survey studies either did not mention at all or mentioned in passing about maintain the anonymity and confidentiality of information obtained from online survey. However, details about the various steps or precautions taken by the researchers to ensure data safety and privacy were seldom mentioned (e.g. de-identified data, encryption process or password protected data storage, use of HIPAA-compliant online survey form/platform etc.). The details and limitations of safety steps taken, and the possibility of data leak should be clearly mentioned/ communicated to participants at the time of taking informed consent (rather than simply mentioning anonymity and confidentiality of information obtained will be ensured, as is the case with offline studies). Moreover, obtaining ethical approval prior to conducting online survey studies is a must. The various ethical concerns unique to online survey methodology (e.g. issues with data protection, informed consent process, survey fraud, online survey administration etc.) should be adequately described in the protocol and deliberated upon by the review boards (Buchanan and Hvizdak, 2009; Gupta, 2017).

In conclusion, there is an urgent need to consider the above described issues while planning and conducting an online survey, and also reviewing the findings obtained from these studies to improve the overall quality and utility of internet-based research during COVID-19 and post-COVID era.

Financial disclosure

The authors did not receive any funding for this work.

Acknowledgments

None.

Conflict of interest

The authors have no conflict of interest to declare.

References

- Akintunde, T.Y., Musa, T.H., Musa, H.H., Musa, I.H., Chen, S., Ibrahim, E., Tassang, A.E., Helmy, M., 2021. Bibliometric analysis of global scientific literature on effects of COVID-19 pandemic on mental health. *Asian J. Psychiatry* 63, 102753. <https://doi.org/10.1016/j.ajp.2021.102753>.
- Andrade, C., 2020. The limitations of online surveys. *Indian J. Psychol. Med.* 42 (6), 575–576. <https://doi.org/10.1177/0253717620957496>.
- Barkay, N., Cobb, C., Eilat, R., Galili, T., Haimovich, D., LaRocca, S., ..., Sarig, T., 2020. Weights and methodology brief for the COVID-19 symptom survey by University of Maryland and Carnegie Mellon University, in partnership with Facebook, arXiv preprint arXiv:2009.14675.
- Buchanan, E.A., Hvizdak, E.E., 2009. Online survey tools: ethical and methodological concerns of human research ethics committees. *J. Empir. Res. Hum. Res. Ethics.* 4 (2), 37–48. <https://doi.org/10.1525/jer.2009.4.2.37>.
- Chandler, J., Sisso, I., Shapiro, D., 2020. Participant carelessness and fraud: consequences for clinical research and potential solutions. *J. Abnorm. Psychol.* 129 (1), 49–55. <https://doi.org/10.1037/abn0000479>.
- Chowdhury, S.R., Motheram, A., Pramanik, S., 2021. Covid-19 vaccine hesitancy: trends across states, over time. *Ideas For India*, 14 April, Available from: (<https://www.idea.sforindia.in/topics/governance/covid-19-vaccine-hesitancy-trends-across-states-over-time.html%20%20>) (Accessed 4 August 2021).
- Determann, D., Lambooi, M.S., Steyerberg, E.W., de Bekker-Grob, E.W., de Wit, G.A., 2017. Impact of survey administration mode on the results of a health-related discrete choice experiment: online and paper comparison. *Value Health.: J. Int. Soc. Pharm. Outcomes Res.* 20 (7), 953–960. <https://doi.org/10.1016/j.jval.2017.02.007>.
- Eysenbach, G., 2004. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). *J. Med. Internet Res.* 6 (3), 34. <https://doi.org/10.2196/jmir.6.3.e34>.
- Gupta, S., 2017. Ethical issues in designing internet-based research: recommendations for good practice. *J. Res. Pract.* 13 (2). Article D1.
- Hlatshwako, T.G., Shah, S.J., Kosana, P., Adebayo, E., Hendriks, J., Larsson, E.C., Hensel, D.J., Erausquin, J.T., Marks, M., Michlielsen, K., Saltis, H., Francis, J.M., Wouters, E., Tucker, J.D., 2021. Online health survey research during COVID-19. *Lancet Digit. Health* 3 (2), e76–e77. [https://doi.org/10.1016/S2589-7500\(21\)00002-9](https://doi.org/10.1016/S2589-7500(21)00002-9).
- Jäckle, A., Roberts, C., Lynn, P., 2010. Assessing the effect of data collection mode on measurement. *Int. Stat. Rev.* 78 (1), 3–20. <https://doi.org/10.1111/j.1751-5823.2010.00102.x>.
- Norman, R., King, M.T., Clarke, D., Viney, R., Cronin, P., Street, D., 2010. Does mode of administration matter? Comparison of online and face-to-face administration of a time trade-off task. *Qual. Life Res.: Int. J. Qual. Life Asp. Treat., Care Rehabil.* 19 (4), 499–508. <https://doi.org/10.1007/s11136-010-9609-5>.
- Sagar, R., Chawla, N., Sen, M.S., 2020. Is it correct to estimate mental disorder through online surveys during COVID-19 pandemic? *Psychiatry Res.* 291, 113251 <https://doi.org/10.1016/j.psychres.2020.113251>.
- Saloniki, E.C., Malley, J., Burge, P., Lu, H., Batchelder, L., Linnosmaa, I., Trukeschitz, B., Forder, J., 2019. Comparing internet and face-to-face surveys as methods for eliciting preferences for social care-related quality of life: evidence from England using the ASCOT service user measure. *Qual. Life Res.: Int. J. Qual. Life Asp. Treat., Care Rehabil.* 28 (8), 2207–2220. <https://doi.org/10.1007/s11136-019-02172-2>.
- Sharma, R., Tikka, S.K., Bhute, A.R., Bastia, B.K., 2021. Adherence of online surveys on mental health during the early part of the COVID-19 outbreak to standard reporting guidelines: a systematic review. *Asian J. Psychiatry* 65, 102799. <https://doi.org/10.1016/j.ajp.2021.102799> (Advance online publication).
- Teitcher, J.E., Bockting, W.O., Bauermeister, J.A., Hoefler, C.J., Miner, M.H., Klitzman, R. L., 2015. Detecting, preventing, and responding to “fraudsters” in internet research: ethics and tradeoffs. *J. Law, Med. Ethics: J. Am. Soc. Law, Med. Ethics* 43 (1), 116–133. <https://doi.org/10.1111/jlme.12200>.
- Zhou, T., Chen, W., Liu, X., Wu, T., Wen, L., Yang, X., Hou, Z., Chen, B., Zhang, T., Zhang, C., Xie, C., Zhou, X., Wang, L., Hua, J., Tang, Q., Zhao, M., Hong, X., Liu, W., Du, C., Li, Y., Yu, X., 2021. Children of parents with mental illness in the COVID-19 pandemic: a cross-sectional survey in China. *Asian J. Psychiatry* 64, 102801. <https://doi.org/10.1016/j.ajp.2021.102801> (Advance online publication).