

# Assessing the Accuracy of Survey Research

Samuel J. Stratton, MD, MPH

Survey research is a frequent method for disaster medical research submitted to *Prehospital and Disaster Medicine*. Surveys utilize questions to gather information from study subjects. By using survey methodology, an investigator can make determinations about the beliefs, attitudes, experiences, and expectations of a study population. The introduction of electronic survey platforms, such as SurveyMonkey (Palo Alto, California USA), has resulted in an upsurge of survey studies and polls. Unlike a census in which all members of a population are studied, the survey uses a fraction or sample of a study population to make conclusions. For a survey to be accurate, the individuals selected to participate in the survey must represent the entire study population in all ways that are relevant to the study objective. Surveys can be conducted one time as a cross-sectional study or repeated over time to form a longitudinal study. In addition to online formats for conducting survey research, an investigator can use telephone, mail, or direct interviews to present survey questions to participants. Survey questionnaires can be designed to generate data that allow for statistical analysis and inference.

While survey research may be intrusive for individuals and can expose personal information, most human subjects or ethics committees consider this type of research exempt from need for detailed review as long as the investigator presents proof that confidential personal information will be protected. Additionally, an investigator must assure that survey subjects participate voluntarily and without coercion. Because appropriately conducted surveys are voluntary, there is risk of lack-of-response by some of those selected as survey participants. Lack-of-response to survey questions leads to one of the major problems in survey research, that of nonresponse error.

Survey research should be conducted using the same principles as other scientific methods with the strictest attention directed to limiting all possible bias. Bias, for purposes of this article, is error and interference within a study that leads to lack-of-validity of the survey outcome(s). For survey-based research, the risks for bias likely are higher than for other common research methods. Therefore, it is important that when considering survey research, the investigator design the research to minimize all forms of bias inherent with survey methodology.

The primary forms of bias found with survey research include:

- Sampling error;
- Statistical margin of error;
- Nonresponse error;
- Response bias; and
- Measurement variability.

While other types of bias may complicate survey research, the five challenges listed above are usually the most important to consider.

Sampling error occurs when the sample of participants selected to represent the study population actually is not representative of that population. This can occur when there is a failure to assure

that the study sample accurately reflects the demographics, culture, education level, experiences, or other significant aspects of the study population. In addition to failure to identify a representative sample of the population, a sample size that is small and does not include all potential representatives of a heterogeneous population in proper proportion will result in sampling error. Sampling error can result in bias or large variance in study results. Sampling error can be minimized best by using a randomized sample of the total population. In addition, the number of subjects included in the sample should be large enough to assure that the expected spread of confidence intervals or other measures of precision are clinically significant. There are a number of online calculators and charts available that allow an investigator to determine the appropriate sample size for various study methods, including survey research.<sup>1</sup>

The term margin of error, when applied to survey research, has a statistical meaning. Margin of error is a measure of the precision of an estimate of a value or data point. Margin of error is reported as a mathematic estimate of the range of values in which the central measure (mean or median) falls. The 95% confidence interval for continuous or categorical data and the 25%/75% quartile range for ordinal data are used most commonly as determinants of the mathematic margin of error for a survey result. It is more subjective and difficult to determine the precision of qualitative findings in surveys, but the degree of theme saturation in the data pool will allow an investigator to prioritize study outcome themes using qualitative techniques. In prioritizing themes, the data derived from a qualitative study can be presented as percentages of respondents who offered the theme with their responses to survey questions. For example, if a qualitative survey were conducted of 50 disaster victims and 90% reported fresh water as a problem and 30% reported food as a problem, the investigator would report both themes (water and food) but identify that fresh water was a more prominent and consistently powerful theme in the data. No matter what type of data are being collected in a survey, sample size determines the precision of findings with a larger or statistically appropriate sample size necessary for survey data to be meaningful.

The most difficult error for most survey research is nonresponse error. Nonresponse error occurs when individuals selected for the study sample are not interviewed or fail to answer survey questions. Nonresponse introduces systemic bias into survey data and almost always has a bigger impact on survey validity than the mathematic margin of error. Nonresponse rarely is random, and depending on the characteristics of nonresponders, this type of error can have significant effects within a data set. Nonresponders may have many different reasons for lack of response to a survey, and missing this data leads to risk of selection bias that can affect outcome results. There is no accepted measure of the effect of nonresponders on survey data, but one logical approach is to assume there is a random or 50% probability that a nonresponder would have answered a question one way or the other. Applying this assumption in the simplest way, assume a sample of 100 potential respondents was asked to answer a yes or no question.

After completion of the survey it was found that there was an 80% response rate, or 20% nonresponse rate. The assumption would then be that the error introduced into the data set due to nonresponse is  $\pm 10\%$  (Nonresponse Error = Nonresponse rate (%) / 2). There are a number of actions that can be taken to decrease nonresponse, including call backs or reminders, investigator conversations with nonresponders to encourage response, providing incentives, and oversampling. A common research technique used to decrease concern about nonresponders is to provide demographic data for the entire population and show that those responding to the survey closely match and mirror the study population.

More difficult to detect and adjust for in interpreting survey data is response bias. Simply defined, response bias is a term for influences and motivations of survey respondents to avoid providing truthful or accurate responses. Response bias is another form of systematic error that is inherent with survey research. Respondents may alter truthful or correct responses to survey questions because they want to please the investigator, they misinterpret questions, or they have concern that they may be identified as having a minority answer. One of the most prevalent and studied areas of response bias is social desirability bias in which a respondent's culture, gender, age, and information received from media and other sources leads to inaccurately answering survey questions. It has been reported that social desirability bias can affect 10% to 70% of data, depending on the topics and focus of a survey.<sup>2</sup> Validating a survey questionnaire prior to administration to the study sample may limit social desirability bias. Validation can be done by having experts or respondents that are not to be included in the study sample complete the survey questionnaire, which allows the investigator to

assess if answers are consistent and plausible. For interview surveys, interviewers must be culturally competent, speak language in a familiar form for the respondents, and be age and gender appropriate for the sample participants.

A final challenge for survey-based research is measurement variability. Measurement variability occurs when respondent answers do not match the data actually needed. This can occur with poorly developed questionnaires or interviews, respondent inability to answer a question, and respondent inability to recall accurate answers. Questionnaires and interviews used for survey research must be developed carefully and validated to avoid measurement variation. Lengthy questionnaires or interviews lead to decreasing motivation of respondents to answer later questions in a questionnaire or interview. Leading questions, confusing phrases, double negative questions, language variation from that of respondents, and implied bias in titles, questions, and conduct of a survey all can result in measurement variation and should be avoided.

For survey research to be considered true science, best method requires control of potential error and that outcome is measured with meaningful precision. Achieving this level of science with survey techniques is extremely difficult and requires attention to limiting all possible sources of bias that can be identified. With attention to the survey research challenges discussed in this article, it is possible to achieve a level of research that is valid and clinically useful. Disaster and emergency health research does not lend itself to control of all potential independent variables, which leads to great difficulty in conducting truly scientific research. On the other hand, survey research is a logical tool for conducting disaster and emergency health research, but it must be done with strict attention to limiting error and bias.

#### References

1. Bartlett JE, Kotrlík JW, Higgins CC. Organizational research: determining appropriate sample size in survey research. *Information Tech, Learning, and Performance Journal*. 2001;19(1):43-50. Available at: [opalco.com/wp-content/uploads/2014/10/Reading-Sample-Size.pdf](http://opalco.com/wp-content/uploads/2014/10/Reading-Sample-Size.pdf). Accessed April 14, 2015.
2. Nederhof A. Methods of coping with social desirability bias: a review. *European J of Soc Psych*. 1985;15(3):263-280. Available at: <http://onlinelibrary.wiley.com/doi/10.1002/ejsp.2420150303/pdf>. Accessed April 14, 2015.

doi:10.1017/S1049023X15004719